UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Pablo Schoeffel

A METHOD TO PREDICT AT-RISK STUDENTS IN INTRODUCTORY COMPUTING COURSES BASED ON MOTIVATION

Florianópolis
2019

Pablo Schoeffel

A METHOD TO PREDICT AT-RISK STUDENTS IN INTRODUCTORY COMPUTING COURSES BASED ON MOTIVATION

Tese submetida ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Santa Catarina para a obtenção do título de Doutor em Ciência da Computação
Orientador: Prof. Dr. Raul Sidnei Wazlawick
Coorientador: Prof. Dr. Vinicius Faria Culmant Ramos

Florianópolis
2019

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Schoeffel, Pablo
A Method to Predict At-risk Students in Introductory Computing Courses Based on Motivation / Pablo Schoeffel ; orientador, Raul Sidnei Wazlawick, coorientador, Vinicius Faria Culmant Ramos, 2019.
197 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Ciência da Computação, Florianópolis, 2019.

Inclui referências.

1. Ciência da Computação. 2. Educational Data Mining. 3. Computing Education. 4. Motivation. 5. Programming. I. Wazlawick, Raul Sidnei. II. Ramos, Vinicius Faria Culmant. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Computação. IV. Título.

Pablo Schoeffel

A METHOD TO PREDICT AT-RISK STUDENTS IN INTRODUCTORY COMPUTING COURSES BASED ON MOTIVATION

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Fabiane Barreto Vavassori Benitti, Dr.ª
Universidade Federal de Santa Catarina

Prof. Cristian Cechinel, Dr.
Universidade Federal de Santa Catarina

Prof. André Luis Alice Raabe, Dr.
Universidade do Vale do Itajaí

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Ciência da Computação.

Digitally signed by Jose Luis Almada Guntzel:53378067004
Date: 2019.10.03 13:29:59 -03'00'

Prof. José Luís Almada Güntzel, Dr.
Coordenador do Programa

Assinado de forma digital por RAUL SIDNEI WAZLAWICK:71172084904
Dados: 2019.10.01 16:33:18 -03'00'

Prof. Raul Sidnei Wazlawick, Dr.
Orientador

Florianópolis, 21 de agosto de 2019.

This work is dedicated to all my family.

ACKNOWLEDGEMENTS

Every process of improvement and maturation requires effort and, in general, brings with it some turbulence. In the case of the doctorate, the daily routine changes, we must abdicate of activities in family, we are absent during periods of class and events, among others. Therefore, for this process to flow without undermining relationships, people around me have to be tolerant and true companions. Fortunately, my family gave me and gives me all this support, tolerating my absence when necessary and giving me the affection and company whenever I needed. My special thanks go to my lovely wife Edna and my sweet children Caio and Sofia, and also to my parents, brother, and mother-in-law, who have always supported me and helped to compensate my absence whenever necessary.

I also thank my advisors in particular. To Prof. Raul Sidnei Wazlawick who gave me security and serenity throughout the process, allowing me to evolve in the conduction of the work, always giving me the necessary support. To Prof. Vinicius Faria Culmant Ramos, for the help and participation in the discussions and evaluations of the work and papers, becoming fundamental in the evolution and development of this thesis.

Thanks also to all the professors, staff and fellow members of the PPGCC, who shared their knowledge and allowed me to be fit for this journey.

I am immensely grateful to UDESC, especially to all my colleagues at UDESC Ibirama, who gave me the clearance period, allowing me to devote entirely to the doctorate.

I thank all the professors who participated in the experiments, allowing the implementation of the research: Prof. Adilson Vahldick (UDESC), Prof. Fernando dos Santos (UDESC), Prof. Paolo Moser (UDESC), Prof. Tiago Schmitz (UDESC), Prof. Cristiano Roberto Franco (FAMEG), Prof. Fernanda dos Santos Cunha (UNIVALI), Prof. Miriam Sejas (UFSC Araranguá), Prof. Esdras Bispo Jr. (UFG).

Thanks also to all those who, somehow, helped me in specific issues, clearing up doubts, discussing, sharing experiences and working together on writing articles: Prof. Adilson Vahldick, Prof. Marcelo de Souza, Prof. Cristiano Roberto Franco, Prof. Paolo Moser, Prof. Paulo Farah.

Anyway, I would like to thank all the friends who have helped reduce tension and anxiety of this period, especially in the final moments. And to all those who in some way have contributed directly or indirectly to the development of this work.

Successful and unsuccessful people do not vary greatly in their abilities. They vary in their desires to reach their potential.

(John C. Maxwell)

RESUMO

Apesar de ser um problema relatado há muito tempo, o alto índice de evasão e reprovações em cursos de computação continua sendo um problema para a área. Nesse contexto, as disciplinas introdutórias de programação estão entre as que mais reprovam, gerando os maiores índices de evasão nos primeiros semestres do curso. Diversos estudos descrevem estratégias para tentar melhorar o ensino de programação e reter mais alunos, assim como outros estudos tentam identificar fatores que estão relacionados com o sucesso ou retenção dos estudantes. Uma das estratégias para intervir no processo de ensino-aprendizagem é identificar previamente os estudantes em risco. Apesar de existir uma forte relação entre a motivação e o resultado dos estudantes, poucos trabalhos utilizam a motivação como um fator para identificar estudantes em risco. Esse trabalho apresenta e avalia um método para identificar características que permitam prever com antecedência estudantes em risco em disciplinas introdutórias de Computação, baseado em quatro componentes principais: fatores pré-universitários, motivação inicial, motivação ao longo da disciplina e percepção do professor. Além disso, é proposto um instrumento para avaliar fatores educacionais que impactam na motivação. O método é baseado em questionários, cujos resultados de validação da confiabilidade e validade, por meio do coeficiente alpha de Cronbach, coeficiente ômega e análise fatorial, mostraram-se satisfatórios. A partir do método criado, nomeado de EMMECS, foram realizados estudos de caso com 173 estudantes de diferentes cursos na área de computação em quatro diferentes universidades no sul do Brasil. Foram então realizadas diversas simulações de predição, utilizando dez diferentes algoritmos de classificação e diferentes configurações de *datasets*. Como resultado, os melhores cenários, utilizando os algoritmos *support vector machine e Adaboost M1*, identificou, em média, mais de 80% dos estudantes que seriam reprovados, desde a primeira semana de aula. Os resultados mostram então que o método proposto é eficaz comparado com trabalhos correlatos e tem como vantagens o fato de ser independente de conteúdo programático, de avaliações específicas, de interação com sistemas de aprendizagem e de permitir a identificação semanal, com bons resultados desde as primeiras semanas.

Palavras-chave: Educational Data Mining. Ensino de Computação. Motivação. Método. Predição. Introdução à Programação.

RESUMO EXPANDIDO

Introdução

Na contramão do aumento da demanda de profissionais de Tecnologia da Informação e Computação, o alto índice de reprovações e evasão continua sendo um problema para a área. No Brasil e no mundo, cursos de Computação estão entre aqueles com maior taxa de evasão, chegando a mais de 30% no mundo e mais de 40% no Brasil, podendo chegar até 70% em algumas instituições. Nesse contexto, o maior índice de insucesso e desistências estão no primeiro ano do curso, principalmente nas disciplinas introdutórias de programação. Estudos mostram índices de reprovação de 30% a 60% nessas disciplinas. Diversos estudos têm sido feitos a fim de criar novas estratégias pedagógicas e métodos de avaliação, entre outras coisas. Porém, outro aspecto importante é identificar potenciais estudantes em risco com antecedência, a fim de realizar intervenções ao longo da disciplina. Um aspecto comumente relacionado com o resultado do estudante é a *motivação*. Porém, verificamos que poucos estudos consideram a motivação como um aspecto a ser analisado em disciplinas introdutórias de Computação, principalmente em estudos de predição. Muitos estudos utilizam como dados de entrada as notas de avaliações ou interações com o sistema de aprendizagem. Porém, isso faz com que o método seja dependente da estratégia de ensino e da ferramenta, além de permitir, geralmente, que a predição seja realizada após algumas semanas decorridas de curso. Ainda, dos estudos existentes que tentam prever antecipadamente estudantes em risco, nenhum estuda a evolução da motivação ao longo do tempo na disciplina. Nesse sentido, esse trabalho visa confirmar se é possível criar um método que utiliza somente dados pré-universitários e de motivação para prever antecipadamente, já nas primeiras semanas, estudantes em risco em disciplinas introdutórias de Computação.

Objetivos

O objetivo geral desse trabalho é identificar o nível de motivação e fatores que afetam a motivação e engajamento dos estudantes de disciplinas introdutórias de Computação e, baseado nesses aspectos, prever com antecedência estudantes em risco de desistência ou reprovação. Para atingir esse objetivo principal, foram definidos aos seguintes objetivos específicos: i) OE1 – identificar fatores que afetam a motivação de estudantes de computação; ii) OE2 – medir a motivação dos estudantes no contexto de disciplinas introdutórias de computação; iii) OE3 – prever com antecedência estudantes em risco; iv)

OE4 – disponibilizar informações relevantes sobre os estudantes para permitir que professores e coordenadores tomem ações com antecedência.

Método de Pesquisa

Quanto ao método de raciocínio, essa pesquisa pode ser classificada tanto como raciocínio indutivo como raciocínio dedutivo. Segundo Trochim e Donnely (2008), raciocínio dedutivo trabalha do objetivo mais geral para o mais específico, enquanto o raciocínio indutivo trabalha de observações específicas, buscando generalizá-las. Inicialmente, esse trabalho usa raciocínio indutivo, pois busca achar padrões de fatores e motivação dos estudantes em observações iniciais, a fim de criar o método EMMECS. A partir daí, faz-se observações específicas para tentar generalizar os resultados. Com relação à estratégia empírica, esse trabalho é classificado como uma pesquisa exploratória, visto que baseia-se no estudo de objetos na sua configuração natural e deixa que os achados emerjam das observações (WOHLIN, RUNESON, *et al.*, 2012, p. 9). Quanto ao tipo de análise de dados, esse trabalho é uma pesquisa mista, que usa tanto dados quantitativos, como os *surveys* e informações acadêmicas, quanto qualitativos, como entrevista com o professor. O posicionamento filosófico é o *pragmatismo* que, segundo Easterbrook et al. (2008), significa que o conhecimento é julgado pelo quanto ele resolve problemas práticos. E foi essa a principal motivação dessa pesquisa. O método de pesquisa utilizado para realizar o trabalho seguiu seis fases principais. A primeira fase visa entender quais fatores relacionados à motivação podem afetar o resultado dos estudantes. Foi planejado identificar esses fatores por meio de um mapeamento sistemático da literatura, seguindo as diretrizes propostas por (PETERSEN, FELDT, *et al.*, 2008) e (KITCHENHAM, BUDGEN, D. and BRERETON, 2010). A segunda fase refere-se à definição do método de coleta dos dados e predição dos estudantes em risco. Essa fase inclui a definição do processo, criação dos instrumentos, além da escolha e configuração dos algoritmos de predição. A terceira fase visa aplicar o método, realizando estudos de caso com estudantes de cursos de graduação em Computação de quatro universidade brasileiras. Para a validação do método, foram priorizadas disciplinas dos primeiros semestres pois, de acordo com a literatura, são os períodos mais críticos relacionados ao insucesso e taxas de motivação dos estudantes. A quarta fase envolve a avaliação dos algoritmos de predição selecionados, a fim de verificar o desempenho e necessidades de ajuste. A quinta fase refere-se ao desenvolvimento de uma ferramenta que, a partir dos dados e resultados de predição obtidos pelo método, visa auxiliar professores e coordenadores na gestão dos estudantes em risco, e apoiar a identificação de fatores que ajudem em ações de intervenção. A sexta e última fase visa validar a ferramenta criada com professores num contexto real, a fim de avaliar questões de usabilidade e intenção de uso da ferramenta. Além disso, nessa fase é realizada entrevista com o professor da disciplina para entender sua experiência com o uso da ferramenta.

Resultados e Discussão

A principal contribuição desse trabalho é o desenvolvimento de um método que permite a identificação com antecedência de estudantes em risco em disciplinas introdutórias de Computação. O método nomeado EMMECS baseia-se em três instrumentos principais, no formato de questionários. Os três instrumentos foram validados quanto a sua confiabilidade e validade utilizando métodos estatísticos como coeficiente Alpha de Cronbach, coeficiente ômega, intercorrelação e análise fatorial. Todos os resultados foram satisfatórios. Cada um dos instrumentos permitiu encontrar fatores relacionados ao resultado e motivação dos estudantes. Em diversos estudos de caso realizados, encontramos evidências da relação do resultado dos estudantes com 5 fatores pré-universitários, 4 aspectos de motivação inicial, 15 fatores educacionais, percepção do professor e motivação ao longo da disciplina. Com relação à validação do método como um todo, os resultados da assertividade de identificação dos estudantes reprovados (recall) foi, no melhor cenário, em média, mais de 80%, chegando a 90%. Um resultado interessante é que o método teve desempenho similar desde as primeiras semanas. A aplicação do método num contexto real mostrou que o EMMECS é uma ferramenta que pode auxiliar significativamente os professores na condução de suas disciplinas e ajudar em intervenções com estudantes específicos.

Considerações Finais

O alto índice de insucesso em disciplinas introdutórias de programação continua sendo um problema e, apesar na relação já comprovada em diversos trabalhos, entre a motivação do estudante com seus resultados, poucos trabalhos analisam esse tema no contexto de cursos de Computação. Diante disso, esse trabalho desenvolveu um método, baseado em questionários, que visam identificar fatores importantes relacionados com a motivação do estudante e, consequentemente, seus resultados acadêmicos. O objetivo do método é utilizar esses atributos coletados para identificar antecipadamente estudantes em risco. O método EMMECS foi aplicado com 173 estudantes em nove turmas de quatro diferentes universidades no Brasil e seus resultados mostraram-se promissores, com taxas de identificação acima de 80%, em média, desde as primeiras semanas. Visto que o método é independente de conteúdo, avaliações ou ferramentas, outros pesquisadores e professores podem utilizar o EMMECS como instrumento para auxiliar na melhoria de disciplinas introdutórias de computação.

ABSTRACT

Despite being a problem reported in a long time, the high rate of dropout and failure in computing courses remains a problem for the area. In this context, the introductory programming courses are among the worst, generating the highest rates of dropout in the first semesters of the program. Several studies describe pedagogical strategies that try to improve teaching programming and retain more students. Other studies try to identify factors that are related to the success or retention of students. One of the strategies to support the teaching-learning process is to identify students at risk in advance. Although there is a strong relationship between the motivation and the students' outcome, few works use the motivation as a factor to identify students at risk. This work presents and evaluates a method to identify features that allow predicting at-risk students in introductory computing courses, based on four main components: pre-university factors, initial motivation, motivation through the course, and professor perception. In addition, it is proposed an instrument to assess educational factors that impact on motivation. The method is based on questionnaires, and the results were validated regarding reliability and validity by the Cronbach's alpha coefficient, omega coefficient, and factor analysis, which we proved to be satisfactory. Using the method created, named EMMECS, case studies with 173 students from different courses in computer science in four different universities in southern Brazil were conducted. We carried out several simulations of prediction, using ten different classification algorithms and different datasets. As a result, the best-case scenarios, using support vector machine and AdaBoostM1 algorithms, we identified on average more than 80% of students that would fail, since the first week of the study. The results show that the proposed method is effective compared with related works and it has as advantages its independence of programmatic content, specific assessments, grades, and interaction with learning systems. Furthermore, the method allows the weekly prediction, with good results since the first few weeks.

Keywords: Educational Data Mining. Computing Education. Motivation. Method. Prediction. Introductory Computing Courses. Programming.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ACM - Association for Computing Machinery
AI – Artificial Intelligence
AMO - Amotivation
AMS - Academic Motivation Scale
ANOVA - Analysis of Variance
API - Application Programming Interface
ARCS - Attention, Relevance, Confidence, and Satisfaction
ARFF - Attribute-Relation File Format
AUC - Area Under Curve
BRASSCOM – Associação Brasileira das Empresas de Tecnologia da Informação e Comunicação
CAE - Correlation Attribute Eval
CER - Computing Education Research
CGPA - Cumulative Grade Point Average
CI - Cost Index
CS - Computer Science
CS1 - Computer Science 1
CSV - Comma Separated Value
EDM - Educational Data Mining
EI - Expectancy Index
EME - l'Échelle scale of motivation in education
EMEr - Extrinsic motivation for external regulation
EMId - Extrinsic motivation for identification
EMIn - Extrinsic motivation by introjection
EMMECS - Evaluation Method of Motivation and Engagement in Computing Students
ENEM – Exame Nacional do Ensino Médio (National High School Exam)
EVC - Expectancy-value-cost
FA - Factor Analysis
GQM - Goal Question Metric
IDE - Integrated Development Environment
IEEE - Institute of Electrical and Electronics Engineers
IGA - Info Gain Attribute Eval
IMES - Intrinsic motivation to experience stimulation
IMK - Intrinsic motivation for knowledge

IMR - Intrinsic motivation for realization

IT - Information Technology

KMO - Kaiser-Meyer-Olkin

KNN - k-Nearest Neighbors algorithm

LA - Learning Analytics

LMS - Learning Management System

LMT – Logistic Model Tree

MAT01 - discrete mathematics

MSLQ - Motivated Strategies for Learning Questionnaire

NN – Neural Networks

ROC - Receiver Operating Characteristic

SEI - Student Engagement Instrument

SES - Student Experience Survey

SLR - Systematic Literature Review

SMO -Sequential Minimal Optimization

SMOTE - Synthetic Minority Over-sampling Technique

SMTSL - Students' Motivation Toward Science Learning

SOFTEX – Associação para Promoção da Excelência do Software Brasileiro

SVM –Support Vector Machine

UDESC - Santa Catarina State University

UES - University Experience Survey

VI - Value Index

SUMMARY

1 INTRODUCTION

1.1 CONTEXTUALIZATION

We can see a lack of professionals in the field of Information Technology (IT) in the face of market demand, both in Brazil (SCHELLER and SCARAMUZZO, 2019) (SOFTEX, 2013) as in the world. This lack of professionals already is felt today, and it is said that it should escalate in the next few years. In Brazil, a deficit of more than 400,000 technology professionals is expected by the year 2024 (SOFTEX, 2013) (BRASSCOM, 2019). In other places around the world, the scenario is similar. For example, in Australia, the demand for technology workers is expected to reach 700,000 until 2020 (AYNSLEY, 2015) and in Europe, that deficit expected is about 800,000 workers (GAREIS, HÜSING, *et al.*, 2014) (HÜSING and DASHJA, 2017).

Against the market demand, there is a lack of interest of teenagers and a high dropout rate of students in the IT field, mainly in the field of computing. Studies around the world describe dropout rates of 19% to 40% in programs of Computer Science and Technology (LISTON, PIGOTT, *et al.*, 2018) (MOONEY, PATTERSON, *et al.*, 2010) (DREW, 2011) (HÜSING, KORTE, *et al.*, 2013). In Brazil, studies show an average rate of 40% or more (DAMASCENO and CARNEIRO, 2018) (PALMEIRA and SANTOS, 2014) (BORBA, 2015) (VERONESE and LEMOS, 2015) (FERREIRA, 2015) (RODRIGUES, 2013). A significant number of students drop out in the first year when the universities lost 29% of Mathematics and Computer Science students (RAMAL, 2017).

In addition, another problem in the computing programs is the high rate of failures in introductory computing courses, which can reach more than 60% (BOSSE and GEROSA, 2015), with an average above 30% (BENNEDSEN and CASPERSEN, 2007) (WATSON and LI, 2014). This high rate of failures can lead to dropout or significantly increase the time to complete the course.

This is even more serious when we realize that the knowledge in computing and programming are seen as fundamental to the future professional skills (EUROPEAN POLITICAL STRATEGY CENTRE (EPSC), 2018) (BUGHIN, HAZAN, *et al.*, 2018), including many initiatives for teaching programming to children, such as "Computer Science for All" (NATIONAL SCIENCE FOUNDATION, 2019), "Code.org" (CODE.ORG, 2019), and "Computing at School" (THE CHARTERED INSTITUTE FOR IT, 2019). But if we already have trouble attracting, retaining, and getting current computing students to succeed, how will we meet that labor market demand that will grow significantly? We need to study and understand the factors that cause this scenario in order to find effective solutions.

There are several reasons that are considered factors for this high rate of dropout and failure, including area-specific factors, such as the difficulty of computer programming (NIITSOO, PAALES, *et al.*, 2014) (BERGIN and REILLY, 2005) and the lack of familiarity with the subject (CARTER, 2006). Another factor addressed by some studies is that many novice students relate STEM disciplines (Science, Technology, Engineering, and Mathematics) as being innovative and interdisciplinary. However, this view often is not confirmed by the first experiences at the university, bringing disappointment and doubt (PETERS and PEARS, 2013).

But are these the only factors that impact the success of students? According to Sinclair et al. (2015), more qualitative data and other measures (such as the expectation of the student and any specific measure) are required for the broad understanding of a Computer Science student.

In this context, other factors associated with the success and retention of students are their motivation and their engagement. According to Entwiltle (2003), motivation is one of the characteristics that influence how students approach their learning, and according to Bruinsma (2004), motivation is important in academic performance. One reason why students drop out is low motivation for study, what can influence learning outcomes (KORI, PEDASTE, *et al.*, 2016).

The lack of motivation can cause a strong discrepancy between potential and achievement in learning. This explains the reason for highly qualified students having poor performance, while students with mediocre potential being among the best (FIGAS, HAGEL and BARTEL, 2013).

Therefore, to improve students' learning and success in computing courses, it is important to understand the factors that keep them motivated and engaged. Although there are differences in the literature, in general motivation and engagement are treated as two separate things. Motivation is the stimulus for the desire to learn something (ABDULLAH and YIH, 2014) or to participate and succeed in the learning process (BOMIA, BELUZO, *et al.*, 1997), while engagement demands time, effort, and student involvement in learning activities (QAA, 2012).

Despite that distinction, many authors deal with motivation and engagement analogously or use concepts that pervade the two definitions. According to Alhazbi (2015), motivation is a key factor that affects the students' effort and consequently impacts their performance. Ngan and Law (2015) describe the motivation of learning as the propensity of student learning and the persistence of effort the student exercises for learning. Due to this, in this study, we deal with motivation and engagement but we call both "motivation".

The transition from medium to high education is a move from a controlled learning environment to a more autonomous mode. Students with a more academic background may be more comfortable with the transition to autonomous learning than those with a vocational background (SAYERS, NICELL and HINDS, 2010).

There are reports in the literature that the CS1 (Computer Science 1) courses are positioned wrongly on the program curriculum. Students, upon entering the university, are faced with many difficulties and in a moment of transition, either because they are in an environment with fewer restrictions or with a different curriculum than they were accustomed (RAMOS, FREITAS, *et al.*, 2015). According to Jenkins (2002), the subject of computer programming is already hard enough to be placed in a moment so unstable for students.

Learn computer programming is not simple for many students and can be an incredibly difficult task (WATSON and LI, 2014). Some relevant factors can influence ownership of skills such as motivation, persistence, confidence, emotional responsibility, problem-solving strategies (WINSLOW, 1996) (GOMES and MENDES, 2007). In addition, students will practice intensively (ROBINS, ROUNTREE and ROUNTREE, 2003) and will be motivated to solve several exercises.

A way to improve this scenario is to understand and measure the variables that can impact on the learning process, in order to enable take actions and do interventions.

Two areas that have recently emerged to study this context and which aims to make use of the potential of the science of analysis and data mining in an educational context (BAKER and INVENTADO, 2014) was the Learning Analytics (LA) and Educational Data Mining (EDM). LA is the measurement, collection, analysis, and reporting of student data and their contexts in order to understand and improve the learning and the environment in which this occurs (LAK, 2011). While EDM aims to “explore data from educational settings to find out descriptive patterns and predictions that characterize learners behaviors and achievements, domain knowledge content, assessments, educational functionalities, and applications” (PEÑA-AYALA, 2014).

One of the strategies often used in LA and EDM is the prediction of students’ performance or dropout. It is important to explore methods that can extract understandable and reliable knowledge of the students, in order to allow the prediction of circumvention with high enough accuracy (BAYER , BYDZOVSKÁ, *et al.*, 2012).

1.2 PROBLEM

Substantial rates of failures undermine introductory computing courses around the world and are increasing over the years (BORNAT, DEHNADI and SIMON, 2008). With regard to the introductory programming courses, Bennedsen and Caspersen (2007) performed studies that indicate an average rate of 33% failures in introductory programming courses around the world. In countries such as Portugal, Germany, and Brazil, these rates are above 50% (WATSON and LI, 2014).

In Brazil, the scenario is even worse. In a review of the literature, Ramos et al. (2015) found seven studies that reported failure rates in introductory programming courses with an average of 45.6%. Another study (BOSSE and GEROSA, 2015) found failure rates above 60%.

A significant portion of the students leaves in the first year when the universities lost 29% of students of Mathematics and Computer Science (RAMAL, 2017). According to several studies (STEPHENSON, 2018) (KORI, PEDASTE, *et al.*, 2015), the major failure and dropout rates occur in the first two years of the course and the main difficulties are related to subjects in the area of programming and Mathematics. Other studies demonstrated that there is a large concern with students' initial motivation (JENKINS, 2001) and their performance in the first year (ALLEN, ROBBINS, *et al.*, 2008).

In the university context, especially in the Brazilian scenario, it is usual comments about unpreparedness with which students arrive at university, especially when it comes to courses in the area of Mathematics or the lack of knowledge about the course. But are we reflecting on what teachers and coordinators can improve?

Therefore, the problem of high rates of dropout and failure in computing courses remains a problem for the field, mainly in Brazil. Against the increase of demand from industry, universities are not able to retain and graduate their students in order to introduce more and good professionals in the labor market. For example, in 2017, the number of graduated students in computing courses in Brazil was reduced by 4.84% (2,034 students) compared to the year 2016 (NUNES, 2018).

The first step in the process of improvement is to understand the problems and prospects of solution alternatives. In this context, aiming to increase the percentage of students successfully completing courses in computing, it's important to measure and understand what causes students to stay motivated and engaged in the program. This possibly will generate better performance and higher success rate.

Several studies tried to identify risk factors for low retention and graduation in computing courses, including knowledge about computing and previous experience, in addition to the potential and performance of the student in high school (HOWLES, 2005) so that interventions could be carried out throughout the process. These studies often use data resulting from assessments or interaction with learning management systems (LMS). The problem of this approach is that, in general, the identification of students at risk is only possible after a significant time, due to the need to perform assessments and collect temporal data from the students' interactions. In addition, many times these data are specific to the teaching plan or dependent on a specific LMS system. Thus, how is it possible to identify in advance the students at risk, using factors that can be measured from the beginning of the course and can be suitable also to other contexts?

Many of the factors that contribute to student retention and success are intrinsic to themselves even before they enter the university, such as interest in the area, the perspective of future, previous knowledge, etc. However, one aspect that seems overlooked in computing courses is that students can be motivated during the course and have an engagement that will enable them to succeed at the end of the journey, even though initially they weren't totally motivated and engaged.

In addition, we identified that despite the motivation being mentioned as a factor directly related to the success or failure of students, few works have treated the subject in the area of computing education.

> "We saw a general absence of studies in computing that used motivation, which we had expected to see. Instead, many articles seemed to be using engagement, perhaps as a proxy. More generally, psychometric data is less frequently utilized, and the increase in usage of self-efficacy and self-regulation data might be signaling growth in that area" (HELLAS, IHANTOLA, *et al.*, 2018).

Of existing works that consider motivation or some motivational aspect of computing students, few follow taxonomies or standard models to measure motivation or identify factors that impact on motivation. In addition, none of the works found consider motivation as something that can change during the course and must be monitored over time.

According to (NIKULA, GOTEL and KASURINEN, 2011) the literature focuses on individual actions aimed at improving the results of first programming courses, such as changing the programming language, synchronizing curricula and course contents, automatically assessing student assignments, emphasizing the role of process in constructing programs, and using visual tools.

Gomes (2000) proposed a set of teaching and learning recommendations and strategies to minimize the introductory programming courses unsuccessfully. However, beyond these recommendations the author emphases that the student has to show strong motivation.

In this sense, the motivation has been a factor strongly related to the students' outcome but it little used by studies that try to identify at-risk students in advance. Thus, how to identify in advance if students are not motivated so that interventions can be performed?

1.3 RESEARCH QUESTIONS

This section presents the research questions and hypotheses that will be tested in this thesis:

RQ01 – Is motivation related to the student's outcome in introductory computing courses?

- H_{A0} – Motivation has no significant relationship with the outcome of students in introductory computing courses;
- H_{A1} – Motivation has a significant relationship with the outcome of students in introductory computing courses;

RQ02 - Is it possible to detect in advance students at risk according to their motivation?

- H_{B0} – No prediction methods can identify at least 80% of failing students until de fourth week in introductory computing courses in different institutions;
- H_{B1} – One or more prediction methods can identify at least 80% of failing students until the fourth week in introductory computing courses in different institutions;

The goal of 80% was defined based on the findings described in Section 3.2.1.

RQ03 – Is there a significant variation of students' motivation during the introductory computing courses?

- H_{C0} – The students' motivation does not change significantly throughout the course;
- H_{C1} – The students' motivation changes significantly during the course and this variation is related to the students' outcome;
- H_{C2} – The students' motivation changes significantly during the course, but this variation is not related to the students' outcome.

RQ04 – Is there motivation factors related to the educational context that impact the students' outcome?

- H_{D0} – No motivation factors related to educational context impact significantly the students' outcome;
- H_{D1} – One or more motivation factors related to educational context impact significantly the students' outcome.

1.4 OBJECTIVES

The general objective of this thesis is to identify the motivation level and factors affecting motivation and engagement of students of introductory computing courses and, based on these aspects, predict in advance at-risk students of dropout or failure. To achieve the main objective, we defined the following specific objectives:

SO1: Identify factors that affect the motivation of computing students;

SO2: Measure the student motivation in the context of introductory computing courses;

SO3: Predict at-risk students in advance;

SO4: Provide relevant information about students to allow professors and coordinators acting.

1.5 RESEARCH METHOD

About the reasoning methods, this research can be classified as either inductive reasoning as deductive reasoning. According to Trochim and Donnelly (2008), "deductive reasoning works from the more general to the more specific. Sometimes this is informally called a top-down approach and inductive reasoning works the other way, moving from specific observations to broader generalizations and theories".

Initially, the research uses the inductive reasoning, since it seeks to find patterns of factors and motivation of students from initial observation to, then, creates a method that can measure these variables, testing hypotheses and validating a theory. From there, we used deductive reasoning, starting from the theory and assumptions that these variables can be used to predict in advance at-risk students, making observations and validating these assumptions. "Most social research involves both inductive and deductive reasoning processes at some time in the project" (TROCHIM and DONNELLY, 2008).

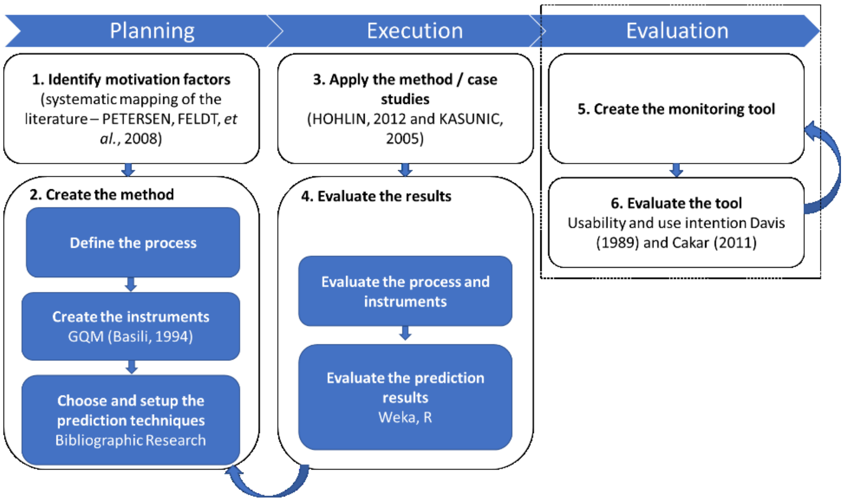
As for the type of empirical strategy, this work is classified as exploratory research. "Exploratory research is concerned with studying objects in their natural setting and letting the findings emerge from the observations" (WOHLIN, RUNESON, *et al.*, 2012, p. 9).

As to the type of data analysis, the work is classified as mixed research, because it uses both quantitative surveys, as qualitative interviews. According to Wohlin et al (2012), for quantitative data, the analysis typically includes descriptive statistics, correlation, development of predictive models, and testing hypothesis. According to Creswell (2003), qualitative research uses open-ended questions so that participants can express their views.

The philosophical position is the pragmatism that, according to Easterbrook et al. (2008), it means that the "knowledge is judged by how useful it is for solving practical problems".

For the development of the work, six phases are defined, as described in Figure 1.

Figure 1: Overview of the research method divided into three stages and six phases



Source: Developed by the author

The first phase aims to understand which features related to motivation can affect the student outcome. We planned to identify the factors that affect the motivation and engagement of computing students from the revision of existing works. For this, we planned a systematic mapping of literature about motivation in computing education, following the guidelines proposed by (PETERSEN, FELDT, *et al.*, 2008): i) definition of research questions (research scope); ii) conduct search for primary studies (all papers); iii) screening of papers for inclusion and exclusion (relevant papers); iv) keywording of abstracts (classification scheme); and v) data extraction and mapping of studies (systematic map). This phase is detailed in Section 3.1.

The second phase refers to the definition of the method for the identification and prediction of computing students' outcome based on motivation and engagement. This phase is composed of three activities:

a) Definition of the process: aims to identify the phases and activities flow of the method according to relevant features identified. For this, we propose to group the relevant features to be collected in specific moments. This activity is detailed in Section 4.1;
b) Creation of the instruments: definition and creation of scales and questionnaires to be used to measure the factors and indexes of motivation. First, we plan to map actual scales that measure student

motivation by doing bibliographic research. For the creation and validation of the questionnaires, we propose to apply a pilot test, according to the proposal of Kasunic (2005). This activity is detailed in sections 4.2, 4.3, 4.4, and 4.5;

c) Choice and configuration of prediction methods: analysis of the context and characteristics of data, to perform a search for the most suitable prediction algorithms to be used and evaluated. For this, we propose to filter data and evaluate all possible combinations. This activity is detailed in Section 5.1.

The third phase aims to apply the method, performing case studies with students of the courses of undergraduate computing programs in four Brazilian universities. For the validation of the method, we prioritize the two first years of study because, according to the literature, these are the most critical periods related to dropout and motivation (JENKINS, 2001) (ALLEN, ROBBINS, *et al.*, 2008). The dependent variable used was the student's outcome in the course (pass or fail), and dropout of course was considered failure. The following steps will be developed (WOHLIN, RUNESON, *et al.*, 2012, p. 58):

d) Case study planning: define the objectives, research questions, methods and selection strategies;
e) Case study protocol: create a “container for the design decisions on the case study as well as field procedures for carrying through the study” (WOHLIN, RUNESON, *et al.*, 2012, p. 60);
f) Preparation and data collection: conduct the data collection techniques;
g) Data analysis: analyze the descriptive statistics, correlation analysis, development of predictive models, and hypothesis testing; and
h) Reporting: communicate the findings of the study, aiming to be the main source of information for judging the quality of the study.

The fourth phase involves the evaluation of pre-selected prediction techniques and algorithms to verify their performance and any necessary adjustments to improve. For this, we used the tool Weka[1], the software environment R[2], and the following metrics:

a) Accuracy;
b) Recall;

[1] www.cs.waikato.ac.nz/weka

[2] www.r-project.org

c) AUC (Area under curve) / ROC Area;
d) F-measure.

The third and fourth phases are detailed in sections 5.1, 5.2, and 5.3. Phases 2, 3 and 4 are iterative because the method is used in pilot projects, allowing adjustments and improvements, according to the results and observations obtained.

The fifth phase refers to the development of a tool that, from the data and results obtained by the method proposed, aims to help professors and coordinators in the management of students at risk, and support in the identification of factors to interventions. This phase is detailed in Section 6.1.

In the sixth phase, it is carried out the validation of the tool with one class in a real context in order to check usefulness and ease of use issues (DAVIS, 1989). In this phase, it is also conducted an interview with the professors in order to understand their experience using the tool. This phase is detailed in section 6.2.

1.6 ORIGINALITY

There are various initiatives and projects concerned with the motivation of students in computing, and most of these works propose or report the use of new approaches and educational tools. However, few studies were found which discuss the factors and aspects to be considered to motivate students in computing courses. Also, studies that carry out a review on this subject were not found.

Similarly, there are several studies on the prediction of performance or dropout. However, there is a lack of standardization in methods to measure motivation, and in the procedures to identify the key factors that may have contributed to that. Few works were found that include motivation as a factor to predict student outcome, being that most of the works found use attributes based on grades, LMS interaction or assessments of specific tests, as shown in Section 3.2. This causes the methods to be specific to a context, and prediction to be late in the course since we need to wait for specific tests, assessments, or grades. In addition, the results obtained by current works that use psychometric data are significantly lower than those that use data from student performance or system interaction.

So, the key differential of this work is to create a simple method easy to apply, independent of context, without using grades, LMS interaction or specific tests, and to consider the motivation as an input, and which can be applied at any time during the course.

Most of the studies found that measure motivation, or consider some motivational aspect to analyze performance or dropout, do the measurement at a particular time (GRAY, MCGUINNESS and OWENDE, 2014) (HIDAYAH, PERMANASARI and RATWASTUTI, 2013). However, motivation and engagement can change over time due to several factors. To confirm this hypothesis, it is important to identify this variation and the impact of this variation on the student's performance. However, we did not find studies that consider the change in motivation over time as a factor for the prediction of students' performance.

One differential of this work is the adaptation and the longitudinal application of a scale to measure student motivation in introductory computing courses. Despite the possibility of using it in all disciplines, the focus of this work and the scope of validation is restricted to introductory computer courses due to particularities and history of high rates of dropout and failure in these courses.

Another differential of this work is the use of pre-university and weekly motivation to predict the success of students in introductory programming courses. Besides facilitating the application, because it is simple and fast, the method allows replication in other contexts, since it is not dependent on any technological instrument nor on the content of the course.

1.7 CONTRIBUTIONS

The main *scientific* contributions proposed by this work are:

a) Consolidation of the factors affecting the motivation of students of computing programs (computer education);
b) Creating a simple and easy to apply method, to identify and measure the motivation of computing students (computer education);
c) Evaluation of machine learning techniques to predict in advance at-risk students in introductory computing courses (EDM - Educational Data Mining, AI – Artificial Intelligence);

The main *technological* contributions proposed by this work are:

a) Provision of a method and tools to identify motivation and factors related to the prediction of at-risk students in introductory computing courses (EDM - Educational Data Mining);

a) Availability of a tool to support the identification of at-risk students and factors affecting failure (learning analytics, informatics in education).

The main *social* contributions proposed by this work are:

a) One social contribution of this work is to increase the effectiveness in learning and the students' retention in introductory computing courses, because the proposed method allows identifying previously students with difficulties, enabling interventions by the professor;

b) Furthermore, it allows the improvement in the quality of education by professors that can identify the main educational factors that may be impacting on student motivation and outcome.

1.8 DELIMITATION AND SCOPE OF WORK

The method proposed in this work was developed and validated specifically in the context of introductory computing courses (algorithms, introduction to programming, CS1, computational logic, etc.). Although the model can be used in other contexts, some instruments would need to be adapted, since they are specific to the area of computing, especially the pre-university factors questionnaire. However, the scope of this work is limited to the validation of the model in the context of computing students. The results of the reliability and validity of the instruments are specific to the context of computing students. We use data from only a few universities in Brazil, being the limited scope and the results can be related to specific features of the contexts discussed.

There are several other factors that can impact on student motivation and performance that are not considered in this work, such as family, job, health, etc. However, our aim in this work is to cover only aspects that can be manipulated in the educational context, allowing intervention by the professor or coordinator.

1.9 WORK STRUCTURE

This work is divided into eight chapters. In Chapter 2 we describe the main theoretical concepts for understanding the work. In Chapter 3, it is shown the state of the art in the area of motivation in computing and prediction of student performance. In Chapter 4, the proposed method (EMMECS) is detailed and validated. In Chapter 5, we present the results of the prediction of at-risk students applying the proposed method. In Chapter 6 we present the functionality and validation of a tool to facilitate the use of the proposed method. In Chapter 7 the main results are discussed. Finally, in Chapter 8, we present the conclusions and future work.

2 THEORETICAL BACKGROUND

In this chapter, we describe all theoretical contextualization, to allow understanding the research carried out. The chapter is divided into four sections. In Section 2.1 it is contextualized the concept and challenges of computing education area. In Section 2.2 data mining strategies for education are presented. In sections 2.3 and 2.4 theories and scales to measure the students' motivation are presented.

2.1 COMPUTING EDUCATION

Interest in Computer Science is growing. As a result, Computer Science (CS) faculties are experiencing unprecedented demand from other areas for computing teaching (COOPER, FORBES, *et al.*, 2016).

This growth is an unparalleled opportunity to expand the reach of computing education. However, this growth is also a unique research challenge, as we know very little about how to best teach our current students. According to Cooper et al. (2014), the expanding field of Computing Education Research (CER) is positioned to address this challenge by answering research questions such as:

a) How should we teach computer science, from programming to advanced principles, to a broader and more diverse audience?
b) How can we ensure that we retain this more diverse audience through inclusive pedagogy and generally more effective teaching?
c) How can teaching approaches and their assessment (regarding student learning) scale effectively? What training should K-12 teachers receive? What methods have been shown to be effective?
d) How can computer science teaching adapt to how different people learn and build on age-related learning progressions?
e) How should computing be taught and integrated into other disciplines?

Copper and colleagues argue that Computer Science departments should lead the way in establishing CER as a foundational research field of Computer Science, discovering the best ways to teach CS, and inventing the best technologies to teach with. This is not only in the best long-term interest of our field but also the long-term interests of society.

In computing education, we tend to focus on coming to understanding algorithms and information representations. The summit focused at a different level. Computing education researchers study ways to better understand how conflicting cultural values and self-belief that impact motivation and tackling these through design practices and tools that foster meaningful learning (COOPER, GROVER, *et al.*, 2014).

Fincher and Petre (2004) classify the CER into ten areas: student understanding, animation/visualization/simulation systems, teaching methods, assessment, educational technology, the transfer of professional practice into the classroom, the incorporation of new development and new technologies into the classroom, transferring to remote teaching ("e-learning"), recruitment and retention of students, and, finally, the construction of the discipline itself.

We understand that this work is related to two of these areas: student understanding, and recruitment and retention of students. Both are detailed as follows, according to Fincher and Petre (2004).

Student understanding. The area of student understanding is characterized by an investigation of students' mental and conceptual models, their perceptions and misconceptions. The kinds of question that researchers find motivating in this area are concerned with why students have trouble with some things, what distinguishes good students from bad students, and what are the differences between how students understand things and how experts understand things. This area of interest encompasses investigations at a wide scale from very broad topics, such as "What design behaviors do students exhibit?" and "How do students learn particular programming paradigms?" to very specific questions such as "How do students learn recursion?" Because of the nature of the evidence required in this area, given that it focuses on internal phenomena, some work has been conceived as CS education research, but there is also a lot of work of this kind that occurs within Psychology or Sociology.

Recruitment and retention of students. Issues of recruitment and retention motivate a research area, including a considerable interest, and body of work, in diversity and gender issues. There are real questions about what makes students come into CS and what makes students stay in CS. And the other side of the coin: what makes them not bother, and what makes them leave. Usually, there's a diversity perspective, but there doesn't have to be. It can be just one of the questions that concern the discipline. For example, what innate abilities contribute to performance in CS? What are the kinds of skills we can engender?

Segundo (COOPER, FORBES, *et al.*, 2016), new researchers to CER are using big data analytic techniques, and borrowing approaches from other computing disciplines, most notably machine learning. How can these "newer" techniques (for example, looking at click-through rates, intermediate forms of student artifacts, and other forms of "big data") be used to answer our questions? Do they substitute for our other research approaches, or do we use them in addition to more traditional approaches?"

In this context, there are fields of study that use technologies of data mining and machine learning in the educational context. In the following, we detail the concepts of Education Data Mining (EDM) and Learning Analytics (LA).

2.2 EDUCATION DATA MINING AND LEARNING ANALYTICS

According to Siemens and Baker (2012), and Moissa, Gasparini, and Kemczinski (2015), although EDM and LA have similar definitions, as they both address the use of educational data to improve the learning environment in which the student is inserted, LA and EDM processes are different.

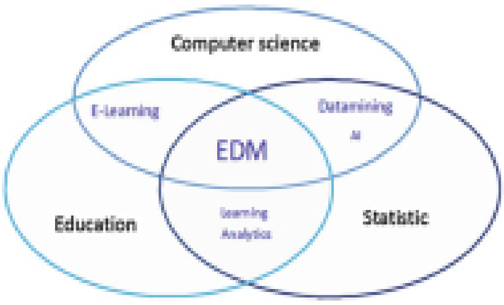
According to Siemens and Baker (2012), there are five differences between these areas: (1) the discovery of information in EDM is automatic and LA takes advantage of human trial; (2) the approach is reductionist in EDM (emphasis in reducing the problems on components and analyze it individually or your relationship with the other) and LA is holistic (understand the system as a whole); (3) LA originated in the semantic web, intelligent curricula, among others, and EDM originated in educational software, in modeling, among others; (4) EDM focuses on automatic adjustment (without human participation) while LA focuses on informing and empowering teachers and students; and (5) considering the techniques and methods, EDM uses social network analysis techniques, influence analysis, sentiment analysis, among others, and LA uses classification, clustering, Bayesian modeling, relationship mining, among others.

According to Romero et al. (2011), EDM focuses on the development of methods for exploring the unique types of data that come from an educational context. These data come from several sources, including data from traditional face-to-face classroom environments, educational software, online courseware, and summative/high-stakes tests.

According to Larusson and White (2014), "learning analytics can be summarized as the collection, analysis, and application of data accumulated to assess the behavior of educational communities".

Elatia, Ipperciel, and Zaïne (2016) say that Educational Data mining (EDM) emerged in the last few years from Computer Science as a field in its own right that uses DM (Data Mining) techniques to advance teaching, learning, and research in higher education. According to Bousbia and Berlamri (2013), EDM is an intersection between Computer Science, Education, and Statistics, while LA is an intersection between only Education and Statistics (Figure 2).

Figure 2: Areas in relation to EDM



Source: (BOUSBIA and BELAMRI, 2013)

Therefore, we realize that there are some differences and overlaps in definitions of EDM and LA. However, regardless of the nomenclature and definitions in the field of study, “data mining techniques can be used to gather the information that can be used to assist educational designers to establish a pedagogical basis for decisions when designing or modifying an environment’s pedagogical approach” (ROMERO, VENTURA, *et al.*, 2011).

One of the applications or tasks that uses data mining techniques is the prediction of student grades and learning outcomes. The objective is to predict a student’s final grades or other types of learning outcomes (such as retention in a degree program or future ability to learn), based on data from course activities. The most frequently used techniques for this type of goal are classification, clustering, and association (ROMERO, VENTURA, *et al.*, 2011).

The ultimate LA and EDM goal is to optimize both student and faculty performance and to allow instructors to judge their own educational efficacy, using both statistical and predictive techniques (LARUSSON and WHITE, 2014). According to Pardo (PARDO, 2014), the prediction phase is addressed by the following questions:

a) Which aspects of the experience need to be predicted?
b) Which factors can be used as input for the prediction algorithms?
c) What kind of prediction technique will be used?
d) How is the accuracy of the prediction going to be measured?
e) How are the predictions reported to the stakeholders?

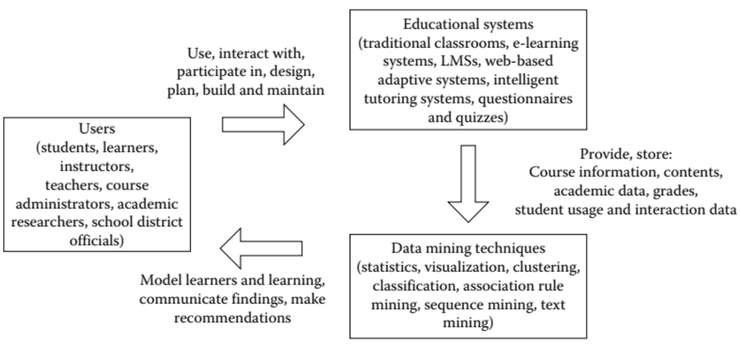
“One of the more advanced uses of analytics that generates huge interest is the possibility that from the pattern of learners’ static data (e.g. demographics; past attainment) and dynamic data (e.g. pattern of online logins; quantity of discussion posts) one can classify the trajectory that they are on (e.g. “at-risk”; “high achiever”; “social learner”), and hence make more timely interventions (e.g. offer extra social and academic support; present more challenging tasks)" (SHUM, 2012).

Romero et al. (2011) proposed to classify EDM objectives depending on the viewpoint of the final user (learner, educator, administrator, and researcher) and the problem to resolve:

a) Learners. To support a learner's reflections on the situation, to provide adaptive feedback or recommendations to learners, to respond to student's needs, to improve learning performance, etc.
b) Educators. To understand their students' learning processes and reflect on their own teaching methods, to improve teaching performance, to understand social, cognitive and behavioral aspects, etc.
c) Researchers. To develop and compare data mining techniques to be able to recommend the most useful one for each specific educational task or problem, to evaluate learning effectiveness when using different settings and methods, etc.
d) Administrators. To evaluate the best way to organize institutional resources (human and material) and their educational offer.

On the other hand, Sachin and Vijay (2012) define three main roles in the educational data mining process, consisting of users, educational systems, and data mining techniques, as shown in Figure 3.

Figure 3: Roles in the data mining process



Source: (ROMERO, VENTURA, *et al.*, 2011)

To achieve the EDM objectives, the majority of traditional data mining techniques including but not limited to classification, clustering, and association analysis techniques have been applied successfully in the educational domain (BOUSBIA and BELAMRI, 2013).

Bousbia and Belamri (2013) proposed a classification of these methods into prediction, clustering, relationship mining, distillation of data for human judgment, discovery with models, outlier detection, social network analysis, process mining, text mining, knowledge tracing, and matrix factorization.

The prediction goal is to develop a model which can infer a single aspect of the data (predicted variable) from some combination of other aspects of the data (predictor variables). The authors define three types of predictions methods: classification (when the predicted variable is a categorical value), regression (when the predicted variable is a continuous value), and density estimation (when the predicted value is a probability density function). An example of EDM application is predicting student's academic success and behaviors (BOUSBIA and BELAMRI, 2013).

According to Sachin and Vijay (2012), classification is used to predict class labels (describes future situation) and it is used to label newly encountered (still unlabeled) patterns from a collection of labeled (pre-classified) patterns.

The idea of classification is to place an object into one class or category, based on its other characteristics. In education, teachers and instructors are always classifying their students based on their knowledge, motivation, and behavior. Assessing exam answers is also a classification task, where a mark is determined according to certain evaluation criteria (ROMERO, VENTURA, *et al.*, 2011).

According to Hämäläinen and Vinni (2011), there are four types of classification problems that have often occurred in the previous research: predicting academic success, predicting the course outcomes, succeeding in the next task, and metacognitive skills, habits, and motivation.

The second group (predicting the course outcome), consists of experiments where the task was to classify the student's success in one course. The objectives were to predict passing/failing a course, dropouts, or the student's score. In most cases, the course was implemented as a distance learning course, where failure and dropout are especially serious problems (HÄMÄLÄINEN and VINNI, 2011).

In general, the authors identified that studies in this context have some characteristics. The data sets are relatively small (50–350 rows, on average 200) because they were restricted by the number of students who take the same course. Usually, the data consisted of just one class of students, but if the course had remained unchanged, it was possible to pool data from several classes. The main attributes concerned were not only exercise tasks and the student's activity in the course, but also demographic and questionnaire data. The original number of attributes could be large (>50), but was reduced to 3–10 before any model was learned. A large variety of classification methods were tried and compared in these experiments. The most common methods were decision trees, Bayesian networks, neural networks, K-nearest neighbor classifiers, and regression-based methods. The average accuracy is only 72% but in the best cases nearly 90%. The most important factors affecting the classification accuracy were the number of class values used (best for the binary case) and at how early a stage the predictions were done (best at the end of the course, when all attributes are available) (HÄMÄLÄINEN and VINNI, 2011).

According to (AGGARWAL, 2015), classification algorithms typically contain two phases:

a) Training phase: a model is constructed from the training instances;
b) Testing phase: the model is used to assign a label to an unlabeled test instance.

The output of a classification algorithm may be presented for a test instance in one of two ways (AGGARWAL, 2015):

a) Discrete label: a label is returned for the test instance;
b) Numerical score: a numerical score is returned for each class label and test instance combination. Note that the numerical score can be converted to a discrete label for a test instance, by picking the class with the highest score for that test instance. The advantage of a numerical score is that it now becomes possible to compare the relative propensity of different test instances to belong to a particular class of importance and rank them if needed. Such methods are used often in rare class detection problems, where the original class distribution is highly imbalanced, and the discovery of some classes is more valuable than others.

Therefore, this work can be considered as being in the area of Educational Data Mining, using the technique of classification through the methods of prediction, more specifically the prediction of course outcomes. As for the end user's viewpoint, the work is directed as much to educators as researchers.

The first step, therefore, is to plan and define data collection, based on the objectives of the research. In this work, the input data are related to the motivation of the student. Therefore, in sections 2.3 and 2.4, we describe some theories and scales to identify and measure the students’ motivation.

2.3 STUDENT MOTIVATION THEORIES

Since the ’50s, several theories of motivation have been created (HERZBERG, 1968) (MASLOW, 1954) to explain what moves people to act. In the educational context, there are various theories and models that try to explain and measure student motivation (ECCLES, ADLER, *et al.*, 1983) (ASTIN, 1984) (TINTO, 1975) (PINTRICH, 1991).

According to Entwistle (2003), the main findings of the research on motivation in higher education can be summarized in:

a) It describes the amount of effort put in activity and its goal;
b) It has some consistency, but may change;
c) It affects, but it is also affected by the level of performance;
d) It appears in contrasting ways;
e) The same author defines three ways of motivation towards learning and study:
f) Extrinsic motivation: focused on satisfaction in finishing the program/course, heavily influenced by pressure and external rewards;
g) Intrinsic motivation: reflects a personal goal and derives from the interest in the subject;
h) Achievement motivation: focuses on personal levels of achievement, shows the degree of student satisfaction in completing successfully a task, proving the student ability to himself, and it is related to fear of failure.

Biggs and Tang (2007) include social motivation as a distinct category of extrinsic motivation and referred to the influence of society and close people.

Several theories and scales, in the academic context, have been created to measure students' motivation (PINTRICH, 1991) (TUAN, CHIN and SHIEH, 2005) (APPLETON, CHRISTENSON, *et al.*, 2006) (GUAY, VALLERAND and BLANCHARD, 2000) (MARTIN, 2007).

According to Brophy (1983), of the several theories of motivation, expectation-value models (ECCLES, ADLER, *et al.*, 1983) provide an understandable framework to understand the students' motivation.

The expectation-value model suggests that motivation consists of two factors that predict the outcomes. The expectation, which reflects how much the student believes he can succeed in a task (related to grades, for example), and the value that reflects how the student thinks that the task is important and worth it (related to future interests, for example) (ECCLES, ADLER, *et al.*, 1983).

One of these theories based on the expectation-value model is the Expectancy-Value Theory (ATKINSON, JOHN W.;, 1957), (ECCLES, ADLER, *et al.*, 1983). These researches posit that individuals' expectancies for success and the importance of the course perceived by students are important determinants of their motivation to perform different achievement tasks (WIGFIELD, 1994). Based on expectancy-value, Kosovich et al. (2014) proposed a practical EVC (Expectancy-value-cost) scale to measure student motivation, including the factor "Cost", which measure the effort needed to perform the tasks.

Below we detail some of the main scales used to measure the students' motivation, including the practical EVC scale.

2.4 STUDENT'S MOTIVATION SCALES

We present, in this section, some works that describe student motivation scales to allow identifying and measuring the level of student motivation.

The Motivated Strategies for Learning Questionnaire (MSLQ) (PINTRICH, 1991) is a questionnaire based on a cognitive vision of motivation and learning strategies. This questionnaire consists of two sections: i) the first with 31 items to evaluate the student's value, expectancy, and affective components; ii) the second section has 31 items to evaluate the use of different cognitive and metacognitive strategies by students, in addition to nineteen items to evaluate the management of different resources by students.

In all, fifteen subscales are evaluated, being the first six regarding motivation:

a) Intrinsic goals orientation: the student realizes that the reason to be participating in the tasks is the challenge and curiosity in mastering the subject;
b) Extrinsic goals orientation: student's perception that the main reasons for participating in activities are not directly related to the task itself, like note, reward, evaluation, competition, etc.;
c) Value of the task: perception of why you're doing the task;
d) Learning control: belief that efforts to learn succeed;
e) Self-efficacy for learning and performance: the expectation of success, confidence that has skills to develop the activities;
f) Anxiety test: negative thinking and emotional/psychological issues of anxiety;
g) Repeat test: read and repeat content for study;
h) Preparation: using strategies like paraphrases, abstracts, analogies, etc.;

i) Organization: select and relate appropriately the information learned;
j) Critical thinking: apply knowledge in new situations to solve problems;
k) Self-regulation: aspects of consciousness, knowledge, and control of cognition, involving activities of planning, monitoring, and regulation;
l) Time and study environment: activities to manage and regulate the time and study environments, such as planning and scheduling time, study site;
m) Effort: keep attention even with difficulties and distractions;
n) Peer learning: collaboration with peers;
o) Search for help: know when and where to come when you need help (tutoring, peers, teachers).

Each subscale has the linked items and each item has one to seven assessment (Likert scale). With this, it is possible to calculate an index of overall motivation and one for each subscale.

The Students' Motivation Toward Science Learning (SMTSL) (TUAN, CHIN and SHIEH, 2005) is a questionnaire to measure the motivation of students for learning science. It is divided into 36 questions with answers from 1 to 5 (Likert scale) and five scales: i) effectiveness – belief in their ability to perform activities well; ii) active learning strategies – the use of various strategies to build new knowledge based on prior understanding; iii) value of learning science – finding the relevance of science in everyday life; iv) performance objectives – compete with other students and gain attention from the teacher; v) stimulating learning environment – curriculum, faculty, and student interaction.

The l'Échelle scale of motivation in education (EME) was developed in Canadian French by Vallerand et al. (1989). This scale assumes the multifactority of the motivational processes. It is composed of 28 items and punctuated on a Likert scale with seven points. EME was subsequently translated into English, originating the Academic Motivation Scale (AMS) (VALLERAND, 1992).

The theoretical or structural factorial model of EME and AMS presents the intrinsic motivation in the way of the subscales: intrinsic motivation for knowledge (IMK), intrinsic motivation for realization (IRR) and intrinsic motivation to experience stimulation (IMES). In addition to the subscale of motivation (AMO), the model also has three other subscales that group the various ways of extrinsic motivation: by identification (EMId), by introjection (EMIn) and by external regulation (EMER).

a) Intrinsic motivation for knowledge (IMK): the student engages in activities aimed at learning, learning for pleasure and satisfaction from exploring or understand something new;

b) Intrinsic motivation for realization (IMR): the student perform or create something, overcoming the limitations known, producing satisfaction and pleasure, leading the student to engage in activities;
c) Intrinsic motivation to experience stimulation (IMES): the student invests in academic activities, in order to experience the stimulating and challenging sensations of sensory or aesthetic nature;
d) Extrinsic motivation for identification (EMId): the student has a reasonable perception level of the importance of his/her actions and acceptance of responsibility, becoming involved with a higher degree of willing and lower pressure sensation/external control;
e) Extrinsic motivation by introjection (EMIn): it is based on external contingencies and controlled by external pressures, such as implicit offers of aggrandizement or implied threats of embarrassment. The student acts according to certain standards or expectations to avoid constraints that cause guilt, shame or seeking positive self-assessment;
f) Extrinsic motivation for external regulation (EMEr): the student feels pressured by others to act in a certain way. This pressure manifests itself in the form of reward or punishment;
g) Motivation (AMO): there is no interest or inspiration, encouraged internally or externally, for which the student act toward an academic goal, expressing indifference or disinterest.

A version of the AMS was translated into Portuguese and validated by Sobral (2003). The questionnaire can be seen in Annex A.

The Student Engagement Instrument (SEI) (APPLETON, CHRISTENSON, *et al.*, 2006) is a questionnaire to measure student engagement, containing thirty items that aim to measure the level of student cognitive engagement and 26 items aimed at examining the psychological engagement of the student's perspective. All items are evaluated on a 4-point Likert scale. The items are divided into five groups of factors: i) student-teacher relationship; ii) control and relevance of activities; iii) peer support for learning; iv) future aspirations and targets; v) family support for learning.

The Student Experience Survey (SES) or University Experience Survey (UES) (WHITELEY, IAROSSI and BRICKNALL, 2015) was created to measure the level of engagement and satisfaction of first and last year students at the Australian universities. It consists of five groups of questions: student engagement, teaching quality, learning resources, student support, and skills development. The engagement group contains seven factors, which generate a scale and an engagement index: i) feeling prepared for studies; ii) sense of belonging to the institution; iii) online or face-to-face discussions; iv) working with other students; v) student interaction outside the class; vi) interaction with different students and; vii) opportunities to interact with local students.

Another existing scale is the EVC scale. The initial version of the EVC Scale included 24 items used in undergraduate general education courses. This larger pool of items was reduced to twelve for use in evaluating an online intervention in high school science. Later, Kosowich et al. (2014) created a brief, ten-item scale, to measure middle school students' expectancy, value, and cost for math and science. The so-named practical EVC scale has the following items divided into expectancy, value, and cost (KOSOVICH, HULLEMAN, *et al.*, 2014):

- E1 - I know I can learn the material in my [math or science] class.
- E2 - I believe that I can be successful in my [math or science] class.
- E3 - I am confident that I can understand the material in my [math or science] class.
- V1 - I think my [math or science] class is important.
- V2 - I value my [math or science] class.
- V3 - I think my [math or science] class is useful.
- C1 - My [math or science] classwork requires too much time.
- C2 - Because of other things that I do, I don't have time to put into my [math or science] class.
- C3 - I'm unable to put in the time needed to do well in my [math or science] class.
- C4 - I have to give up too much to do well in my [math or science] class.

Despite the importance given to the motivation and engagement in the success of students, few works in this context were found in the area of computing. A systematic mapping of literature (SCHOEFFEL, WAZLAWICK and RAMOS, 2018a) found studies that joined computing and motivation, identifying 32 relevant studies. However, the authors found that little more than half of the studies (53%) use some model/protocol previously proposed. Of these, only two equal models/protocols were used in two different studies. The other thirteen studies were based on thirteen other different previous models/works; that is, none of these was reused.

In order to compile the results of these scales, we evaluated a set of important criteria for the scope of this investigation: measure motivation, course-oriented or initial motivation, brief (less than 10 items or 10 minutes), based on some theory of motivation, as shown in Table 1.

Table 1: Comparison of motivation scales

Motivation scale	Measure Motivation	Course-oriented	Initial Motivation	Based on theory	Brief
MSLQ	Yes	No	No	No	No
SMTSL	Yes	Yes	No	No	No
AMS	Yes	No	Yes	No	No
SEI	Yes	No	No	No	No
EVC	Yes	Yes	No	Yes	Yes

Source: Developed by the author

As our goal was to measure initial motivation and motivation throughout the course, the scales that best fit the criteria used were: i) Academic Motivation Scale (AMS) (VALLERAND, 1992) translated into Portuguese by Sobral (2003), which is used to measure the initial motivation of students; and ii) an adaptation of the EVC Scale (Expectation-Value-Cost) proposed by Kosovich et al. (2014), that is used to measure the motivation of students through the course.

As the goal of this work is to use the scale repeatedly during the course, we adapted the practical scale proposed by Kosovich et al. (2014). We named the new scale "EVC Light". We reduced and simplify the questionnaire from ten to six items, in order to agile and simplify the process. The proposed scale is detailed in Section 4.3.

3 STATE OF THE ART

This chapter presents the studies that contribute and perform research similar to this work, which allows to understand the key differences and to compare the results of the proposed work. The chapter is divided into 2 sections. Section 3.1 describes the related work on motivation and computing, detailing a systematic mapping of the literature conducted on the subject. Section 3.2 describes studies on the prediction of students' performance or outcome and a systematic mapping of literature performed on the prediction of success in introductory computing courses.

.

3.1 SYSTEMATIC MAPPING ABOUT STUDENT MOTIVATION IN COMPUTING

No literature review or systematic mapping studies were found about motivation or engagement of computing students. Systematic reviews were found about motivation in software engineering but focused on professionals (BEECHAM, BADDOO, *et al.*, 2007) (FRANÇA, GOUVEIA, *et al.*, 2011). The literature presents some reviews on learning analytics or, more specifically, on the prediction of student performance or dropout, not focused on computing, which mentions some important features to help to predict dropout, including motivation (CRISP, TAGGART and NORA, 2015) (SHAHIRI, HUSAIN and RASHID, 2015). So, we conducted a systematic mapping study in order to do the categorization and analysis of several studies, allowing it to be a source of more complete and grouped information, focusing on factors that affect the motivation of students in computing.

3.1.1 Research Method

We conducted a systematic mapping study on the motivation of students in computing undergraduate programs. It has followed the guidelines proposed by (PETERSEN, FELDT, *et al.*, 2008) and (KITCHENHAM, BUDGEN, D. and BRERETON, 2010). The main research question is "what are the factors that impact students' motivation in computing undergraduate programs?". We defined four analysis questions to guide our findings:

- AQ1 – What are the categories of motivation and/or engagement evaluated?
- AQ2 - What are the assessed factors/aspects that impact on motivation and/or engagement of students in computing undergraduate programs?

- AQ3 – What are the factors that most impact on motivation and/or engagement of students in computing undergraduate program?
- AQ4-What are the protocols or models used for assessment of student motivation/engagement?

3.1.1.1 Search mechanism and terms

The search was performed in the following databases: ScienceDirect (Elsevier), ACM Digital Library, IEEE Explore, SpringerLink, PNAS - Proceedings of the National Academy of Sciences, Web of Science, Scopus, MathSci (AMS), CiteSeer, Wiley Online Library, IOPscience Institute of Physics - IOP), World Scientific Electronic & Communication (ProQuest), Science (AAAS), SciELO.ORG, Computers and Applied Sciences Complete (EBSCO), Institution of Civil Engineers, Cognitive Sciences Eprint Archive, DOAB: Directory of Open Access Books, ArXiv.org, Oxford Journals, Computer and Information Systems (ProQuest), AIP Scitation - American Institute of Physics, LSM2 - Logic Substitution Model (IIASA), Cambridge Journals Online, Academic Search Premier – ASP (EBSCO), and ACM Computing Reviews.

We used the following search terms and expressions: computing undergraduate programs or courses: (“computer science” OR “programming” OR “computer science” OR “CS1” OR “software engineering” OR “software development” OR “computing” OR “information systems” OR “coding”), motivation/engagement: (“motivation” OR “motivate” OR “engagement” OR “engage”).

3.1.1.2 Inclusion and exclusion criteria

We included (inclusion criteria) studies with students in computing undergraduation programs, with discussion or evaluation of relevant factors, published at any year until 2016, and written in English.

We did not include (exclusion criteria) studies that do not have relationship with computing undergraduate programs, subjects not related to computing, studies that describe reports of application of a specific approach or education strategy even if it includes motivation as one factor analyzed, studies that evaluate specifically the issue of gender in computing students (because this factor is already covered in other studies), studies that evaluate only e-learning courses, studies that address elementary, middle, high, or postgraduate students, and other matters which have no direct relation with the motivation in computing undergraduate programs and courses.

3.1.1.3 Quality analysis

In addition to the inclusion and exclusion criteria that must be clear, Kitchenham (2007) points out the importance of checking the quality of selected studies and reduce the bias of the search. In this section, we present the quality criteria used in this study and the presence or not of these criteria in the papers evaluated.

For each of the papers evaluated in this work, some data has been cataloged on a spreadsheet with the goal of presenting clearly their main features: citations (Google Scholar), education area, type of vehicle, type of strategies, sample, and context.

The average number of citations was 12.59. Twenty-four studies were published in journals or conferences related to Computer Science, while seven works are related to other areas and a study was not the source of the publication information. About the research strategies, the studies were divided as shown in Table 2.

Table 2: Research strategies of studies mapped by SMS

Research strategy	**Studies**
Survey	13
Case study	7
Bibliographic research	4
Quasi-experimental	3
Documental research	1
Experiment post-test only	1
Interview	1
Method proposal	1
Multi-methods	1

Source: Developed by the author

Of selected works, it can check on Table 2 that the type of research project more used is the use of questionnaires (surveys), followed by the case study. Together, these types of research covering more than 60% of the studies. Another finding is the low number of experiments, with only four studies using post-test or quasi-experiments, and no papers used pre and post-test experiment with a control group, for example.

The sample average of the researches was 148.85 individuals, excluding bibliographic and documentary research papers. Considering only survey researches, this average rises to 188.77 individuals.

We also evaluate the quality of selected studies considering three aspects: problem relevance, contribution, and results. For each aspect was created a question and response options of 0 to 4, with 0 (non-existent), 1 (very negative), 2 (Slightly negative), 3 (Slightly positive), 4 (very good):

a) Is the problem to be resolved clear and motivated?
b) Is the contribution clear?
c) Do the results of the experiments convince and validate the proposal?

The average of the evaluations can be checked in Table 3. The list with the evaluation of all the studies is available in Appendix K.

Table 3: Quality of selected studies

Problem/ Motivation	**SD [a]**	**Contribution**	**SD**	**Results**	**SD**	**Mean**	**SD**
3,6	0,66	3,6	0,72	2,9	1,00	3,00	0,55

a. Standard deviation

Source: Developed by the author

Based on selected studies, the main problem identified is the demonstration and validation of results. Almost half of the studies (14) were evaluated negatively, indicating that have shown no qualitative results, do not validate the results statistically, or not exploring all the data collected.

3.1.1.4 Selection process

The protocol was evaluated using the technique of peer review. Two researchers in the area reviewed and validated the protocol. The search was performed during December of 2016.

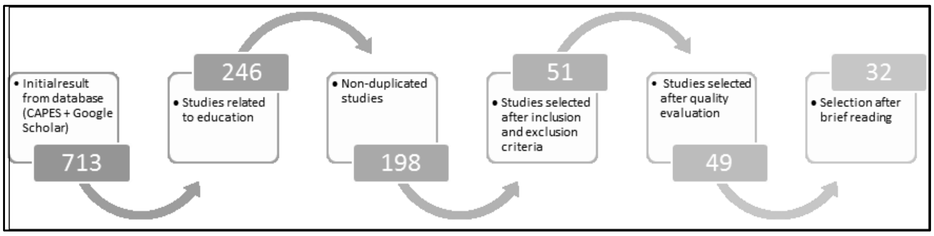
As a preliminary result, following the application of the search terms in the databases, we obtained 713 studies. From this first selection, studies not related to computing education were rejected and after title and summary analysis 246 studies remained. After the rejection of duplicated studies, 198 studies were left. Of these, the filters (inclusion and exclusion criteria) were applied resulting in 51 studies.

We excluded by the criteria of quality other two papers which had no citations, resulting in 49 selected studies. From this selection, we have conducted a brief reading of all studies to evaluate in more detail the results of each article that were not clear or explicit in the abstract.

This analysis has eliminated another seventeen studies, for the following reasons: i) two studies were not related to learning (addressing motivation for professionals and experience in the industry); ii) seven studies with full paper not found or not existing (e.g. panels on conferences); iii) two studies about teaching on middle, high or fundamental level; iv) one about teaching in another undergraduate program; v) one about gender issues; vi) one comparison of international surveys; vii) two case studies with proposals for improving strategies of motivation, but without analyzing of motivation itself; viii) one study repeated in other paper.

Finally, it has remained 32 studies. Figure 4 shows an overview of the section process. After data tabulation, extraction and reading, we consolidated the results found in the following section. The list of all selected papers and more details are available in Appendix A.

Figure 4: Summary of the selection process



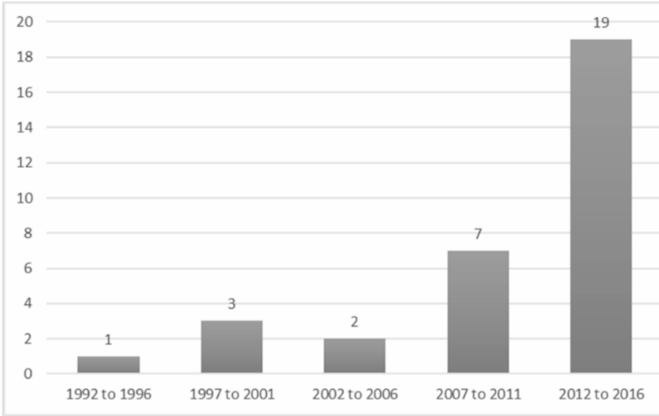
Source: Developed by the author

3.1.2 Results

Most of the 32 selected studies are recent, having significant growth in the 2010s. We highlight that the period from 2013 to 2016 has seventeen selected studies (53% of the total). Considering the last ten years, we have approximately 80% of the selected studies, as shown in Figure 5. This demonstrates that the subject is a current concern in the computing field.

Figure 5: Chart of studies by year/period

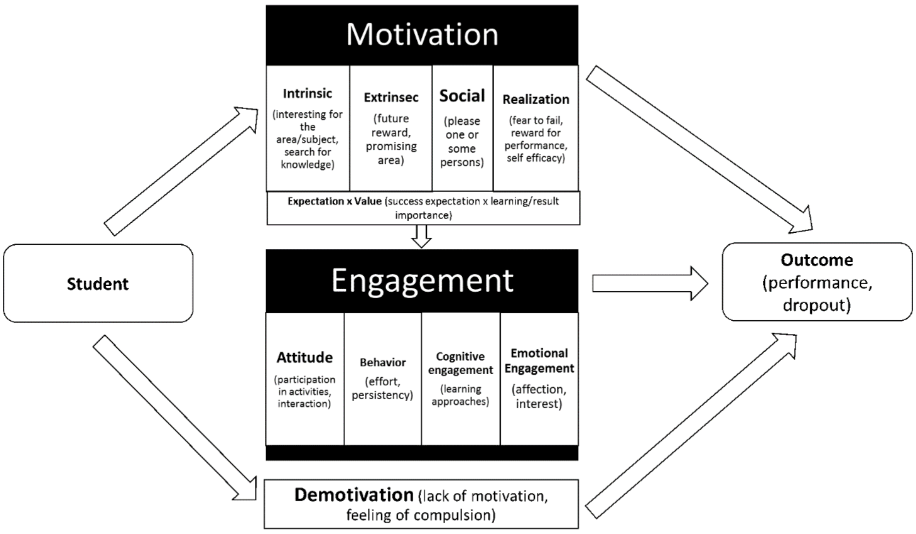
Year	Studies
1992	1
1999	**1**
2001	2
2002	**1**
2006	1
2007	**1**
2010	3
2011	**3**
2012	2
2013	**5**
2014	4
2015	**4**
2016	4



Source: Developed by the author

We have verified that motivation/engagement is divided into some standard categories. From the evaluation of the studies, we have created a taxonomy based on the classifications of Jenkins (2001) and Kanaparan, Cullen, and Mason (KANAPARAN, CULLEN and MASON, 2017) as shown in Figure 6.

Figure 6: Proposal for a categorization taxonomy of for motivation and engagement in computing students



Source: Developed by the author

The proposed taxonomy defines that there are three aspects that may impact on the results: motivation, engagement, and demotivation. Motivation involves the reasons/motives for the student to take the undergraduate program. Engagement is the practical result in student activities, which is impacted by motivation. Demotivation is the absence of a specific motivation or the feeling of doing the program simply by obligation.

Motivation involves the following aspects (JENKINS, 2001):

- *Intrinsic motivation*: the main motivation is the interest and taste for computing itself, obtaining knowledge;
- *Extrinsic motivation*: the main motivation is the career and rewards arising with the successful end of the undergraduate program;
- *Social motivation*: the main motivation is to please someone (family, friends, professor, etc.), rising status in the classroom;
- *Realization*: the main motivation is personal satisfaction to achieve a good performance, fear of failure;

The motivation can still be evaluated under the expectation and the value/importance that the student perceives about the course or program. Expectation refers to performance expectation and value refers to how important, interesting, and useful the task is (PINTRICH, 1991).

On the other hand, engagement can be classified as (KANAPARAN, CULLEN and MASON, 2017):

- *Attitude*: participation in activities, interaction with professors and peers;
- *Behavior*: effort and persistence of the student;
- *Cognitive engagement*: adaptation to teaching/learning approaches used;
- *Emotional engagement*: affection and interest shown by the student for the program or course, professor, classmates, etc.

We still have the demotivation or no particular motivation, with answers like "just to pass".

Table 4 shows the results by category, identifying the number of studies that use each category to classify motivation. The amount of studies is greater than 32 because each study can deal with more than one category. None of the categories is present in all studies; moreover, the most used category is present in only fourteen studies (43.8%). This demonstrates the lack of standardization in the analysis of motivational categories. Another important point, as Table 4 shows, is that demotivation or lack of motivation is the most cited category, which shows that researchers are concerned about identifying factors that demotivate the students. A significant number of studies (11) classifies motivation into intrinsic or extrinsic.

Table 4: Studies by categories of motivation/engagement

Category	Studies	% Studies
Demotivation/lack of motivation	14	43.8%
Intrinsic motivation	13	40.6%

Extrinsic motivation	11	34.4%
Realization	10	31.3%
Attitude/engagement	8	25.0%
Motivation in General	7	21.9%
Social Motivation	5	15.6%
Expectation x value	2	6.3%
Others	8	25.0%

Source: Developed by the author

To classify the motivation factors found, we divided them into five categories: student, professor/teaching, program/course/content, environment/university, and social, as well as their respective subcategories. All of them are described below and the list of all the factors and where these factors were found can be seen in Appendix A.

Table 5 describes the factors of each category. Category "student" groups factors related to the behavior and characteristics of the students. Category "professor/teaching", brings together factors related to professors and the teaching-learning process. Some factors are duplicated because they are part both of the "students" category as well as the "professor/teaching" category. Factors of the categories "program/course/content, environment/university, and social (external)" refer to the structural aspects of the program and facilities the university, besides the influence of factors external to the university environment.

Table 5: Motivation factors of the "student" category

Factor	Description	Sub-factors
Student category		
Student-professor interaction	It refers to the frequency and quality of the professor-student relationship. It is related to the support and openness to discussions by the professor, communication, and coexistence between professor and student, as for the number of emails and information exchanged between students and professors.	- Quality of interactions between student and professor - Access to activities and materials in the learning management system - Amount of digital interactions between student and professor - Inclusive learning community - Feedback and help from the professor - Promotion of discussions
Prior knowledge	Refers to the knowledge acquired by the students prior to their entry into the university, from their general and broad knowledge, until their ability to programming and computing in general.	- Performance in high school - The broad and social vision of the area - Perception of the area and nature of the course - Previous knowledge in programming - Quantitative reasoning
Future aspirations	Refers to the student's desire for the future, related primarily to the professional future, but also to the application of the content in the real world.	- Career/ employment opportunities - Use of computing in other areas - Be the basis for future specializations

Participation	Refers to the effectiveness of student participation in class, involving presence, extra classroom studies, and participation in activities.	- Presence in class - Extra classroom study - Completion of the proposed activities
Reason to be attending	Refers to the initial reason to choose and attend computing.	- Reason to enter the course
Experience	Refers to the time of prior experience in computing.	- Professional experience - Programming experience - Experience in the course
Confidence	Refers to how confident the students are with what they are learning, related to the efficacy and self-regulation.	- Self-confidence
Gender	Refers to the influence of gender on performance and motivation of the students.	- Gender
Behavior	Related to the fears, attitudes, and mentality of students.	- Attention - Growth mindset (think that can improve) - Fear - Behavioral engagement (effort, persistence) - Emotional engagement (affection, interest)
Learning	Refers to the pleasure to learn and apply knowledge, beyond the perception of their learning level.	- Self-efficacy (perception to be able to learn) - Learning at high levels (Bloom's taxonomy) - Pleasure in learning - Desire to apply knowledge
Satisfaction/ entertainment	Refers to how much pleasure and fun are the activities.	- Satisfaction/feeling happy - Fun (teaching-learning process enjoyable/funny)
Practice/ study	Refers to the way of study, involving time management, available resources and practical activities outside the university environment.	- Activities outside the university context - Studying right (time management, problem-solving) - Mature study practices
Independence	Freedom and autonomy of the students to participate in the discussions and definitions of the activities.	- Independence of students (participation in decisions, activities)
Professor/teaching		
Teaching-learning strategies	Refers to how the professor addresses teaching practice, types, and varieties of pedagogical activities.	- Interesting examples - Fun environment - Diversity of approaches - Relevant to the market and updated information - Improving team spirit - Active learning, practice - Collaborative learning - Promoting reflection
Student-professor interaction	Described in the student category	
Level of difficulty	Refers to the adequate difficulty level of contents (neither easy nor difficult).	- Adequate difficulty level - Self-efficacy (perception of being able to learn)

Defined and relevant goals	Refers to the explicit definition of course goals and how important they are.	- Clear learning goals - Understanding of contents relevance
Professor issues	Refers to items related to training, abilities, and gender of the professor	- Training of professors - Education background of professors - Gender
Satisfaction/ entertainment	Described in the student category	
Independence	Described in the student category	
Reward and recognition	Professors recognize and value tasks well performed	- Reward and recognition
Punishment	Refers to the existence or not of punishment, in addition to its application form.	- Punishment
Challenges	Existence of challenging and intriguing goals.	- Challenging goals
Assessment	The flexibility of the assessment to meet all kinds of students.	- Flexible assessment
Other categories		
Course/ content	Refers to the actual content of the course, its curriculum, and syllabus.	- Course content - Distaste for programming - Challenging and interesting area - Type of the course (class shift, emphasis) - The course aimed at the market
Programming language	Refers to the programming language used in teaching.	- Programming language(s) used in teaching
University environment	Refers to the learning environment, the support provided by the university and students' socialization.	- Socialization within the class - Adequate support environment - Student interaction with the University
Virtual environment	Impact of having or not an adequate virtual environment.	- Adequate virtual environment
Social influence	Refers to the influence of some friend or family member and to social pressure to choose or remain in the course.	- Influence of family/friends - Society pressure

Source: Developed by the author

Table 6 shows the frequency of factors described in selected studies, for each category, as well as the number of studies that mention at least one factor in the category.

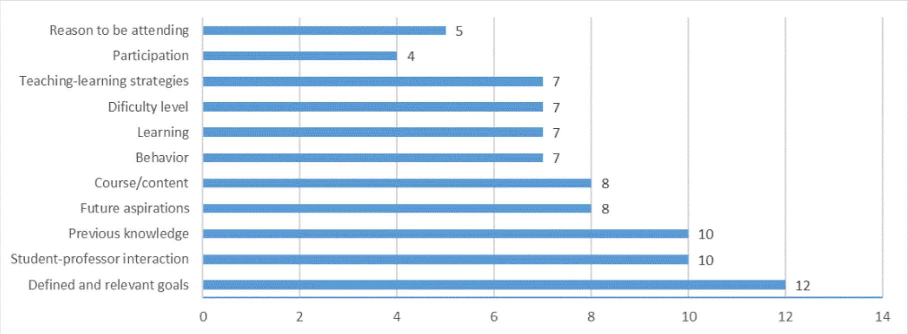
Table 6: Impact factors on motivation/engagement

Category	**Studies**	**% Studies**
Student	26	83.9%
Professor/teaching	20	64.5%
Course/content	10	32.3%
Environment/university	6	19.4%
Social (external)	5	16.1%

Source: Developed by the author

As Figure 7 shows, grouping all the factors by incidence, only three of the most cited factors can be linked directly with intrinsic motivation: learning (7), student behavior (7), and prior knowledge (10). The other are factors that receive interference from external agents, such as teaching strategies (12), student-professor interaction (10), future aspirations (8), course/content (8), difficulty level (7), and defined and relevant goals (7).

Figure 7: Most cited factors chart



Source: Developed by the author

This result seems to go against the belief that the intrinsic factors are the most important because most of the mentioned factors refer to aspects that may be changed, improved, and controlled externally. This allows us to make two observations that need to be investigated with greater control: i) current studies are not measuring all the factors that affect the motivation of students; ii) there are factors related to the educational context that may influence the change of motivation through the course or program.

To evaluate the impacts for each factor in each study, we have classified the results described according to a Likert scale (I4 - strong impact proven statistically with degree of confidence of 95% or more, I3 - strong impact not proven statistically, I2 - weak impact statistically proven with degree less than 95%, I1 - weak impact unproven statistically, I0 - no impact).

Table 7 shows that not all the factors mentioned were evaluated. This occurs because some factors were mentioned as impacting on motivation, but only descriptive or empirically, and no study has shown experimental results about them. These factors are: i) student behavior; ii) challenging goals; iii) programming language; iv) university environment; v) confidence; and vi) professors issues.

Table 7: Summary of the factors and their impact on motivation/engagement

Factor	Qty of Sub factors	Studies	Average Impact	I4	I3	I2	I1	I0
Teaching strategies	10	6	2.80	2	6	0	2	0
Participation	10	4	2.30	2	5	0	0	3
Reasons to enroll	6	4	3.83	5	1	0	0	0
Social influence	4	4	3.25	3	0	0	1	0
Learning	4	4	1.75	1	1	0	0	2
Gender	7	3	3.43	3	4	0	0	0
Future aspirations	3	3	3.67	2	1	0	0	0
Course/content	3	3	2.67	1	1	0	1	0
Prior knowledge	3	3	2.00	1	0	1	0	1
Experience	3	3	1.67	1	0	0	1	1
Student-professor interaction	4	2	2.75	2	1	0	0	1
Defined and relevant goals	3	2	3.67	2	1	0	0	0
Reward and recognition	3	2	1.67	1	0	0	1	1
Level of difficulty	2	2	2.00	0	1	0	1	0
Satisfaction/ entertainment	2	2	1.00	0	0	0	2	0
Punishment	1	1	4.00	1	0	0	0	0
Practice/study	1	1	4.00	1	0	0	0	0
Virtual environment	1	1	4.00	1	0	0	0	0
Initial motivation	1	1	4.00	1	0	0	0	0
Course/discipline type	1	1	3.00	0	1	0	0	0
Evaluation	1	1	3.00	0	1	0	0	0
Independence	1	1	1.00	0	0	0	1	0

Source: Developed by the author

Notice that the factors whose impacts have been assessed do not repeat in many studies, and "teaching strategies" is the most mentioned factor (6 times). As a result, of the factors with more than one mention, the only ones with unanimity in the responses of strong impact were "reasons to enroll", "future aspirations", "defined and relevant objectives" and "gender". Interestingly, these factors are assigned to intrinsic motivation, extrinsic motivation, teaching-learning strategies, and previous characteristics, respectively.

Regarding some factors such as "experience", "prior knowledge", "learning" and "participation" there are differences among studies. Let's see the example of the "experience" factor, which was evaluated by three studies, one of them with a strong impact proven (KORI, PEDASTE, *et al.*, 2016), another with weak impact (MAMONE, 1992), and a third with no impact (SAYERS, NICELL and HINDS, 2010).

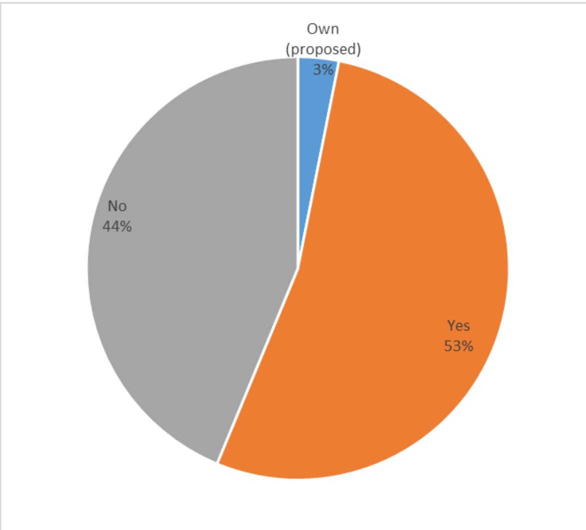
Other factors indicate a positive impact, but with caveats: "teaching strategies", "social influence", and "program/course/content". As an example, the "social influence" factor was rated by four studies, and in three of them it had a strongly positive impact (LAW, LEE and YU, 2010) (ZAINAL, SHAHRINA, *et al.*, 2012) (RAO and WANG, 2014), but on the other had a weak impact (MCCARTNEY, BOUSTEDT, *et al.*, 2016). This indicates that these factors need to be investigated with more accuracy.

Therefore, few studies (17) with demonstrated results were found, and there are still many divergences (factors that present different results depending on the study).

Notice that most of the seventeen studies that have shown results used factors related to students and professors. However, only a few have evaluated factors related to program/course, environment, and social aspects, although some of those factors have shown positive results strongly impacting the motivation in studies.

Figure 8 shows a chart summarizing the number of studies that used or are based on some evaluation model/protocol of motivation proposed earlier, even if partially.

Figure 8**:** Graphic of use of research models/protocols



Source: Developed by the author

A little more than half of the studies (53%) use some model/protocol previously proposed. However, only two models/protocols were used in two different studies: the ARCS (Attention, Relevance, Confidence, and Satisfaction) model (KELLER, 1987) and the model described by Law and other (LAW, SANDNES, *et al.*, 2009). The latter was used in two studies by the same authors who created the template. The other 13 studies were based on 13 other different previous models/works, that is, none of these was reused.

3.1.3 Discussions

We carried out a systematic mapping study to identify factors that affect motivation and engagement in computing undergraduate students. We identified 32 relevant studies. We identified that a significant number of studies assess motivation using the categories of intrinsic and extrinsic motivation, as well as many studies seek to identify demotivation or lack of motivation of the students. Nevertheless, we have noticed that there is no standardization in this categorization since the category most used was mentioned by only 43% of the studies. We have identified several factors, which were grouped into five categories: student, professor/teaching, program/course/content, environment/university and social (external) and respective 29 subcategories (factors). As a result, a large portion of the studies deals with the factors related to student and professor/teaching, but the minority considers factors related to program/course/content, university environment, and social aspects. This demonstrates that the current studies are not extensive, often being limited to work with only some specific factors.

It can also be demonstrated a lack of standards in studies on the motivation of computing students, that assess different sets of factors (considering that the most used factor – teaching strategies – was cited by less than one-third of the studies), what may be due to too specific or limited studies. According to (RYAN and DECI, 2000) and (JENKINS, 2001), students with intrinsic motivation have better results and greater success. As the intrinsic motivation refers to the fact that the students are attending the course or program because they like it and are interested in computing, factors like "reason for attending" and "learning" should be the most frequent. However, we have noticed that these factors are reported only by a little more than 20% of studies. The most mentioned factors were related to the teaching-learning strategies (37.5%), professor-student interaction (31.3%) and prior knowledge (31.3%). This indicates that the current studies are considering and concerned about the impact of the learning-teaching process in students' motivation, suggesting that the motivation/engagement may vary throughout the undergraduate program as students interact with the educational components.

These results indicate that there are issues not fully covered in the literature about the factors that impact on motivation/engagement of computing students. The results also indicate that the current studies related to computing students are not following a standard with respect to the model or protocol used, which can hinder replication studies and comparison of results. The possible reason for that is that standards and models covering all aspects and factors identified are lacking.

Another finding was that the existing studies mostly aimed at assessing the motivation/engagement in a specific moment of the student cycle. However, we have understood that it is important to monitor the motivation/engagement of the student during all the undergraduate periods. Another interesting aspect we have identified was that, despite the literature saying that intrinsic motivation has a greater impact on the student's results, many studies are being carried out considering extrinsic aspects, previous and social factors, and self-realization, among others. This can demonstrate the concern of the computing community with aspects that can be somehow be affected by the academic institutions involved. On the other hand, we found only a few papers presenting controlled studies and evaluations on the impact of the factors in the motivation of students.

It is possible to create several actions and tools to improve the teaching-learning quality and students' retention, based on the mapping of motivation factors. For example, identifying that students are dissatisfied with the "Level of difficulty" factor may indicate a lack of prior knowledge of some subject or that professor is overcharging beyond that expected. Another example, a poorly evaluation for "Teaching-learning strategies" factor may indicate that professors are not varying their teaching strategies, and this may be doing classes boring and may not suit all student profiles. One solution would be to train professors in new pedagogical approaches, for example. From the definition of these variables, it is possible to construct computational tools, using data mining, machine learning, and statistics to support decision making and to discover information and connections for predicting and advising people's learning.

From the mapping of motivation factors, it is also possible to create specific instruments for measurement and control. For example, tools can be created to evaluate pre-university factors in order to check what actions in schools or entry selection for freshmen can be made. Instruments can be created to identify satisfaction with the factors related to the program, to verify the need and critical points for curricular reform. An instrument can be created to identify factors related to professors and classes, in order to plan actions for training and defining teaching strategies. Instruments can be created to identify factors that make students drop out.

3.2 STUDENTS OUTCOME PREDICTION

In this section, we describe some concepts and related works about student's outcome prediction. In section 3.2.1, we describe some studies related to student prediction.

In section 3.2.2, we describe a mapping study of literature that we conducted about student prediction in introductory computing courses.

3.2.1 The State of Art in Student Prediction

The term ‘prediction’ is generally used to characterize models (based on EDM techniques) designed for predicting new outcomes or scenarios based on new observations (BOUSBIA and BELAMRI, 2013).

A general problem that faces higher education administrators is mining institutional data to identify predictors. This usually begins with a goal in mind, such as understanding student dropout so that it might be mitigated or addressed (EUBANKS, EVERS JR. and SMITH, 2016).

Therefore, the prediction is one of the most used techniques in EDM. Given this, it is important to know and understand what has been researched in this area, which are the techniques and methods used, as well as the results obtained. There are some literature reviews in EDM or LA, which we describe below.

Papamitsiou and Economides (PAPAMITSIOU and ECONOMIDES, 2014) conducted a systematic literature survey of LA and EDM. They examined case studies conducted between 2008 and 2013. The work identified 209 studies, but it was limited to forty key studies. Ten of these selected studies were about the prediction of dropout and retention, and the other ten studies were about the prediction of performance. The issue of motivating engagement in learning activities and consequently increasing students’ satisfaction and retention was explored for only four papers, all of them about the prediction of dropout and retention.

Dejaeger *et al*. (2012) used a questionnaire that aggregated constructs in order to capture the concepts of perceived trainer performance (PTP), perceived training efficiency (PTE), perceived ease of learning (PEL), and perceived usefulness of training (PUT). They also used data of Teacher Assistant Evaluation (TAE), containing teacher’s information. They applied at the end of the course and the best accuracy rate was 75% and the best AUC rate was 0.908.

Giesbers et al. (2013) investigated the relationship between available tools used, student motivation using AMS scale, participation, and performance on a final exam in an online course in economics (N = 110). As results, motivation to accomplish was the only predictor nearing significance in predicting the final exam score.

The study conducted by Guo (2010) based on the student survey results collected from 43 courses in eleven semesters from 2002 to 2007, statistical analysis and NN (Neural Networks) techniques are incorporated for establishing some dynamic models for analyzing and predicting student course satisfaction. Each dataset consists of three items of unsatisfactory (U), neutral (N), and satisfactory (S) rates from student surveys and six items of number of students (NS) enrolled, high distinction (HD), distinction (D), credit (C), pass (P) and fail (F) rates extracted from the course records. As results, he used a neural network to predict the course satisfaction with correlation rate = 0.943.

Guruler *et al*. (2010) conducted a study, using the input data: the registered information for state (registered city, date of birth, etc.), high school information (diploma degree, education type, foreign language, etc.), Turkish university entrance exam degree and university placement information, family's living conditions and financial status. They classified the data using decision trees, in order to predict GPA values. The authors did not inform the sample used.

Ilhantola *et al*. (IHANTOLA, VIHAVAINEN, *et al.*, 2016) provided an overview of the body of knowledge regarding the use of educational data mining and learning analytics focused on programming teaching and learning. They selected 76 papers distributed over the period from 2005 to 2015. Of the papers, 59 (78%) were considered as a case study, where results reported data collected in a natural setting, such as in-class activity. This contrasts with only eleven studies (14%) that were classified as experimental research, in which formalized experimental environments were set up in order to collect the data. Only eight studies (11%) formally referenced and utilized a specific theory or model in the development of a tool and/or the analysis of the data collected. Only fifteen (20%) studies presented work that could be considered longitudinal in nature, where data was collected from students over multiple offerings of a course or courses. Only eight studies (11%) focused on dropout risk and performance, showing approaches for identifying students that are at risk of dropping out of class, as well as measuring performance.

We also found three main literature reviews about student's performance prediction.

Shahiri *et al*. (SHAHIRI, HUSAIN and RASHID, 2015) conduct a systematic review to answer the following goals:

a) To study and identify the gaps in existing prediction methods.
b) To study and identify the variables used in analyzing students' performance.
c) To study the existing prediction methods for predicting students' performance.

The review found thirty studies from 2002 to 2015. Ten of thirty papers have used cumulative grade point average (CGPA) as their main attributes to predict students' performance. Other ten studies use internal assessment, such as assignment mark, quizzes, lab work, class test, and attendance. Students demographics, including gender, age, family background, and disability, are used by nine studies. External assessment is identified as a mark obtained in the final exam for a particular subject and used for nine studies, high school background for four studies, and social interaction network for five studies. There are also several researchers in other studies who have used psychometric factor to predict students' performance in four studies.

A psychometric factor is identified as student interest, study behavior, engagement time, and family support. Several researchers have used these attributes to produce clear, simple and user-friendly systems. It helps the professor to evaluate students' achievement based on their personal interest and behavior. However, these attributes are rarely applied in predicting students' performance by several researchers because it focuses more on qualitative data and it is also hard to get valid data from respondents (SHAHIRI, HUSAIN and RASHID, 2015).

The methods mostly used are Decision Tree (10), Neural Network (8), Naive Bayes (4), Support Vector Machine (3), and K-Nearest Neighbor (3). We grouped the prediction results of studies mentioned by Shahiri, Husain, and Rashid. Table 8 shows that works using psychometric and demographic data as input had worse results.

Table 8: Accuracy results from Shahiri, Husain, and Rashid

	Grade and assessments	**Psychometric and demographic data**
Average	81,14%	67,57%
Median	80,00%	69,00%
Std. Deviation	8,67	9,86

Source: Developed by the author

The accuracy rate was from 50% to 98%. However, all studies with accuracy above 80% used internal or external assessment as attributes. We based on these results to define the research goals. Most studies used data collected after the end of the course or program.

Hellas et al. (HELLAS, IHANTOLA, *et al.*, 2018) presented a systematic literature review in the area of predicting student performance. They selected 171 studies from 2010 until 2018. They identified 29 input features, with performance in the course of interest, engagement in the course of interest, and performance in previous courses standing out as the most common data being used in predictions. The psychometric factors, such as motivation, self-regulation, and self-efficacy are used only in 3.06%, 2.60%, and 1.21%, respectively. However, the authors did not mention the prediction results.

Hellas *et al*. also proposed a checklist with items that each article focusing on predicting student performance should include. This is a minimum requirements list, which essentially outlines the need for explicit and clear methodology and results:

a) Define what is being predicted. If the value describing performance (e.g., course grade) consists of multiple items (e.g., course exam, course assignments), describe the contribution (weight) of each item when the performance value is calculated.
b) Define the factors used for prediction. Describe them in such detail that a reader that is not familiar with that particular context understands them. Provide links to, or if not possible, include the scales and surveys that have been used when collecting data.
c) Define the methodologies used for prediction and link the methods used by referencing appropriate papers. Unless you propose a novel method, formal proofs, etc. are not required. If you use feature selection, include details on them.
d) Define the data. Explain where the data comes from, if it is self-reported or automatically collected, and if students are compensated for participating. Moreover, if the data contains students from a course, discuss the number of students in the course, describe how many were excluded from the analysis and why, and provide descriptive statistics that outline the data. Be specific on whether the data is from a single course or single institution, and also discuss if a separate data set is used for validating the prediction results.
e) Provide the results. Perform and report on the tests necessary to test the required attributes of the data. Name the analyses being performed and report all the relevant statistics to allow for interpretation of the results.
f) Discuss the reasons why specific factors, performance metrics, and methods were chosen (or omitted).
g) Reflect upon the results and consider why the methods and factors used did work or did not work. What are the particular context-specific issues that may influence the outcomes?
h) Describe threats to validity and limitations. Note situations in which a model or approach might be applied as well as where it is not valid.

Anoopkumar and Rahman (ANOOPKUMAR and RAHMAN, 2016) conducted a review of data mining techniques and factors used in Educational Data Mining to predict student improvement. The authors identified 69 papers, of which 25 were related to student performance evaluation methods. The work did not compile the results and did not group the attributes or techniques used, but it briefly describes each study. Only three of 69 selected papers use some psychology attribute to predict student's performance. Table 9 shows a summary of these three main works.

Table 9: Summary of the related works in student prediction

	(SHAHIRI, HUSAIN and RASHID, 2015)	**(ANOOPKUMAR and RAHMAN, 2016)**	**(HELLAS, IHANTOLA, *et al.*, 2018)**
Context	Data mining techniques to predict students' performance	Data mining techniques to predict students' performance	Predicting students' performance in computing, informatics, and engineering
Sample	30	69	171
Period	2002-2015	2005-2015	2010-2018
Input features	student demographics (9) external assessment (9) extra-curricular activities (5) high school background (4) social interaction network (5) psychometric factor (4)	Not informed	29 input features into performance in the course of interest, engagement in the course of interest, performance in previous courses, general demographic data, psychometric factors, and in data extracted from logs.

Shahiri, Husain and Rashid (2015) describe a summary of the results of the prediction accuracy of the studies found. The accuracy was average of 77.75%, being the average accuracy for works that use grades or some specific assessment 81.14% (66% to 97%) and for works that use only general factors (pre-university and psychometrics) 67.57% (50% to 83%).

The work uses only psychometric factors that had the best result of prediction (average of 78.33%) was described by Sembiring *et al*. (2011). Sembiring and colleagues used a questionnaire to collect the real data describes the relationships between behavioral of students (psychometric factors) and their final academic performance and outcome. The variables used in the questionnaire were interest, study behavior, engage time, believe, and family support. The number of students was 1000 with three different majors in a college of computer system and software engineering. The sources of collected data were personal records, students' academic record, and course records. As the dependent variable, they grouped all grades into five groups: excellent, very good, good, average, and poor. They obtained good results using J48 Decision Tree as shown in Table 10.

Table 10: Results obtained by Sembiring *et al.*

Performance Prediction	Training		Testing	
	Best Accuracy (%)	Average Accuracy (%)	Best Accuracy (%)	Average Accuracy (%)
Excellent	100.00	99.67	100.00	92.00
Very Good	100.00	100.00	93.33	75.67
Good	100.00	100.00	73.33	61.00
Average	100.00	99.70	80.00	69.33
Poor	100.00	99.70	96.67	93.67

Fonte: (SEMBIRING, ZARLIS, *et al.*, 2011)

How we found a few works about prediction for introductory programming courses, we conducted a systematic mapping study of literature in order to identify the current studies in this context.

3.2.2 Systematic mapping study of student prediction in introductory computing courses

Therefore, although there are several studies that report case studies of prediction, we found only one review in the context of programming or introductory computing courses (HELLAS, IHANTOLA, *et al.*, 2018). However, this work refers to studies by the year 2015 and does not present compiled results, to which new approaches can be compared. Given that, in order to explore and better understand the existing works, we conduct a mapping study of the outcome prediction of students in introductory programming courses.

We searched into the databases ACM Library and IEEE Explorer. The following criteria were defined: case studies and experiments, papers that demonstrate results, face-to-face education or distance education, published in the last six years (from 2013 to 2019), written in English or Portuguese and published as a full paper. We detail, in Table 11, the structure used to create the search terms.

Table 11: Search terms of mapping study of prediction

	Criteria Details	**Search terms**
Population	Students	student / academic

Intervention	Methods/ techniques for prediction	Predict*
Outcome	performance / success	performance / *success / fail* / risk / achievement / outcome
Context	Programming courses	Programming / cs101 / introductory computer science

Source: Developed by the author

Initially, we found 414 works (IEEE Explorer – 198 and ACM Library – 216). We include also some papers that are not included in the original search bases. These papers are from Google Scholar search and SBIE (Brazilian Symposium of Informatics in Education) and are marked in bold in Appendix B. With the papers originally found we executed three stages of selection:

a) After analysis of the title, 143 papers were selected;
b) After removing duplicate items, 98 papers left;
c) After analysis of the content, 45 valid papers remained.

Table 12 shows a summary of papers by year of publication. We can notice that success prediction of students' outcome in programming courses is a subject that has been maintaining a constant interest in the scientific community in recent years.

Table 12: Number of selected papers by year

Year	Papers
2013	3
2014	6
2015	8
2016	7
2017	8
2018	11
2019	2

Source: Developed by the author

The full list of selected papers can be seen in Appendix B. We seek to answer three analysis questions in this mapping study, which are detailed below.

AQ1 – What are the main attributes used to predict student outcome?

In order to facilitate the analysis, we have grouped the input attributes for the prediction, as found in the selected works, into nine sub-groups and three groups: performance-related attributes (course/program/tasks), LMS and programming toll interaction, and psychometric and demographic attributes. Psychometric and demographic attributes are not specific to a particular course and are independent of a specific assessment tool.

Table 13: Input features used in selected works

Performance-related	Papers	Psychometric and demographic	Papers	Interaction	Papers
Grade or assessments	16	Demographic data	14	LMS interaction	13
Activities performance	8	Pre-university data	14	Programming behavior	4
		Psychometric factors	7		
		Learning style and skills	6		
		Study behavior	2		

Source: Developed by the author

As shown in Table 13, we realized that the main attributes used for prediction are related to performance (grades/assessment), demographic data, pre-university data and interaction with Learning Management System (LMS). However, what is the relationship between these attributes and the results? Below we detail the results in accordance with each group of attributes used.

AQ2 – What is the performance of the prediction?

We consider 31 of the 44 selected studies that demonstrate some metric of prediction performance (the others indicate only the index of correlation between the attributes and the variable being predicted). Of these 31 studies, the average accuracy is of 76.67%, being between 48% and 93.3%. In Table 14, the results are shown according to the attribute type used.

Table 14: Accuracy according to the attribute type

Group	Attributes	Accuracy Average	Papers
Performance	Grade, assessments, or activities performance	79.82%	19
Interaction	LMS interaction or programming behavior (interaction with IDE)	76.79%	8
General (demographic and psychometric)	Demographic data, Pre-university data, psychometric factors, learning style, soft skills, or study behavior	75.96% (68.80% to 86%)	4

Source: Developed by the author

Although some works use more than one attribute group, we consider only the priority attributes group, following the order of the attributes of greater specificity to the more general: performance → interaction → general. We realize that studies that use attributes based on grades and student interaction with LMS have the best results.

Only four of the papers analyzed (Table 15) use only general attributes, independent of specific activities in the course.

Table 15: Papers found using only general features

Author(s)	Year	Sample	Techniques	Best Result
(NINRUTSIRIKUN, WATANAPA, *et al.*)	2016	85	Artificial Neural Network, Support Vector Machine, and the c1assic Naive Bayes.	78.33%
(AYUB and KARNALIM)	2017	41	J48 (written assessment)	70.73%
(AZIZ and AHMAD)	2014	399	Naïve Bayes, Rule-Based (OneR), Decision Tree (J48)	68.80%
(QUILLE and BERGIN)	2018	692	Naïve Bayes	86.00% (recall)

Source: Developed by the author

Of these studies, the average results (accuracy or recall) in the prediction of the students' outcome is 75.96%, being 86.00% the best result. These results are a little higher than the results described by Shahiri, Husain, and Rashid (2015), but corroborating with the fact that the use of grades and specific assessments of the course increase the prediction accuracy. However, this also hinders the replication of studies and previous prediction, because they often rely on waiting for the first grades and assessments. In addition, these type of input attributes is dependent on the course content and teaching methods.

AQ3 – In which moments data collection for prediction occur?

Only nine studies collected data longitudinally, evaluating the change and evolution of prediction over time. Usually, in these cases, the accuracy increases over time, as shown in Table 16.

Table 16: Weekly prediction of related works

Week	1ª	2ª	3º	4ª	5ª	6ª	End	Features input type
(VIHAVAINEN, 2013)		64%					78%	Interaction, programming behavior
(UMER, SUSNJAK, *et al.*, 2017) (AUC)	82,9%	86,1%	87,2%	87,1%	87,8%	88%	NI	Grade, activities performance
(WATSON and LI, 2014)			only variance is explained				75,6%	Interaction, programming behavior

(DETONI, ARAUJO and CECHINEL, 2014)	35%	48%	52%	57%	64%	68%		LMS interaction
(HUNG, WANG, *et al.*, 2017)	84,1%	84,1%	84,1%	84,1%	84,1%	84,1%	89,8%	Grade, LMS interaction
(PARDO, HAN and ELLIS, 2016) (MSE)		15,8	15,5	15,1	14,5	29,2	32,0	LMS interaction, pre-university
(SANTOS, PITANGUI, *et al.*, 2016)	67,7%	68,1%	70,6%	74,2%	75%	74,6%	84,7%	LMS interaction
(KLOFT, STIEHLER, *et al.*, 2014)	72%	74%	83%	83%	84%	85%	85%	LMS interaction
(MACHADO, CECHINEL and RAMOS, 2018)	10%	50%	60%	75%	85%	75%	95%	LMS interaction (recall) [a]

a. Approximate values, because the results are displayed only in graphic format

Source: Developed by the author

Therefore, the best prediction accuracy considering the first two weeks are 84.14% and 86.16%, respectively. However, both use grade as an input attribute. None of the works found uses psychometric data to make a longitudinal analysis during the course.

This mapping study shows that still exists concern of the community in understanding the factors and prediction mechanisms to identify at-risk students in introductory computing courses. However, despite this, we realized that much of the existing work in this field using information from students' performance on specific tasks or assessments, or in interaction with LMS. Few works seek to use psychometric information and create methods that can easily be replicated in other contexts.

A justification for this is that the best prediction results are from works that use this type of attribute input (performance or interaction). But it is not possible to create a prediction method based on psychometric data, as the motivation? The lack of studies and the low number of evidence leaves us this doubt and gaps, which we try to fill in this work, through the EMMECS method proposed that will be detailed in Chapter 4.

4 EMMECS

In this chapter, we present the method created to identify in advance at-risk students and motivation factors, describing its details. The chapter is divided into five sections. Section 4.1 is an overview of the method. Sections 4.2, 4.3, 4.4, and 4.5 detail each of the instruments used and presents their validation.

4.1 THE METHOD OVERVIEW

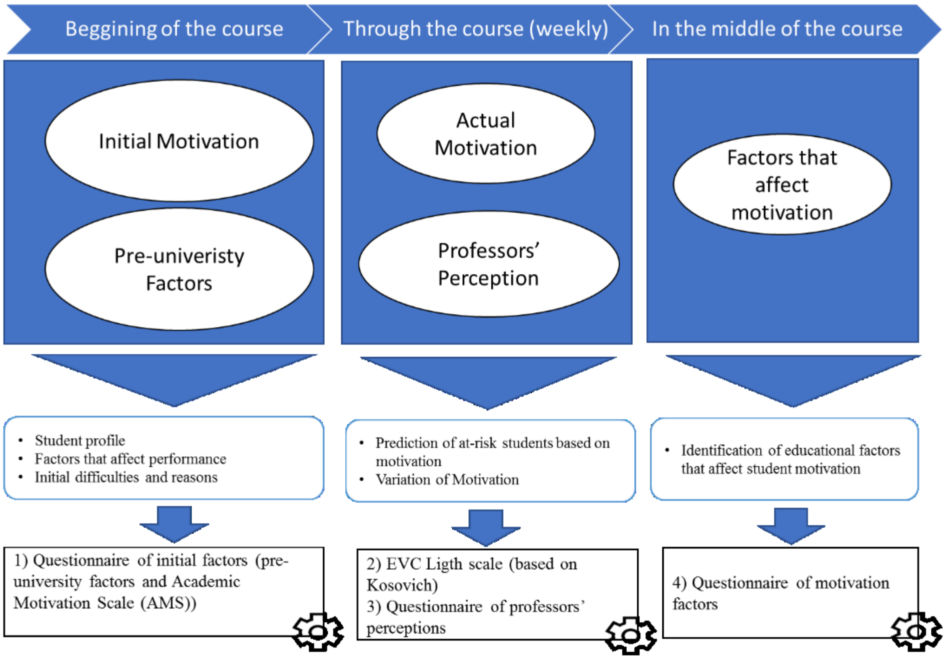
The method named Evaluation Method of Motivation and Engagement of Computing Students (EMMECS) aims to predict at-risk students in introductory programming courses based on motivation and pre-university factors. Another goal is to be able to map out the factors that are affecting the motivation and hence the students' failure.

The method was created to be easy and simple to apply, and it considers the student motivation longitudinally. To do that, the method proposes some instruments in three distinct moments of a computing course:

a) At the beginning of the course, when it is necessary to collect some demographic data and pre-university factors that can impact on motivation, engagement and student outcome. In addition, it must be identified the initial motivation of students, that is, the reasons why they chose to take the course;
b) During the course (weekly), it must be identified the student motivation to the course. To do this, we propose an EVC light-scale, detailed in Section 4.3; During the course (weekly), the professor must point out his/her perception related to each student's interest and motivation;
c) In the middle of the course (once each semester), it aims to identify some factors related to student interaction with the educational environment (professors, course, university environment, etc.) that can impact on students’ motivation and outcome. Data collection using this instrument can be anticipated if the professor wants to identify earlier possible factors related to at-risk students;

Figure 9 shows an overview of the proposed method, containing the process, the information gathered, and the instruments used.

Figure 9: General scheme of the proposed method



Source: Developed by the author

The following sections report in detail the proposed four instruments to collect data and conduct the prediction.

4.2 PRE-UNIVERSITY FACTORS AND INITIAL MOTIVATION

Studies demonstrate that there is a large concern with students' initial motivation (JENKINS, 2001) and their performance in the first year (ALLEN, ROBBINS, *et al.*, 2008). However, two things should be discussed more deeply: i) does the initial motivation have a direct and exclusive relationship with student outcome and performance? ii) which are the other possible factors that may influence the initial motivation of the student?

Studies relating previous factors with performance have divergent results. For example, Wilson (BRUINSMA, 2004) found no correlation of performance with previous programming experience, previous non-programming computer experience, and gender, but he found a correlation with math background. However, Rountree, Rountree, and Robins (2002) found a relationship between performance and math background, but they have not found a correlation with knowledge of a programming language.

Similarly, few studies correlate performance with the initial motivation, that is, the reason for attending the course or undergraduate program. Studies that correlate performance with motivation also show divergent results. For example, Lishinski *et al.* (LISHINSKI, YADAV, *et al.*, 2016) found no correlation between intrinsic and extrinsic goal orientation on students'' performance. Rountree, Rountree, and Robins (ROUNTREE, ROUNTREE and ROBINS, 2002) also found no correlation between the factor "how keen they were to take studies" and performance. However, Pasini et al. (PASINI, SOLITRO, *et al.*, 2016) demonstrate that theoretical exam grades were positively correlated with motivational antecedents.

4.2.1 Initial Factors Questionnaire

In order to find out the previous factors that impact on computing students' motivation and performance, we conducted a systematic mapping study about computing students' motivation (Section 3.1), in which we found seventeen previous factors, divided into four categories:

a) Personal/demographic data: gender (MAMONE, 1992) (ATES, 2011);
b) Taste and knowledge of the area and course: broad and social vision of computing (PETERS and PEARS, 2013), taste for technology and programming (PETERS and PEARS, 2013), correct perception of the area and professionals (PAYNE, 2013), and knowledge of computing and undergraduate program goals (PAYNE, 2013);
c) Informatics/programming experience: work experience in the area (KORI, PEDASTE, *et al.*, 2016) (JENKINS, 2001) (MAMONE, 1992), prior knowledge of computing (ZAINAL, SHAHRINA, *et al.*, 2012) (KORI, PEDASTE, *et al.*, 2016), and programming experience in basic education (PETERS and PEARS, 2013);
d) Prior school performance: student ability in computing (SOERJANINGSIH, 2001), educational history (SAYERS, NICELL and HINDS, 2010) (ROBEY, VON KONSKY, *et al.*, 2006) (NIKULA, GOTEL and KASURINEN, 2011) systemic vision, and knowledge about mathematics (PETERS and PEARS, 2013).

Studies also show divergent results about factors that impact on students' motivation. As an example, the "social influence" factor was rated by four studies, and in three of them it had a strong positive impact (LAW, LEE and YU, 2010) (RAO and WANG, 2014) (ZAINAL, SHAHRINA, *et al.*, 2012), but on the other one it had a weak impact (MCCARTNEY, BOUSTEDT, *et al.*, 2016).

Therefore, although there are several studies regarding the previous factors that impact on student performance and motivation in computing courses, there are few studies that measure the correlation between motivation, performance and outcome, and pre-university factors.

We worked out a questionnaire divided into five groups and 25 factors, as shown in Table 17. Each item has options following a five-point Likert scale.

Groups 2, 3, and 4 of the questionnaire were based on a compilation of factors extracted from the literature. Group "reasons to choose computing" is a light scale adapted from Vallerand (1992) and Jenkins (2001). The entire questionnaire, in Portuguese, can be shown in Appendix C.

Table 17: Questionnaire groups and factors

Group	Factor
1. Personal and demographic data	1A – Gender 1B – Quota 1C – Entrance exam position 1D – Year of entrance 1E – Age 1F – Entrance way
2. Taste and knowledge of the area	2A – Taste for programming and technology 2B – Knowledge about the undergraduate program goals 2C – Knowledge about the undergraduate program content 2D – Correct perception about computing professionals
3.Computing and programming experience	3A – Knowledge and experience in computing 3B – Knowledge and experience in computer programming 3C – Programming experience in high school
4. Prior school performance	4A – General educational performance 4B – Prior math performance
5. Reasons to choose computing	Reason1 – Program content Reason2 – Interest for learning Reason3 – Career/job Reason4 – Parents influence Reason5 – Friends influence Reason6 – Challenge to success Reason7 – Challenges of computing area Reason8 – Lack of other options Reason9 – Not having passed another course Reason10 – Family pressure

Source: Developed by the author

In addition to these factors, the instrument includes the AMS scale (VALLERAND, BLAIS, *et al.*, 1989) to measure the motivation of the students. The original version of the scale was developed in Canadian French (L'échelle of motivation in Éducation-EME) by (VALLERAND, BLAIS, *et al.*, 1989) assuming the multifactored motivational processes. It is composed of 28 items and punctuated on a Likert scale with seven points. EME was subsequently translated into English, originating the AMS (Academic Motivation Scale) (VALLERAND, 1992). A Portuguese version was translated and validated by (SOBRAL, 2003). The entire AMS questionnaire used, in Portuguese, can be shown in Annex A and it is detailed in Section 4.2.

The theoretical or structural factorial model of EME and AMS presents the intrinsic motivation in the form of the subscales: intrinsic motivation - knowledge (IMK), intrinsic motivation - accomplishment (IMA) and intrinsic motivation - stimulation (IMS). In addition to the subscale of amotivation (AMO), the model also contains three other subscales that group the various forms of extrinsic motivation: identify regulation (EMId), introjected regulation (EMIn) and external regulation (EMER) (VALLERAND, 1992).

To identify the type of motivation of students, we use statistical tests to verify the relationship between the variables and the type of motivation, according to the AMS scale. To measure the intensity of motivation in general, facilitating the interpretation and analysis of the results, we calculated a motivation index based on the AMS subscales, grouped in intrinsic motivation, extrinsic motivation, and amotivation (see Equation 1).

$$MI = (\Sigma IMK + \Sigma IMA + \Sigma IMS)/3 + (\Sigma ER + \Sigma IR + \Sigma IdR)/3 - \Sigma AM \quad (1)$$

Where:

MI = Motivation index

ΣIMK = Sum of intrinsic motivation – knowledge

ΣIMA = Sum of intrinsic motivation – accomplishment

ΣIMS = Sum of intrinsic motivation – stimulation

ΣER = Sum of external regulation

ΣIR = Sum of introjected regulation

ΣIdR = Sum of identified regulation

ΣAM = Sum of amotivation

4.2.2 Questionnaire Evaluation

We conducted two surveys to evaluate the questionnaire: i) a pilot test in one university for initial validation and ii) a second survey in ten universities.

The pilot was based on the application of a questionnaire (survey) to 64 freshmen students of the Bachelor Program on Software Engineering at Santa Catarina State University (UDESC) - Brazil. The survey was applied during the second semester of 2016 and the first semester of 2017. The development and implementation of this survey were based on the process described by Kasunic (KASUNIC, 2005).

To assess the level of internal consistency of the questionnaire, we calculated the Cronbach's Alpha that resulted in 0.768, which can be considered satisfactory. "*There are* *different reports about the acceptable values of alpha, ranging from 0.70 to 0.95*" (TAVAKOL and DENNICK, 2011).

In the second survey, 159 students from ten universities participated in the research, being five public universities and five private universities, belonging to the following programs: Computer Science, Software Engineering, Information Systems, Analysis in Development of Systems, Systems for the Internet, and Information and Communication Technology. More than 80% of the respondents belong to five universities, according to Table 18.

Table 18: Respondents by university

University	Type	Students	%	Program
University A	Public	39	24.5%	Software Engineering
University B	Public	28	17.6%	Information and Communication Technology
University C	Public	27	17.0%	Computer Science
University D	Private	26	16.4%	Information Systems
University E	Public	18	11.3%	Computer Science
University F	Private	6	3.8%	Information Systems
University G	Public	5	3.1%	Systems for the Internet
University H	Private	5	3.1%	Information Systems
University I	Private	3	1.9%	Computer Science
University J	Private	2	1.3%	Analysis in Development of Systems

Source: Developed by the author

The reliability of the questionnaire is measured by Cronbach's alpha coefficient, which measures the internal consistency. As a result, we found that the questionnaire can be considered reliable (Cronbach's alpha of pre-university factors = 0.7816, Cronbach's alpha of AMS = 0.9522 and Cronbach's alpha of whole questionnaire = 0.9421).

4.2.3 Results and Correlations

We identified that, in general, the students had greater motivation in the subscales: external regulation, introjected regulation, and intrinsic motivation – accomplishment, according to Table 19.

Table 19. Motivation by subscale

Subscale	Mean	SD
Amotivation	5.534	3.086
External regulation	15.654	4.519
Introjected regulation	15.622	4.252
Identified regulation	12.377	4.907
Intrinsic motivation - knowledge	11.251	4.208

Intrinsic motivation - accomplishment	15.691	4.093
Intrinsic motivation - stimulation	13.037	4.486

Source: Developed by the author

In order to understand the possible factors that impact the initial motivation and also the performance of students, we define some questions to be answered, with relation to the pre-university and initial motivation factors.

AQ1 – Personal data (gender and age) and the way of entering the university are related to the initial motivation of the students?

Of the respondents, 86% were male, and 14% female. Most of the respondents entered by university exam (*vestibular*) (34%) and national high school exam (ENEM) (57.9%). The average age of the respondents was 21.9 years. Little more than half (50.9%) was classified in the first call and 29.5% were selected by quota, primarily by public school quota or social quota (24.5%).

In order to answer the first analysis question (AQ1), we examined the correlation between the demographic data (gender and age), student input data (way of entering, quota, and entrance exam position), and each subscale of the initial motivation.

Table 20 presents the analysis according to gender. Although we perceived a greater level of motivation for men, we could not statistically prove this variation, both in the analysis of the motivation index and subscales.

Table 20: Results according to gender

	Female (n=22)		**Male (n=137)**		**p-value**
	Avg	*t*	*Avg*	*t*	
Motivation Index	19.969	9.688	22.725	7.404	0.124
Amotivation	5.863	4.246	5.481	2.875	0.592
External regulation	15.000	4.265	15.759	4.565	0.466
Introjected regulation	12.045	5.075	12.430	4.897	0.734
Identified regulation	14.363	5.350	15.824	4.036	0.135
Intrinsic motivation - knowledge	14.363	4.884	15.905	3.931	0.101
Intrinsic motivation - accomplishment	11.954	5.028	13.211	4.388	0.224
Intrinsic motivation - stimulation	9.772	4.888	11.489	4.058	0.076

Source: Developed by the author

We checked the correlation between age (how many years old are each student) and each motivation subscale (Table 21). According to (COHEN, 1988), a correlation is considered significant when the absolute value of the correlation is greater than 0.29. Thus, we found no correlation between age and initial motivation.

Table 21: Correlation between age and motivation subscales

Subscale	Age
Motivation Index	-0.059
Amotivation	-0.144
External regulation	-0.222
Introjected regulation	-0.174
Identified regulation	-0.172
Intrinsic motivation - knowledge	-0.032
Intrinsic motivation - accomplishment	-0.022
Intrinsic motivation - stimulation	0.020

Source: Developed by the author

To fill the program vacancies, the university offers admission accordingly to the candidates' admittance grade. We call this process as entrance exam position. Initially, some candidates are passed considering the classification order. If at least one of those candidates does not enroll, the university makes a new admission offer for each vacancy remaining. This is repeated until all vacancies are filled, or when the waiting list is empty.

In relation to the entrance exam position, although the students who entered in the first call had higher levels of motivation, we could not statistically prove this variance, except for the subscale "Intrinsic motivation – stimulation" (Table 22).

Table 22: Motivation based on entrance exam positions

	First (n=81)	Second (n=48)	Other calls (n=24)	Does not apply (n=6)	p-value
Motivation Index	23.427	21.576	20.902	19.611	0.310
Amotivation	5.605	5.458	5.542	5.167	0.985
External regulation	16.247	14.917	15.042	16.000	0.372
Introjected regulation	12.753	11.875	12.500	10.833	0.666
Identified regulation	16.395	14.854	15.083	13.500	0.105
Intrinsic motivation - knowledge	16.358	15.458	13.875	15.833	0.069
Intrinsic motivation - accomplishment	13.568	12.833	12.125	11.167	0.355
Intrinsic motivation - stimulation	11.778	11.167	10.708	7.000	0.049

Source: Developed by the author

In the studied programs there were basically two ways of entering: *vestibular* or university entrance exam, a test applied by each institution that allows access to a number of students sorted by their final grade, and the National High School Exam (ENEM), applied throughout the national territory by the Brazilian Federal Government to those that are concluding high school. This exam allows the enrollment of other candidates sorted by their final grade. The exceptions are classified in this work as other ways of entering (profile analysis, high school grades, transfer, etc.).

In relation to the way of entering, we identified that in general students who entered through the vestibular had higher rates of motivation, while students who did not enter via *vestibular* or ENEM had a lower motivational rate. However, we found significant variance only for the subscales "Introjected regulation" and "Intrinsic motivation – stimulation", as shown in Table 23.

Table 23: Motivation based on the way of entrance

	ENEM (n=92)	Vestibular (n=54)	Other (n=13)	p-value
Motivation Index	22.449	23.080	17.666	0.431
Amotivation	5.760	5.500	4.076	0.183
External regulation	15.652	16.129	13.692	0.219
Introjected regulation	12.402	13.388	8.000	0.0015
Identified regulation	15.695	16.037	13.384	0.126
Intrinsic motivation - knowledge	16.173	15.185	14.384	0.181
Intrinsic motivation - accomplishment	13.239	13.407	10.076	0.0437
Intrinsic motivation - stimulation	15.652	16.129	13.692	0.219

Source: Developed by the author

In Brazilian public universities, there are some reserved vacancies or quotas to specific social and ethnic groups. With respect to quotas, we found no significant variance, as shown in Table 24.

Table 24: Motivation based on quota

	No (n=112)		Yes (n=47)		p-value
	Avg.	*t*	*Avg.*	*t*	
Motivation Index	2.321	7.548	22.397	8.402	0.956
Amotivation	5.60	3.002	5.361	3.306	0.649
External regulation	15.785	4.317	15.340	5.005	0.572
Introjected regulation	12.232	4.924	12.723	4.902	0.566
Identified regulation	15.625	4.042	15.617	4.761	0.991
Intrinsic motivation - knowledge	15.803	4.037	15.425	4.256	0.597
Intrinsic motivation - accomplishment	13.098	4.236	12.893	5.078	0.794
Intrinsic motivation - stimulation	11.241	4.206	11.276	4.256	0.961

Source: Developed by the author

AQ2 – Is the taste for technology related to the initial motivation of first-year students in computing?

The students answered about how much they like games, hardware, and programming. Most students have confirmed that they like and use games (86.8%), like hardware and equipment (72.3%), and like programming (65.4%). In general, the more intense the taste of the three areas, the higher the rates of motivation. Table 25 shows a summary of the analysis of variance, identifying in bold the significant variances.

Table 25: Motivation based on taste for computing

	Games	Hardware	Programming
Motivation Index	0.0403	0.0003	0.0015
Amotivation	0.5860	0.3450	0.9370
External regulation	0.0225	0.1500	0.2690
Introjected regulation	0.2570	0.0539	0.0985
Identified regulation	0.0004	0.0003	0.0082
Intrinsic motivation - knowledge	0.0067	0.0041	0.0001
Intrinsic motivation - accomplishment	0.1080	0.0143	0.0005
Intrinsic motivation - stimulation	0.0242	0.0008	0.0002

Source: Developed by the author

AQ3 – Is knowledge of the course and area related to the initial motivation of first-year students in computing?

Regarding the knowledge about the program and the area, the students answered about their perception of knowledge about: the objectives and differentials of the program, the content of the program (curriculum), and the job market (Table 26).

Table 26: Motivation based on knowledge about the program

	Course objectives and differentials	Course contents	Job market
Motivation Index	0.085	0.003	0.001
Amotivation	0.022	0.425	0.165
External regulation	0.775	0.360	0.273
Introjected regulation	0.345	0.003	0.028
Identified regulation	0.103	0.022	0.005
Intrinsic motivation - knowledge	0.070	0.018	0.013
Intrinsic motivation - accomplishment	0.008	0.000	0.000
Intrinsic motivation - stimulation	0.775	0.0135	0.003

Source: Developed by the author

AQ4 – Is previous experience in computing related to the initial motivation of first-year students in computing?

We analyzed the relationship between the experience and prior knowledge of the student in computer science and general computing, programming, and programming in high school with the initial motivation (Table 27).

Table 27: Motivation based on prior experience

	Experience in computing	Programming in high school	Experience in programming
Motivation Index	0.080	0.218	0.001
Amotivation	0.124	0.200	0.009
External regulation	0.126	0.293	0.089
Introjected regulation	0.135	0.082	0.043

Identified regulation	0.128	0.512	0.004
Intrinsic motivation - knowledge	0.015	0.335	0.016
Intrinsic motivation - accomplishment	0.023	0.573	0.011
Intrinsic motivation - stimulation	0.121	0.355	0.044

Source: Developed by the author

We found significant variances mainly in “experience in programming”. An interesting fact is that students who report having a lot of experience with software development have fewer levels of initial motivation than others (Table 28).

Table 28: Motivation based on programming experience

	1 (n=64)	2 (n=37)	3 (n=31)	4 (n=15)	5 (n=12)	p-value
Motivation Index	21.677	21.360	25.290	26.622	15.972	0.001
Amotivation	5.703	4.648	5.483	4.933	8.250	0.009
External regulation	15.953	14.351	16.387	17.400	14.000	0.089
Introjected regulation	12.031	11.702	14.161	13.866	9.833	0.043
Identified regulation	15.765	14.432	17.000	17.466	12.666	0.004
Intrinsic motivation - knowledge	15.218	14.594	17.354	17.533	15.000	0.015
Intrinsic motivation - accomplishment	12.203	12.432	14.741	15.533	11.833	0.010
Intrinsic motivation - stimulation	10.968	10.513	12.677	12.866	9.333	0.044

Source: Developed by the author

AQ5 – Is the previous school performance related to the initial motivation of first-year students in computing?

Students evaluated their perception of their general performance in high school and their performance in Math. Table 29 shows the relationship between these factors and the initial motivation.

Table 29: Motivation based on prior school performance

	Math performance	General performance
Motivation Index	0.178	0.009
Amotivation	0.166	0.508
External regulation	0.75	0.031
Introjected regulation	0.252	0.146
Identified regulation	0.294	0.008
Intrinsic motivation - knowledge	0.044	0.001
Intrinsic motivation - accomplishment	0.056	0.002
Intrinsic motivation - stimulation	0.209	0.092

Source: Developed by the author

In general, students with a better perception of performance have better levels of initial motivation. However, for performance in math, we only found significant variance in subscale "Intrinsic motivation – knowledge". For overall performance, in four of the seven scales, we found a significant variance.

AQ6 – Is the context (university and program) related to the initial motivation of first-year students in computing?

In order to verify whether the context can impact on the initial motivation of the students, we analyzed the variance of motivation for different universities, programs, and type of university (public or private).

Table 30 shows that there was a significant variance only in the "Intrinsic motivation – knowledge" subscale, both in relation to different universities and programs.

Table 30: Motivation based on university and type of program

	University	Program
Motivation Index	0.710	0.100
Amotivation	0.642	0.940
External regulation	0.356	0.085
Introjected regulation	0.836	0.634
Identified regulation	0.315	0.062
Intrinsic motivation - knowledge	0.002	0.000
Intrinsic motivation - accomplishment	0.811	0.245
Intrinsic motivation - stimulation	0.852	0.543

Source: Developed by the author

The difference occurred basically in two public universities and their two respective programs. According to Table 31, University A and University B have students with a lower level of intrinsic motivation – knowledge. So this seems to be a punctual result, not a feature of the whole sample and that can be generalized.

Table 31: Motivation based on university

University	Type	Motivation	SD	n
A	Public	14.71795	4.495312	39
B	Public	12.85714	5.197680	28
C	Public	17.03704	2.192320	27
D	Private	16.69231	3.608537	26
E	Public	16.72222	2.886185	18
F	Private	17.16667	2.228602	6
G	Public	17.60000	3.286335	5
H	Private	17.20000	3.701351	5
I	Private	18.33333	1.527525	3
J	Private	17.00000	2.828427	2

Source: Developed by the author

Analyzing the variance between public and private universities, we found a significant difference in two subscales (external regulation and intrinsic motivation – knowledge), although in most of them the levels of initial motivation of students in public universities were smaller than students in private universities (Table 32).

Table 32: Motivation based on the type of university

Subscale	Private	Public	p-value
Motivation Index	23.746	21.840	0.174
Amotivation	5.476	5.556	0.887
External regulation	16.905	15.205	0.036
Introjected regulation	12.358	12.385	0.975
Identified regulation	16.619	15.265	0.077
Intrinsic motivation - knowledge	16.952	15.240	0.020
Intrinsic motivation - accomplishment	13.048	13.034	0.987
Intrinsic motivation - stimulation	11.786	11.060	0.339

Source: Developed by the author

We can observe in Table 33 that, in addition to the intrinsic reasons (interesting in learning, taste for the area), the prospect of a professional future has significant importance for students to choose the computer programs.

Table 33: Answers per reasons to attend

Reasons to attend	Negative	Neutral	Positive
Interest in learning more about computing	5.0%	15.7%	79.2%
To like and feel that the computing area is interesting	6.3%	18.9%	74.8%
Challenges that computing activities promote	25.8%	27.0%	47.2%
Challenge to get through and succeed in the course. Professional challenge	27.7%	24.5%	47.8%
Parental influence	72.3%	19.5%	8.2%
Influence of friends, teachers or other people	60.4%	20.8%	18.9%
Family or social pressure to attend a higher education program	77.4%	11.3%	11.3%
Career and employment perspective	4.4%	13.2%	82.4%
To be the one or one of the few program options for the current reality	65.4%	20.8%	13.8%
Have not passed in another course that was desired	83.0%	7.5%	9.4%

Source: Developed by the author

Comparing the motivational index with the reasons for choosing the course, we identified that students who had higher levels of initial motivation had the following reasons: "Interest in learning more about computing" (p-value=8.91×10^{-9}), "To like and feel that the computing area is interesting" (p-value=8.05×10^{-8}), "Challenges that computing activities promote" (p-value=1.57×10^{-8}), "Personal challenge in getting through and achieving success in the course" (p-value=2.61×10^{-10}) and "Career and employment perspective" (p-value=3.83×10^{-8}).

4.2.4 Conclusions

This section we evaluated a questionnaire to identify pre-university factors and initial motivation and we also evaluated the impact of these attributes in a class of first-year computing students, divided into five groups: personal and demographic data, taste and knowledge of the area, computing and programming experience, prior school performance, and initial motivation.

The motivation was measured using the AMS (Academic Motivation Scale), which assesses and divides the motivation into 3 groups and 7 subscales: amotivation, intrinsic motivation (knowledge, accomplishment, and stimulation) and extrinsic motivation (external regulation, introjected regulation, and identified regulation).

As a result, we have identified that the demographic data and the way of entering the university had no significant relation with the motivation rate. Besides these, we found no significant variation of initial motivation according to the knowledge of the objectives and differentials of the program, previous experience in computing, experience with computing at school, and performance in math.

On the other hand, we found a significant variation of the initial motivation according to taste for the area (games, hardware, and software development), knowledge of the program content, knowledge of the area's job market, previous experience in programming, and general school performance. It was interesting to identify that students with greater programming experience have lower levels of initial motivation.

We also identified that the main reasons for students to choose computing programs are related to the perspective of employment and career, interest in learning and taste for the area. Regarding the motivation subscales, we have identified that students had a higher level of motivation in external regulation, introjected regulation, and intrinsic motivation – accomplishment. The lowest level of motivation was in intrinsic motivation – knowledge.

These results indicate that, for the sample used, despite the taste and knowledge of the area is the group with the greatest impact on the initial motivation, the first-year students in computing are widely extrinsically motivated, with emphasis on the concern with career and professional future. On the other hand, students who already have more experience in computing are less motivated, which may indicate a need for professional qualification for people already positioned in the job market, but they do not have the same initial excitement of the others.

4.3 EVC LIGHT SCALE

Related studies that measure the motivation or consider some motivational aspect to predict performance or dropout perform this measurement at a specific time (GRAY, MCGUINNESS and OWENDE, 2014) (HIDAYAH, PERMANASARI and RATWASTUTI, 2013). However, can motivation change over time? To confirm this hypothesis, it is important to identify this variation and its impact on the student's performance or outcome. However, we did not find studies that consider the change in motivation over time as a factor for predicting students' performance or outcome.

In this context, we developed a light instrument to measure the student's motivation over time computing courses based on the practical EVC scale proposed by Kosovich *et al*. (2014). We also evaluate the relation of motivation with the student's performance. To achieve this goal, we defined six analysis questions (AQ) as follows:

- AQ1: Is there evidence for internal consistency of the questionnaire?
- AQ2: Is there evidence of the convergent and discriminant validity of the questionnaire?
- AQ3: How do underlying factors influence the responses to the items of the questionnaire?
- AQ4: Is there a relationship between motivation index and student performance?
- AQ5: Is the variation of motivation over time relevant?

Below we describe how the instrument was created and validated, as well as details of the instrument and results of its application.

4.3.1 Research Method

We describe the validation of an instrument to measure motivation in computing students across the introductory programming course. The instrument was built based on (KOSOVICH, HULLEMAN, *et al.*, 2014).

In order to perform an evaluation of the questionnaire, we conducted a case study as follows:

Preparation: definition of the study goals.

Execution: i) apply the instrument to computing students; ii) collect and organize data from case studies.

Analysis: i) internal consistency reliability (omega coefficient); ii) convergent and discriminant validity (intercorrelation of the scale items and item-total correlation); iii) factorial validity (factor analysis).

In the preparation phase, we defined the study goal to analyze the questionnaire in order to evaluate the reliability and construct validity from the viewpoint of students in the context of computing programs. We decomposed the goal into analysis questions to be assessed based on the data collected in the case studies.

Several studies in the literature that suggest different sample size to factorial analysis, from 50 to 1000 (SAPNAS and ZELLER, 2002) (COMREY, 1992) (HAIR , ANDERSON , *et al.*, 2010). For this work, we based on (HAIR , ANDERSON , *et al.*, 2010), that suggests the minimum sample size should be 100 to reach acceptable effect size of 0.29 as defined by (COHEN, 1988).

The execution phase was based on the application of a questionnaire (survey) to 245 students of four different universities from Brazil. The development and implementation of this survey were based on the process described by Kasunic (2005).

The analysis phase aims to measure reliability and validity of the instrument. Reliability is the agreement between two efforts to measure the same trait through maximally similar methods. Validity is represented in the agreement between two attempts to measure the same trait through maximally different methods (CAMPBELL and FISKE, 1959).

About the reliability, according to (DUNN, BAGULEY and BRUNSDEN, 2014), if the scale is multidimensional or demonstrates some divergence from unidimensionality, it is recommended to use a coefficient that allows that the scale is split into subscales, each as omega coefficient.

Overall, the main advantages of omega over alpha can be summarized as follows (DUNN, BAGULEY and BRUNSDEN, 2014):

a) Omega makes fewer and more realistic assumptions than alpha.
b) Problems associated with inflation and attenuation of internal consistency estimation are far less likely.
c) Employing "omega if item deleted" in a sample is more likely to reflect the true population estimates of reliability through the removal of a certain scale item.
d) The calculation of omega alongside a confidence interval reflects much closer the variability in the estimation process, providing a more accurate degree of confidence in the consistency of the administration of a scale.

In terms of construct validity, convergent and discriminant validity are the two subcategories of construct validity (TROCHIM and DONNELLY, 2008). “Convergent validity refers to the extent to which different methods of measuring the same trait yield similar results” (CARMINES and ZELLER, 1979). In contrast, discriminant validity refers to the extent to which similar or identical methods measuring different traits lead to different results (CARMINES and ZELLER, 1979). To analyze the convergent and discriminant validity of the questionnaire, we calculated the intercorrelations of the items and item-total correlation. Intercorrelation refers to the degree of correlation between the items (CARMINES and ZELLER, 1979) (DEVILLES, 2003). The higher the correlations among items that measure the same factor, the higher the validity of individual items and, hence, the validity of the instrument as a whole. Item-total correlation is analyzed in order to check if any item in the questionnaire is inconsistent with the averaged correlation of the others, and thus, can be discarded (CARMINES and ZELLER, 1979) (DEVILLES, 2003).

In addition, we used factor analysis (FA) to determinate how many factors underlie the set of items of the questionnaire. FA consists of a variety of statistical methods for discovering clusters of interrelated variables. Those items that are more highly correlated with each other than with the other items define each factor. When the factor loading is higher, the particular item contributes more to the given factor (CARMINES and ZELLER, 1979).

To evaluate the quality of the instrument we defined the following analysis questions (AQ):

Reliability

AQ1: Is there evidence for internal consistency of the questionnaire?

We measured the internal consistency of the questionnaire by calculating the omega coefficient, which measures the reliability of the multidimensional questionnaire.

Construct Validity

AQ2: Is there evidence of the convergent and discriminant validity of the questionnaire?

To establish evidence of the convergent and discriminant validity of the items of the questionnaire, we calculated the intercorrelations of the items and correlation item-total.

Intercorrelations of the items. The first quality we seek in a set of scale items is that they should be highly intercorrelated. One way to determine how intercorrelated the items are is to inspect the correlation matrix (DEVILLES, 2003).

Item-total correlation. This test aims to evaluate the correlation of each item with all the other items. Each item of the instrument is considered consistency if it has a medium or high correlation with all the other items (DEVILLES, 2003). On the other hand, a low item-total correlation of an item undermines the validity of the scale, and, therefore, should be eliminated. For this analysis, we used the method of corrected item-total correlation, which compares one item with every other on the instrument, excluding itself.

In general, the correlations are medium to high considering reference values as defined by (COHEN, 1988), considering a correlation satisfactorily, if the correlation coefficient is greater than 0.29.

AQ3: How do underlying factors influence the responses on the items of the questionnaire?

Prior to the extraction of the factors, several tests should be used to assess the suitability of the respondent data for factor analysis. These tests include Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy, and Bartlett's Test of Sphericity (WILLIAMS, ONSMAN and BROWN, 2010). The KMO index measures the sampling adequacy with values between 0 and 1. An index value near 1.0 supports a factor analysis and anything less than 0.5 is probably not amenable to use factor analysis (BROWN, 2015). Bartlett's Test of Sphericity should be significant ($p < 0.05$) for factor analysis to be suitable (WILLIAMS, ONSMAN and BROWN, 2010).

To obtain the number of factors retained in the analysis, we used the Kaiser-Guttman criterion, because it is the most commonly used method. This method states that the number of factors is equal to the number of eigenvalues greater than 1 (BROWN, 2015). The eigenvalue refers to the value of the variance of all the items, which is explained by a factor (TROCHIM and DONNELLY, 2008).

We also used the Parallel Analysis. In a parallel analysis, actual eigenvalues are compared with random order eigenvalues. Factors are retained when actual eigenvalues surpass randomly ordered eigenvalues (WILLIAMS, ONSMAN and BROWN, 2010).

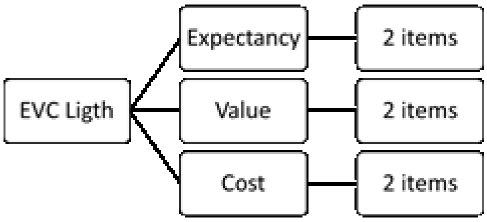
After identifying the number of underlying factors, we determined which items are loaded into which factor. In order to identify the factor loadings of the items, we used the Varimax with Kaiser Normalization rotation method, because it is the most widely accepted (COHEN, 1988).

4.3.2 The Instrument

To create the proposed motivation scale, we used a questionnaire based on the EVC practical scale proposed by (KOSOVICH, HULLEMAN, *et al.*, 2014). We used this scale because it is simpler and smaller, getting easier the longitudinal evaluation over time in the course.

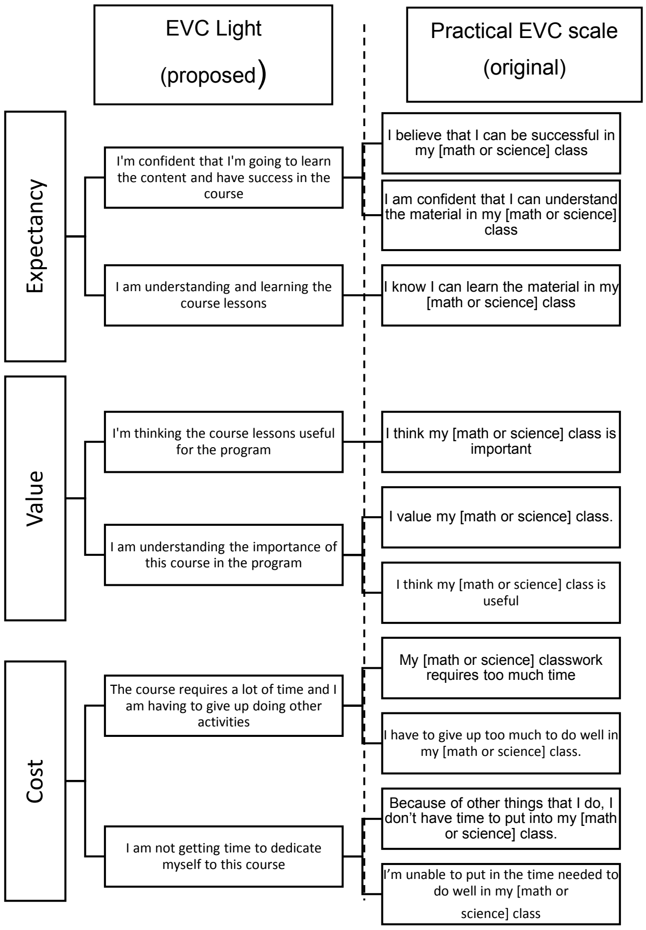
The instrument proposed is composed of three factors: expectancy, value, and cost. Each factor consists of two items, as shown in Figure 10.

Figure 10: Structure of EVC Light proposed



Source: Developed by the author

Figure 11: Relation between original practical EVC scale and EVC Light



Source: Developed by the author

We wrote the questionnaire in Portuguese (Appendix D) and we used the items described below, which are classified into the factors: expectancy (item 1 and 4), value (item 2 and 5), and cost (item 3 and 6):

1. I'm confident that I'm going to learn the content and have success in the course;
2. I'm thinking the course lessons are useful for the program;

3. The course requires a lot of time and I am having to give up doing other activities;
4. I am understanding and learning the course lessons;
5. I am understanding the importance of this course in the program;
6. I am not getting time to dedicate myself to this course.

Each question has five response options, on a 5-points Likert scale from 1 to 5. Based on the questionnaire answers, we define some indexes of motivation for each factor analyzed, as in (1), (2), (3), and (4). These indexes are converted to values between 0 and 1:

$$EI = \frac{(q1 + q4 - 2)}{12} \quad (1)$$

$$VI = \frac{(q2 + q5 - 2)}{12} \quad (2)$$

$$CI = \frac{(q3 + q6 - 2)}{12} \quad (3)$$

$$EVC\ Index = \frac{(EI + VI - CI) + 1}{3} \quad (4)$$

In the equations, *q*1, *q2*, *q3*, *q4*, q5, and *q6* are the answer values for each question, *EI* is the Expectancy Index, *VI* is the Value Index, and *CI* is the Cost Index.

4.3.3 The Questionnaire Evaluation

With the objective of evaluating the impact of motivation measured by the instrument to the students' outcome, we defined the following analysis questions:

Motivation Analysis

AQ4 – Is motivation index related to students' outcome?

AQ5 – Is the variation of student motivation over time relevant?

The survey involved 245 undergraduate students during the first and second semester of 2018, as shown in Table 34. In total, students from four different universities participated in the research (two public universities and two private universities), belonging to following programs: Software Engineering, Information Systems, Computer Science, Internet Systems, and Technologies of Information and Communication.

Part of the questionnaires was applied online, using Google Forms tool. Another part of the questionnaires was applied in paper, mainly on the first application, when in most cases the author monitored and coordinated the survey to clear any doubts from students. In, general, each application took about five minutes. In other cases, the course professor applied the questionnaires.

Table 34: Participants by program and university

University	Course [A]	Period	Students
A	Introductory Programming	20181	53
A	Programming I	20181	21
A	Programming II	20181	38
A	Discrete Mathematics	20181	43
A	Introductory Programming	20182	47
A	Programming I	20182	17
B	Algorithms	20182	30
C	Algorithms and Programming	20182	16
D	Algorithms and Programming Techniques	20182	24

[a.] The courses “Introductory Programming”, “Algorithms”, “Algorithms and Programming” and “Algorithms and Programming Techniques” are distinct names but similar to CS01.

Source: Developed by the author

We conducted the survey during the first and second semester of 2018. In the first semester, we conducted a pilot in four classes from one of the universities. In the second semester, we expanded the research to other scenarios (courses, programs, and universities).

4.3.3.1 Reliability

AQ1: Is there evidence for internal consistency of the questionnaire?

The omega total coefficient was satisfactory (0.90), considering the three original factors.

4.3.3.2 Construct Validity

AQ2: Is there evidence of the convergent and discriminant validity of the questionnaire?

According to Table 35, all correlations had positive values into the same dimension and considered acceptable (above 0.29).

Table 35: Intercorrelations of the items

Item		it1	it4	it2	it5	it3
Expectancy	it1	-				
	it4	0.737	-			
Value	it2	0.444	0.422	-		
	it5	0.367	0.436	0.688	-	
Cost	it3	-0.227	-0.285	-0.034	-0.070	-
	it6	-0.247	-0.294	-0.067	-0.071	0.677

Source: developed by the author

Table 36 shows that items three and six, related to cost, had worse values. To evaluate the positive dimensions (Expectancy and value), we calculated the correlation between these two factors. Table 35 shows that all correlations are strong. This clearly shows us there are two distinct dimensions in the questionnaire.

The degree of correlation between the items determines the degree of convergent and discriminant validity. However, these results do not determine how many factors underlie the set of the questionnaire. With this objective, we performed a factor analysis, answering the analysis question AQ3.

Table 36: The item-total correlation for all items

Item	Total
it1	0.2753464
it2	0.4380430
it3	0.1295154
it4	0.2485915
it5	0.3973791
it6	0.0755557

Source: developed by the author

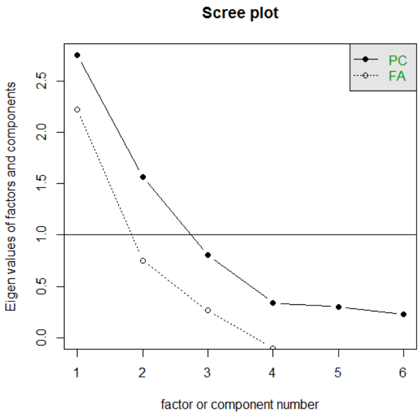
AQ3: How do underlying factors influence the responses to the items of the questionnaire?

In order to identify the number of factors (quality factors or dimensions) that represents the responses of the set of the 6 items of the questionnaire, we performed a factor analysis. Based on the original definition of the questionnaire, we assume that it is influenced by 3 groups of factors (expectancy, value, and cost).

Analyzing the set of items of the questionnaire, we obtained a KMO index of 0.64, indicating that factor analysis is appropriate in this case.

Following the Kaiser-Guttman criterion, our results show that 2 factors should be retained, explaining 58.10% of the data. The scree plot (Figure 12) shows the eigenvalue for each factor number (representing each item).

Figure 12: Factor analysis of adjusted questionnaire



Source: developed by the author

Table 37 shows the factor loadings of the items associated with the two retained factors. The highest factor loading of each item, indicating to which factor the item is most related, is marked in bold. The Chi-Square statistical test proves that two factors are sufficient (p-value = 3.14e-87).

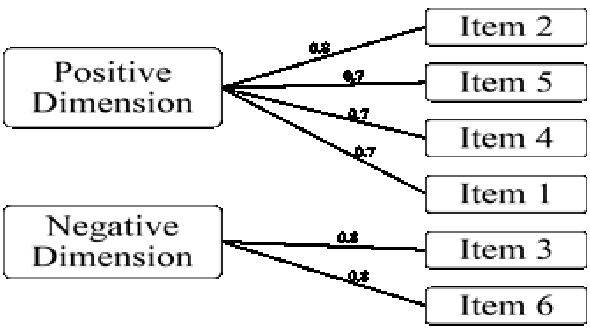
Table 37: The Factor Loadings

Item	**Factor1**	**Factor2**
it1	0.793	-0.199
it2	0.631	
it3		0.818
it4	0.804	-0.227
it5	0.586	
it6	-0.111	0.833

Source: developed by the author

Therefore, in order to group similar factors, we suggest that the best group is joining "Expectation" and "value" in a single dimension (positive values), as shown in Figure 13.

Figure 13: Factor analysis graph of the adjusted questionnaire



Source: developed by the author

4.3.3.3 Motivation analysis

AQ4 – Is there a relationship between student motivation and student performance?

To verify this relation, we analyzed the variance between the average of the motivation indexes calculated according to the final student's status in the course (pass or fail), using the ANOVA (Analysis of Variance) method (Table 38).

These results confirmed the relation between motivation and results, showing that successful students have greater motivation index. As expected, we can perceive a positive relationship between the motivation index and student outcome. In general, as the students' outcome is better, also is better their motivation indexes.

Table 38: The relation between motivation factors and status

Factor	Status	Mean	SD	n	p-value
Expectancy	Pass	0.7999	0.1870	127	0.000011
	Fail	0.6857	0.2271	118	
Value	Pass	0.8603	0.1712	127	0.995
	Fail	0.8601	0.1974	118	
Cost	Pass	0.4501	0.2683	127	0.025
	Fail	0.5324	0.3028	118	
EVC	Pass	0.7366	0.1468	127	0.0007
	Fail	0.6711	0.1529	118	

Source: developed by the author

We also analyzed the variation of the motivation indexes according to the student's final grade level, using the ANOVA method, as shown in Table 39.

We can identify that failure students have significantly lower expectancy indexes and higher perception of the cost. Similarly, as shown in Table 39, students with lower grades also have lower expectancy indexes and higher perceived cost. We were unable to identify the significant variation of the value factor neither with respect to the student's status nor to the level of the grade.

So, we can see that the students' performance had a significant variation compared to the levels of expectancy and cost. However, the value factor had no significant variation compared to the performance.

Table 39: The relation between motivation factors and final grade

Factor	Grade	Mean	SD	n	p-value
Expectancy	A	0.8374	0.1928	41	2.7×10^{-6}
	B	0.7995	0.1849	69	
	C	0.7333	0.1553	30	
	D	0.6762	0.2335	105	
Value	A	0.8821	0.1382	41	0.383
	B	0.8430	0.1947	69	
	C	0.9026	0.1332	30	
	D	0.8508	0.2037	105	
Cost	A	0.3639	0.2519	41	0.00741
	B	0.4794	0.2795	69	
	C	0.5749	0.2575	30	
	D	0.5214	0.3016	105	
EVC	A	0.7851	0.1456	41	0.00028
	B	0.7210	0.1424	69	
	C	0.6888	0.1242	30	
	D	0.6689	0.1583	105	

Source: developed by the author

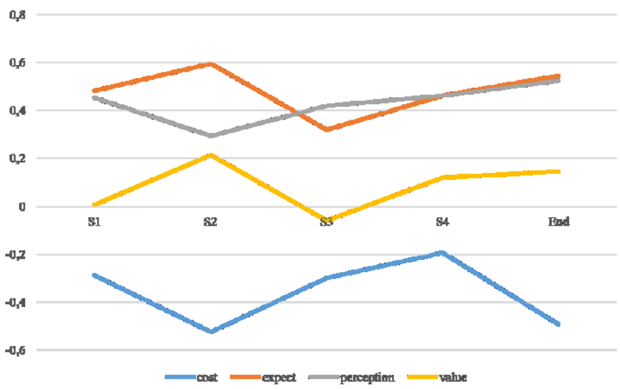
AQ5 – Is there a relationship between the variation of motivation and student performance?

In order to identify the variation of motivation, we analyzed the difference between the motivation indexes each week, using the t-student test for paired samples.

In general, we identified a small variance, negligible. For the few situations with significant variance, there were divergent results. For example, in one week, the index variation is positive, and in another one, it is negative.

To calculate the variance, we analyzed the difference between the subscales index of each week (Figure 15). We found significant variation between week three and week two ($p = 0.046$) and between the end of the course and week four ($p = 0.003$). Both variations were positive. To analyze the variance in student outcome, we compared the results of passed and failed students. We did not find significant differences or any relevant pattern in the variation.

Figure 14: Correlation variation between motivation and final grade



Source: developed by the author

Similarly, we analyzed the evolution of the motivation indexes over the weeks and the variance according to the students' final status (pass or fail). We have identified that there is a significant difference in the subscales "expectancy" and "cost", especially in the first week and after the fourth one. We can perceive, as shown in Table 38, that the expectancy decreases significantly after the fourth week, mainly in the case of students failed. The *value* factor does not change significantly between failed and passed students, but it also decreases after the fourth week. About the *cost* factor, the passed students do not change their perception. However, the failed students decrease their cost perception between the first and fourth week, increasing after the fourth week. We suppose this variation is related to the moment of the first assessment.

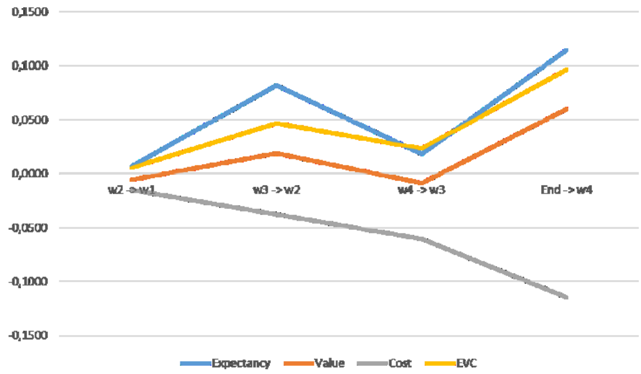
Table 40: Relation between motivation factors and status by week

		Expectancy		Value		Cost	
Week	**Status**	**Avg**	**p-value**	**Avg**	**p-value**	**Avg**	**p-value**
1	Pass Fail	0,877 0.687	4.19 e^{-7}	0.915 0.875	0.235	0,439 0,603	0.007
2	Pass Fail	0.760 0.690	0.055	0.846 0.851	0.893	0,405 0,552	0.001
3	Pass Fail	0.790 0.745	0.166	0.847 0.901	0.065	0,436 0,504	0.142
4	Pass Fail	0,828 0.689	0.010	0.909 0.913	0.897	0,438 0,542	0.224
Others	Pass Fail	0.722 0,557	8.3 e^{-6}	0.823 0.809	0.667	0,455 0,604	1.45 e^{-4}

Source: developed by the author

We also analyzed the correlation between the weekly indexes of motivation and the final grade of the students (Figure 14). We have identified a medium correlation between subscale "expectancy" and the final grade, during all periods.

Figure 15: Relation between grade and variation of the motivation indexes (all students)



Source: developed by the author

4.3.4 Discussion

With respect to the AQ1, we identified that the instrument is consistent and reliable, according to the omega coefficient.

With respect to the AQ2, we confirmed the convergence and discriminant validity between items, considering two distinct dimensions: positive factors (expectancy and value) and negative factor (cost).

With respect to the AQ3, we confirmed the existence of two dimensions that influence the responses, and not three as initially defined. Thus, we decided to join the first two factors (expectancy and value) in a single dimension.

With respect to the AQ4, we proved the relation between motivation and performance of students. We also identified that expectancy and cost factors have had significant variation depending on performance and outcome. We believe that the lack of a significant relation between value perceived and performance and outcome should be considered in future research.

With respect to the AQ5, although there is a correlation between motivation over time and students' performance and outcome, we found no significant difference between the weekly variation of the motivation indexes and the students' performance and outcome. We identified that the change in motivation among passed and failed students is similar in the first week and at the end of the course, having a small approximation in the second and third week, and returning in the fourth week.

We identified that, in general, students strongly perceive the importance of the course in the curriculum (value), but the expectancy of the successful student in the course gradually decreases through the semester. On the other hand, the perception of the effort required for the course is relatively high, increasing even more in the fourth week of classes.

Thus, we assume that students demotivate from the programming courses, mainly due to the level of difficulty encountered. In the researched sample, only about 17% already knew computer programming before entering the course and less than 24% had computing in basic education. Still, almost half (48.8%) of students assume they have regular or poor performance in Math.

These two situations are supposed to be reasons that contribute to the difficulty in programming and consequently, reduce the expectancy and increase the perception of cost through the course. These two factors corroborate the findings of Gomes (GOMES, 2000) who identified that students with difficulty in programming have enormous difficulties in resolving problems and a deficit of mathematical knowledge. In addition, the author concluded that the programming tasks, usually proposed by the teachers, present levels of difficulties misadjusted to the student's cognitive level, indicating a very large level of difficulty.

Similarly, Gomes (GOMES, 2000) identified that the personal perception of capacity and accomplishment has a relationship with the performance of the programming students. This result corroborates the findings in this study, where we perceived a strong relationship between the expectancy and students' performance.

4.3.5 Conclusions

This section evaluated a new instrument to assess students' motivation in computing courses. The goal was to create and to validate a simple and easy to apply scale, in order to identify at-risk students.

The EVC Light scale, adapted from (KOSOVICH, HULLEMAN, *et al.*, 2014), proved to be consistent and valid. It contains only six questions about expectancy, value, and cost factors. The results of the survey performed with 245 computing students also showed that motivation indexes had a significant relation to student performance (status and final grade), mainly related to the expectancy and cost factors.

In general, students with better results have a greater expectancy and lower perception of course cost, since the first weeks.

So far, those results were expected. However, we realized that the variation along the weeks is not significantly different among students with better and worse performance. With these results, we accept the null hypothesis that there is no correlation between motivation variation and performance of freshmen computing students. We identified that the results in the first few weeks tend to have similar behavior during all weeks of the semester.

A positive aspect of the instrument is to allow identifying motivation or lack of motivation and, consequently, at-risk students since the first weeks of the course. Another positive aspect is the simplicity of the instrument, which can be applied in a few minutes, consisting of a simple questionnaire that needs only six responses in a Likert scale, which greatly simplifies its application in class and that, compared to other models, seems to be less costly.

On the other hand, we have two possible assumptions to justify the lack of variation in the motivation indexes, which must be checked in future works: i) other previous factors are more important to the students' outcome; and ii) the teaching strategies and pedagogical actions not have an impact on the students' motivation.

4.4 PROFESSOR PERCEPTION

Given the significant role of teacher judgments in assessing students' academic achievement, a number of studies have examined the accuracy of these perceptions (BEGENY, ECKERT, *et al.*, 2008).

In an investigation of fourth-grade elementary students in Austria, a group of underestimated students was contrasted with a group of accurately estimated and overestimated students (URHAHNE, CHAO, *et al.*, 2011). The underestimated and overestimated students were classified based on teacher judgments of test performance in Math, learning motivation, academic self-concept, and text anxiety. Research results revealed that underestimated students had the same test performance as non-underestimated students but showed lower expectancy for success, displayed lower academic self-concept and felt more test anxiety. Two other investigations on motivation and emotions of under and overestimated students could confirm these findings (URHAHNE, CHAO, *et al.*, 2011). In another study, the pattern that emerged was that underestimated students had lower expectancy for success, a lower self-concept, and more test anxiety than overestimated students (ZHOU and URHAHNE, 2013).

Many questionnaire instruments exist for assessing students' motivation, primarily as self-report, but fewer instruments are available for assessing teachers' perceptions of their students' motivation. This is an important construct because it serves as the basis for internal feedback for teachers to self-regulate their motivational efforts (HARDRÉ, DAVIS and SULLIVAN, 2008).

Teachers' ability and effort perceptions are also closely related to children's academic achievement and motivation (UPADYAYA and ECCLES, 2015).

Hardré et al. (2008) proposed instruments to assess teachers' perceptions of student motivation. The Perceptions of Student Motivation (PSM) questionnaire combine the two methods of assessing motivation, as they must ask teachers to self-report their perceptions of students' motivation, based on their observations of students' behaviors that would reasonably result from motivation or the lack of it. The general motivation subscale assesses teachers' perceptions of the strength of students' academic motivation (their interest, effort, and engagement) based on behaviors that teachers may observe in their classrooms (7 items):

a) The students in this class really try to learn;
b) My students work at learning new things in this class;
c) My students generally pay attention and focus on what I am teaching;
d) The students in this class generally do class-related tasks and assignments willingly;
e) The students in this class don't put forth much effort to learn the content;
f) My students are often distracted or off task, and I have to bring them back to focus on the topic or work at hand;
g) In general, my students are genuinely interested in what they are asked to learn in my class.

As this work goal is to create an easy and simple method to measure the motivation several times across the course, we proposed a reduced questionnaire to the professor's perception of students' motivation. While Perceptions of Student Motivation (PSM) questionnaire to aim to assess the motivation of the entire class, we propose a questionnaire to assess the individual student motivation and interest variation, using only one question with a Likert-type 7 point scale, being negative and positive variation as shown in Table 41.

Table 41: Professor's perception about students' motivation questionnaire

Question	**Negative variation**			**No variation / Not perceived**	**Positive variation**		
"What is your perception about the motivation, performance and interesting of each student? (If you do not have clear perception, answer zero)"	-3	-2	-1	0	+1	+2	+3

Source: developed by the author

This questionnaire is answered by the professor for each student weekly. We identified that positive evaluations and extreme negative evaluation (-2) affect the student' outcome (Table 42).

Table 42: Regression model of professors' perception

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
Intercept	2.70848	0.41796	6.480	$2.36e^{-10}$
week	2.70848	0.11822	2.032	0.04277
perception[T.-2]	-2.64708	1.00293	-2.639	0.00859
perception[T.-1]	-0.85281	0.59126	-1.442	0.14988
perception[T.0]	-0.35660	0.45123	-0.790	0.42977
perception[T.1]	2.02864	0.47836	4.241	$2.69e^{-05}$
perception[T.2]	2.83745	0.53317	5.322	$1.61e^{-07}$
perception[T.3]	2.33728	0.45437	5.144	$3.98e^{-07}$

Source: developed by the author

To verify the result changes over time, we analyzed weekly the different perceptions between passed and failed students, and the correlation with the final grade. Table 43 shows that there is a significant difference in students' outcome according to professor perception in every week and a moderate correlation between the professor perception and the student's grade.

Table 43: Weekly results of professors' perception

	Week 1		**Week 2**		**Week 3**		**Week 4**		**Week 5 – end**	
	Mean	p	Mean	p	Mean	p	Mean	p	Mean	p
Pass	1.19	$1.1e^{-5}$	1.97	$6.9e^{-4}$	1.52	$9.3e^{-4}$	0.687	0.0243	1.21	$2.1e^{-9}$
Fail	0.30		0.55		0.26		-0.29		-0.78	
Grade	0.433	$2.3e^{-7}$	0.397	$8.3e^{-5}$	0.408	$9.9e^{-6}$	0.244	0.0103	.708	$6.2e^{-9}$

Source: developed by the author

These results indicate that the perception of the professor is related to the result of the student, since the first few weeks of school. So, professor perception can be a candidate attribute to predictions of students at risk.

4.5 MOTIVATION FACTORS

In addition, to predict at-risk students it is important to identify the reasons or factors that may be contributing to the demotivation and failure of the students. Therefore, we propose in the EMMECS method an instrument that is applied to identify those factors.

There are several reasons that can be considered as factors for demotivation and consequently increasing failures, including specific factors in the area, among which is the students' difficulty with computer programming (NIITSOO, PAALES, *et al.*, 2014) (BERGIN and REILLY, 2005). And the lack of familiarity with the subject (CARTER, 2006). Another factor addressed by some studies is that many novice students relate STEM disciplines (Science, Technology, Engineering, and Mathematics) as being interdisciplinary and innovative. However, this view is often not confirmed by the first experiences in the university, bringing disappointment and doubts (PETERS and PEARS, 2013).

But are these the only factors that impact student success? According to (SINCLAIR, BUTLER, *et al.*, 2015), more qualitative data are required and other measures (such as student expectation and some specific measure) are necessary for the wide understanding of the experience of the computer science student.

We consider that the first step towards an improvement process in this scenario of high rates of failure of students in computer courses is to understand the current problems and to search for potential alternatives of solution. In this way, aiming at increasing the number of computing students graduating, it is important to understand what makes students stay motivated and engaged in the course, which will possibly produce better performance and highest passing rate.

There are various initiatives concerning the motivation of computing students; most of them propose or report the use of new educational approaches and tools (NAVARRO and VAN DER HOEK, 2009) (SERRANO-CÁMARA, PAREDES-VELASCO, *et al.*, 2013) (DEBDI, PAREDES-VELASCO and VELÁZQUEZ-ITURBIDE, 2014). However, when the issues and factors that influence motivation and engagement are investigated, few studies converge or use categories that may be followed by other researchers, hindering the dissemination and replication of their studies.

In this context, this section proposes and validates an instrument to identify factors that can affect the motivation and engagement of computing students. The instrument was defined based on a compilation of literature works (SCHOEFFEL, WAZLAWICK and RAMOS, 2018c). As the main contributions of this section, we have the compilation of various factors that can affect the motivation and engagement of students, and the reliability and validation of an instrument so that other researchers can use it for new studies.

In fact, the factors were the basis for the creation of the method and, at the same time, it was used to be a complement to the method, to identify reasons for failures, although it has no direct relation with the prediction. Therefore the validation of this instrument was performed before, but the instrument was described later because the method application cycle was followed.

4.5.1 Research Method

This section aims to validate an instrument to identify motivation factors in computing students. The instrument was built based on a group of factors compiled from several works of literature and published by (SCHOEFFEL, WAZLAWICK and RAMOS, 2018c).

In order to perform an evaluation of the questionnaire, we conduct a case study following the same process described in Section 4.3.1. The only difference was the use of Cronbach's alpha coefficient to calculate the internal consistency reliability.

4.5.2 The Instrument

It was worked out a questionnaire with 48 items, including the motivation factors of the literature review (SCHOEFFEL, WAZLAWICK and RAMOS, 2018c), demographic and entrance data of the students at the university and a light scale for the measurement of actual motivation. The items are divided into six groups as shown in Table 44: personal and demographic data, the general perception of motivation, university perception, student behavior, course perception, and class/professor perception.

Groups 3, 4, 5 and 6 of the questionnaire were based on a compilation of factors extracted from the literature. Each item has options following a Likert scale of five points (SA – strongly agree, A - agree, N-neutral, D - disagree, SD – strongly disagree).

The group "general perception of motivation" is a light scale adapted from Vallerand (1992) and Jenkins (2001).

Table 44: Questionnaire groups and factors

Group	Factor
1. Personal and demographic data	1A – Gender 1B – Quota 1C – Entrance exam position 1D – Age 1E – Way of entering
2. General perception of motivation	2A – General level of satisfaction with the program

	2B – Reasons to continue studies 2C – Level of intention to continue studies 2D – Reasons to dropout 2E – Self-efficacy
3. Perception about University	3A – Adequate student support 3B – Adequate learning resources 3C – Adequate LMS (learning management system) 3D – Level of satisfaction of faculty 3E – Graduation and qualification of faculty
4. Student behavior	4A – Feeling of being prepared for the study 4B – Interaction with students outside of the academic environment 4C – Sense of belonging to the University 4D – Interaction with different students 4E – Participation in discussions with students and professors 4F – Group work with other students 4G – Attendance 4H – Commitment to activities and deadlines 4I – Studying in the correct way and proper time managing for activities 4J – Trying to do more than requested 4k – Doing the best to stand out in the class
5. Perception about program	5A – Installations of industrial importance and updated 5B – Alignment with the job market 5C – Appropriate type of program and courses (the focus of course in computing, schedules, etc.) 5D – Appropriate curriculum (syllabus) and program of courses (contents) 5E – Ease of insertion in the labor market and prospects for the future 5F – Balance between areas of knowledge allowing a systemic vision 5G – Proper teaching quality
6. Perception about course and professor	6A – Active learning 6B – Fun 6C – Challenges 6D – Peer learning 6E – Diversity of pedagogical approaches 6F – Team spirit 6G – Practice outside the classroom 6H – Utility and future application of the contents 6I – Participation of the student in decision making 6J – Reward for the effort 6K – Information provided 6L – Adequate difficulty level 6M - Clarity in the goals of the course 6N – Students with difficulty are not exposed 6O - Gender distribution of faculty

Source: developed by the author

4.5.3 Questionnaire Evaluation

The execution phase was based on the application of a questionnaire (survey) to 112 students of the Bachelor Program on Software Engineering at the University of Santa

Catarina State in Ibirama - Brazil. The questionnaires were applied in the period from November 27th to December 1st, 2017, at all semesters of the program.

With the objective of evaluating the quality of the instrument to measure motivation factors in computing students, we defined the following analysis questions (AQ):

Reliability

- AQ1: Is there evidence for internal consistency of the questionnaire?

Construct Validity

- AQ2: Is there evidence of the convergent and discriminant validity of the questionnaire?
- AQ3: How do underlying factors influence the responses on the items of the questionnaire?

We analyzed each of the analysis questions as defined in the research methodology.

4.5.3.1 Reliability

AQ1: Is there evidence for internal consistency of the questionnaire?

The value of the Cronbach coefficient for the entire questionnaire was satisfactory (0.8904). Analyzing by questionnaire group, we can see the following results of reliability: i) general perception of motivation – 0.6429; ii) perception about university – 0.6503; iii) student behavior and engagement – 0.8416; iv) perception about the program – 0.8037; and v) perception about classes and faculty – 0.9006.

We can see that two groups have gotten the coefficient a little below the acceptable. Regarding the group "general perception of motivation", the fact that we use different scales for the items answers may have impacted on the result. Regarding the group "perception about the university", we identified that there are 2 items in this group that are related to faculty (career satisfaction, faculty training, and qualification) and, because of this, the internal consistency of the group may have been changed. After these findings, we moved these items to the group “perception about classes and faculty”.

4.5.3.2 Construct Validity

AQ2: Is there evidence of the convergent and discriminant validity of the questionnaire?

See in Table 44 to Table 47 the results of item intercorrelation of every questionnaire group.

Table 45: Intercorrelation– perception about the university

	3A	3B	3C	3D	3E
3A	1.000				
3B	0.285	1.000			
3C	0.295	0.383	1.000		
3D	0.227	0.247	0.367	1.000	
3E	0.107	0.252	0.272	0.482	1.000

Source: developed by the author

According to Table 45, the number of significant correlations was 4 (40%) to the "perception about university" factor group. However, all other correlations had positive values and most of them close to the minimum value considered acceptable (0.29).

Table 46: Intercorrelation– student behavior

	4A	4B	4C	4D	4E	4F	4G	4H	4I	4J
4A	1									
4B	0.262	1								
4C	0.348	0.427	1							
4D	0.112	0.471	0.364	1						
4E	0.277	0.274	0.356	0.432	1					
4F	0.270	0.416	0.381	0.459	0.393	1				
4G	0.403	0.157	0.231	0.039	0.262	0.262	1			
4H	0.304	0.150	0.155	0.234	0.255	0.224	0.414	1		
4I	0.505	0.178	0.197	0.151	0.281	0.267	0.348	0.460	1	
4J	0.504	0.094	0.174	0.158	0.391	0.250	0.518	0.351	0.481	1
4K	0.461	0.152	0.163	0.067	0.289	0.210	0.405	0.287	0.544	0.703

Source: developed by the author

According to Table 46, the number of significant correlations was 25 (45.5%), in bold, for the "student behavior" factor group. However, all other correlations had positive values and most of them close to the minimum value considered acceptable (0.29).

Table 47: Intercorrelation– the perception of the program

	5A	5B	5C	5D	5E	5F	5G
5A	1.000						
5B	0.412	1.000					
5C	0.283	0.537	1.000				
5D	0.160	0.474	0.646	1.000			
5E	0.446	0.425	0.387	0.421	1.000		
5F	0.449	0.356	0.492	0.404	0.489	1.000	
5G	0.410	0.357	0.407	0.273	0.262	0.426	1.000

Source: developed by the author

According to Table 47, the number of significant correlations was 17 (81%), in bold, to the "perception about program" factor group. However, all other correlations had positive values and most of them close to the minimum value considered acceptable (0.29).

According to Table 48, the number of significant correlations was 72 (68.6%), in bold, for the "perception about courses and professors" factor group. However, all other correlations had positive values and most of them close to the minimum value considered acceptable (0.29).

Table 48: Intercorrelation of items – the perception of courses and professors

	6A	6B	6C	6D	6E	6F	6G	6H	6I	6J	6K	6L	6M	6N
6A	1.000													
6B	0.326	1.000												
6C	0.486	0.644	1.000											
6D	0.370	0.440	0.571	1.000										
6E	0.380	0.366	0.285	0.373	1.000									
6F	0.285	0.250	0.313	0.360	0.467	1.000								
6G	0.375	0.331	0.389	0.441	0.567	0.512	1.000							
6H	0.325	0.283	0.304	0.349	0.465	0.408	0.521	1.000						
6I	0.342	0.276	0.330	0.271	0.448	0.484	0.595	0.509	1.000					
6J	0.301	0.290	0.211	0.210	0.364	0.266	0.426	0.460	0.471	1.000				
6K	0.173	0.251	0.202	0.249	0.250	0.394	0.423	0.402	0.385	0.450	1.000			
6L	0.281	0.111	0.228	0.414	0.318	0.347	0.502	0.394	0.504	0.304	0.370	1.000		
6M	0.308	0.239	0.168	0.209	0.387	0.263	0.467	0.435	0.332	0.452	0.388	0.473	1.000	
6N	0.087	0.019	0.004	0.182	0.302	0.228	0.331	0.292	0.252	0.177	0.324	0.402	0.421	1.000
6O	0.224	0.001	0.072	0.169	0.217	0.307	0.361	0.363	0.415	0.449	0.385	0.499	0.488	0.377

Source: developed by the author

We identified that, in general, most of the items that are part of the same constructor have an acceptable correlation. Similarly, no negative correlation within the same constructor was found. Therefore, we observe evidence that the items of each constructor are correlated, indicating evidence of convergent validity.

On the other hand, we found some correlations between items of different groups (16%), more frequently between items in the group related to the university and program (34%), as well as between program and faculty/courses (52%). For example, item 8.7 (proper teaching quality) has a significant correlation with all items in the university group. This can indicate that there are some factors group overlapping.

Item-total correlation. Only items 6.1 (0.2787), 7.2 (0.2537), 7.4 (0.199), 7.6 (0.1973), and 7.8 (0.2239) have an item-total correlation below 0.29, but with positive and not substantially low values. In addition, the value of Cronbach's alpha if the items were removed did not increase or did increase in a non-significant way.

The degree of correlation between the items determines the degree of convergent and discriminant validity. But these results do not determine how many factors underlie the set of the questionnaire. With this objective, we performed a factor analysis, answering the analysis question AQ3.

AQ3: How do underlying factors influence the responses to the items of the questionnaire?

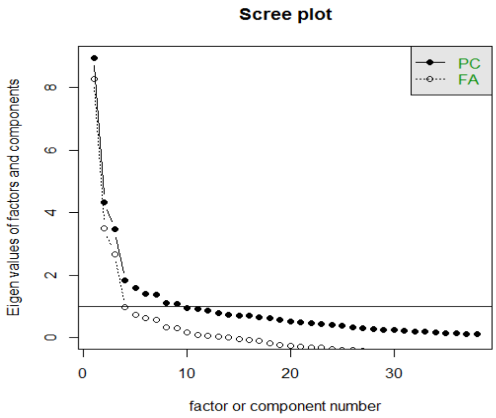
In order to identify the number of factors (quality factors or dimensions) that represents the responses of the set of the 38 items of the questionnaire, we performed a factor analysis. Based on the original definition of the questionnaire, we assume that it is influenced by four groups of factors (perception about the university, student behavior, perception about program and perception about classes and faculty).

Analyzing the set of items of the questionnaire, we obtained a KMO index of 0.7822471 and a significance level of 0.000, indicating that factor analysis is appropriate in this case.

Following the Kaiser-Guttman criterion, our results show that 6 factors should be retained, explaining 59.65% of the data. The scree plot (Figure 16) shows the eigenvalue for each factor number (representing each item).

We also use a parallel analysis. In a parallel analysis, actual eigenvalues are compared with random order eigenvalues. Factors are retained when actual eigenvalues surpass randomly ordered eigenvalues (WILLIAMS, ONSMAN and BROWN, 2010).

Figure 16: Factor analysis of adjusted questionnaire

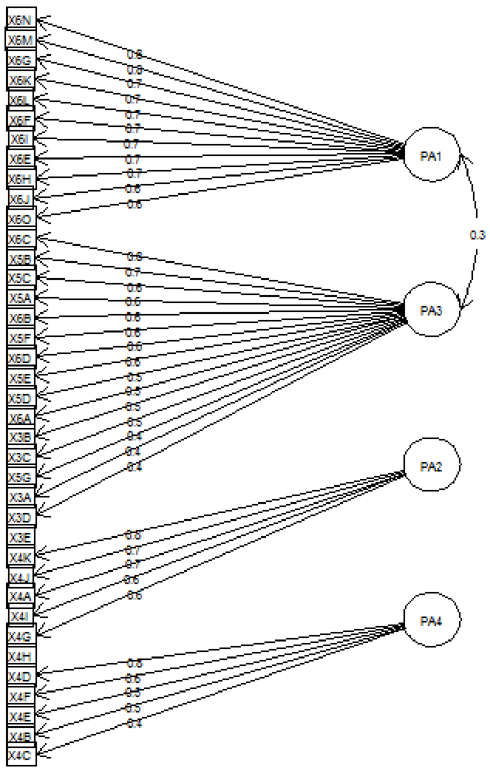


Source: Developed by the author

Figure 16 shows the scree plot considering the values of the principal components analysis and parallel analysis. Although the scree plot indicates six components above one, the parallel analysis shows three or four factors. By comparing to the three alternatives (3, 4, 5 and 6 factors), the accumulated value is 39.5%, 43.3%, 46.1%, and 49%, respectively. However, by analyzing the graph of the factorial analysis, by using six factors we realize that the sixth factor has only one variable loading on it. Thus, this probably represents an over-extraction and let's look at the five-factor solution. By using five factors we realize that the fifth factor has no variable. Because of this, the use of four factors seems to be more indicated.

After identifying the number of underlying factors, we determined which items are loaded into which factor. In order to identify the factor loadings of the items, we used the Varimax with Kaiser Normalization rotation method, because it is the most widely accepted and used rotation method (COHEN, 1988).

Figure 17: Factor analysis graph of the adjusted questionnaire



Source: Developed by the author

In this way, the four original factors are maintained, but there is a different distribution of the items in the groups. Factor 1 (PA1) contemplates twelve items, all related to the group “perception of classes and faculty”, related more to the behavior of professors. Factor 2 (PA2) brings together sixteen items, being seven related to the group “perception of the program”, five related to the group “perception of the university” and four related to the group “perception of classes and faculty”. These last four items are related to the student’s learning experience. Factor 3 (PA3) brings together five factors, all of which relate to “student behavior and engagement”.

It is important to verify that these items are more related to individual behavior than with aspects of the interaction between colleagues and the university. Factor 4 (PA4) includes the other six items related to the student’s behavior and engagement group, and such items are related to student interaction with colleagues and the environment. Another important aspect is that two items (3E and 4H) have no significant correlation with any factor. Fig. 2 shows a factor analysis graph of the adjusted questionnaire.

Table 49 shows the factor loadings of the items associated with the four retained factors. The highest factor loading of each item, indicating to which factor the item is most related, is marked in bold. The Chi-Square statistical test proves that four factors are enough (p-value = 0.000000708).

Table 49: Factor Loadings

	3A	3B	3C	3D	3E	4A	4B	4C	4D	4E	4F	4G	4H	4I	4J	4K	5A	5B	5C
F1		0.112		0.147	0.134			0.128								0.147			
F2	0.376	0.476	0.412	0.361	0.295			0.237		0.199					0.119		0.615	0.649	0.608
F3					0.197	0.608	0.205	0.214		0.341	0.261	0.559	0.239	0.592	0.777	0.853			
F4				0.157		0.309	0.529	0.444	0.722	0.601	0.623	0.233	0.281	0.210	0.181		0.130		

	5D	5E	5F	5G	6A	6B	6C	6D	6E	6F	6G	6H	6I	6J	6K	6L	6M	6N	6O
F1	0.183	0.147	0.319	0.217	0.120			0.196	0.709	0.678	0.750	0.676	0.698	0.620	0.719	0.698	0.749	0.758	0.569
F2	0.547	0.567	0.619	0.450	0.529	0.626	0.787	0.606	0.250	0.211	0.363	0.309	0.222	0.310	0.129	0.155	0.116	-0.167	
F3			0.145	0.132		0.130											0.141	0.165	
F4				-0.209	0.100			0.108		0.204	0.110			-0.166	-0.139	0.103			

Source: Developed by the author

4.5.4 Factors That Impact on Student Motivation

This section describes the impact of motivational factors on the performance and motivation of Software Engineering students, to understand what keeps students motivated and engaged with the course, possibly increasing the performance and approval rate. In order to achieve this goal, the following seven research questions have been developed: AQ1 – Are personal data (gender, age) and way of entering in the program related to motivation and performance? AQ2 – Are motivation factors related to the students' satisfaction and intent to continue studies? AQ3 - Is the perception about university related to the motivation and performance of the students? AQ4 – Is the perception of the student's behavior in the studies (engagement) related to motivation and performance? AQ5 - Is the perception of the program related to the motivation and performance of the students? AQ6 - Is the perception about classes and faculty related to the motivation and performance of the students?, and AQ7 – Are students' performance and performance perceptions related to motivation?

Of the respondents, 64% work full-time during the day, 11% work partial-time as a scholarship student, 74% are male, 10% female and 16% have not informed. Most respondents entered via the university exam (Vestibular) (78.6%) and the national high school exam (ENEM) (13.4%). The average age of respondents is 21.55 years old, although more than half (50.9%) is 20 years old or less. Just over half (50.9%) entered in the first call and 26.8% are quota students, primarily by public school quota (25.9%).

Table 50: Results by demographic data

	Gender	Entry	Age	Exam position	Quota
General satisfaction	0.703 [a]	0.324 [a]	0.697 [b]	0.892 [a]	0.052 [a]
Motivation to continue	0.396 [a]	0.341 [a]	0.514 [b]	0.617 [a]	0.408 [a]
Average grade	0.811 [b]	0.814 [b]	0.118 [c]	0.002 [b]	0.842 [b]
Approval rate	0.538 [b]	0.960 [b]	0.128 [c]	0.002 [b]	0.158 [b]
Type of motivation	0.556 [a]	0.624 [a]	0.088 [b]	0.1445 [a]	0.360 [a]

[a.] Pearson's chi-square test / [b] ANOVA one way / [c] Spearman's coefficient

Source: Developed by the author

In order to answer the first analysis question (AQ1), we analyzed the correlation between the demographic data (gender, age), student entrance data (way of entering, quota and entrance exam position), motivation, and motivation factors (Table 50). We verified that only the entrance exam position was related to the average grade and approval rate[3], and no variable had significant variance according to the type of motivation. We identified that first-call students have better performances, as shown in Table 51.

Table 51: Results by entrance exam position

Exam position [a]	General Average			Approval rate		
	Avg.	**SD**	**Qty**	**Avg.**	**SD**	**Qty**
1	7.855	0.939	47	0.952	0.080	47
2	6.711	1.685	18	0.854	0.179	18
Other	6.660	1.670	10	0.850	0.118	10
NA	6.786	1.220	7	0.817	0.152	7

[a.] ANOVA one way

Source: Developed by the author

We analyzed the correlation between the overall grade average, success rate index and the type of students' motivation. We did not find significant correlation of motivation with grade average ($p = 0.94$) and approval rate ($p = 0.58$). The only strong correlation found was between the success rate and the overall average (Spearman's coefficient, $r = 0.884$).

Considering the motivation to continue studies, 55% of students never thought of dropout, 40.5% already thought or still think of dropping out, but intend to continue and 4.5% seriously think about giving up. We identified a variance in the type of motivation, according

[3] Approval rate is the ratio between the number of passed courses and the number of failed courses from a student

to the intention to drop out (p = 0.041). Students intrinsically motivated have a lower level of intention to drop out, following by extrinsically motivated and demotivated (Table 52).

Table 52: Motivation related to the intention to drop out

	IM	EM	IEM	PD	TD	Pearson	Fisher
No	35	14	4	8	0	0.0481	0.0417
Yes	5	10	1	9	6		

IM – intrinsic motivated, EM – extrinsic motivated, IEM – intrinsic/extrinsic motivated, PD – partially demotivated, TD – totally demotivated

Source: Developed by the author

About possible reasons to drop out the program, 48.2% of students have no reason, because they do not think about dropout, 17% indicated the difficulty in reconciling studies and work, 15.2% informed that the distaste for programming can be a reason for quitting, 9.8% the lack of affinity with the course, 4.5% the difficulty in Math, in addition to 13.4% that reported other reasons.

Students who have no reason to drop out are more intrinsically motivated and less demotivated. The main reasons why students continue in the program are the search for knowledge (64.3%), taste for the field (64.3%), and achieving good positioning in the job market (46.4%). These results show that the main motivation to continue the studies is intrinsic, related to the attraction by the field, but extrinsic motivation related to the job market is also another important factor.

We identified seven factors that impact on the intention of dropout: i) sense of belonging to the university (p = 0.0029); ii) balance between areas of knowledge allowing systemic vision (p = 0.0453); iii) fun (enjoyable activities) (p = 0.0191); iv) academic challenges (activities that promote reflection, creation of strategies, etc.) (p = 0.0454); v) peer learning (exchange of knowledge with colleagues) (p = 0.0213); vi) promotion of team spirit (class union, group activities, collaboration) (p = 0.04268); and vii) usefulness and future application of the content learned (p = 0.0029). Students that evaluated negatively these factors have a higher level of intention to drop out. Then we analyze the results of each group of motivation factors.

4.5.4.1 University-related factors

Regarding the university environment, and the AQ3 question, we evaluated the variance of the grade average, approval rate and students' type of motivation, according to the perception of the factors related to the university environment.

Table 53: Results of perception about the university

	Grade average		Approval rate		Type of motivation	
Question	**F value**	**Pr (>F)**	**F value**	**Pr (>F)**	**X^2**	**Pr (>F)**
3A Adequate student support	0.255	0.775	1.458	0.237	7.517	0.351
3B Adequate learning resources	0.122	0.947	0.243	0.866	40.238	0.001
3C Adequate LMS (learning management system)	2.121	0.086	2.803	0.031	16.947	0.656
3D Level of satisfaction of faculty	0.335	0.716	0.352	0.704	9.8767	0.873
3E Faculty Graduation and qualification	2.023	0.118	2.314	0.082	32.722	0.036

Source: Developed by the author

We found three significant variances, as shown in Table 53: i) approval rate according to perception of LMS support (3C); ii) type of motivation according to the perception of adequate learning resources (3B); and iii) type of motivation according to the perception of qualification of faculty (3E). The students who have negatively evaluated their learning/infrastructure resources are more often demotivated, but students who have positively evaluated the qualification of faculty have lower intrinsic motivation and a higher rate of demotivation. The students who have evaluated the virtual environment very positively, have a lower approval rate. We suppose that unmotivated students blame the bad infrastructure of the university for their failure and motivated students are more critic about the professor and the learning process.

4.5.4.2 Factors related to student behavior

Students were questioned about their academic performance. About 60% of students indicate good performance and expect to complete the program without problems. As expected, students with a better perception of performance have a higher grade average ($p = 0.00224$). The approval rate factor had similar behavior, but the statistic confidence was lower ($p = 0.0596$). We found no significant variance of motivation type according to the perception of performance ($p = 0.454$).

We found significant variance in the overall grade average for five factors, as shown in Table 54: i) feel prepared for the studies (4A); ii) often work in a group with other students (4F); iii) attendance in class (4G); iv) commit the activities and deadlines (4H); and v) do the best to stand out in class (4K).

Table 54. Results of factors related to student engagement

	Grade average		Approval Rate		Type of motivation	
Item	*F value*	*Pr(>F)*	*F value*	*Pr(>F)*	*X square*	*p-value*
4A – Feeling of being prepared for the study	3.231	0.045	2.709	0.073	7.8008	0.801
4B – Interaction with students outside of the academic environment	1.729	0.168	0.863	0.464	33.092	0.007
4C – Sense of belonging to the University	0.459	0.765	1.093	0.366	52.458	0.000
4D – Interaction with different students	0.767	0.516	0.714	0.546	12.961	0.879
4E – Participation in discussions with students and professors	0.52	0.721	0.624	0.647	11.356	0.936
4F – Group work with other students	3.017	0.035	3.531	0.019	17.474	0.356
4G – Attendance	2.299	0.066	1.272	0.288	17.555	0.617
4H – Commitment to activities and deadlines	5.212	0.003	1.578	0.201	9.066	0.697
4I – Studying in the correct way and proper time managing	1.771	0.143	1.433	0.231	14.998	0.777
4J – Trying to do more than requested	1.785	0.157	1.269	0.291	11.071	0.805
4k – Doing the best to stand out in the class	3.047	0.034	1.272	0.290	10.325	0.849

Source: Developed by the author

Students who agree with item "work often in the group with other students" (4F) have better grades. In addition, only two factors were related to the variance of motivation type: i) interact with students outside the academic environment (4B); ii) sense of belonging to the university (4C). For example, we identified that partially or totally demotivated students evaluated more negatively the "sense of belonging to the university" factor.

4.5.4.3 Program-related factors

We observed a low number of students who strongly agree with the program-related factors, showing that even in positively evaluated aspects (AGREE), students believe that improvement is needed.

Table 55: Program-related factors

	General grade average		Success rate		Type of motivation	
Item	*F value*	*Pr(>F)*	*F value*	*Pr(>F)*	*X-square*	*p-value*
5A – Installations of industrial importance and updated	0.620	0.649	1.356	0.257	36.748	0.002
5B – Alignment with the job market	0.196	0.899	0.393	0.758	31.052	0.001

5C – Appropriate type of program and courses (focus of course in computing, schedules, etc.)	0.252	0.908	0.257	0.905	47.710	0.000
5D – Appropriate curriculum (syllabus) and program of courses (contents)	0.603	0.615	1.452	0.234	45.448	0.000
5E – Ease of insertion in the labor market and prospects for the future	1.148	0.341	1.823	0.133	26.167	0.052
5F – Balance between areas of knowledge allowing a systemic vision	0.694	0.598	0.849	0.499	65.991	0.000
5G – Proper teaching quality	0.389	0.816	1.513	0.207	20.323	0.206

Source: Developed by the author

Regarding the AQ5 question, we verified that there was significant variance in the type of motivation for all items except for the item related to the quality of teaching offered, and we found no significant variance for the general grade average and success rate (Table 55).

4.5.4.4 Factors related to classes and faculty

Regarding the factors related to faculty, didactics, teaching strategies, evaluation, among others, intrinsically and extrinsically motivated students have evaluated more positively than demotivated students. Analyzing the AQ6 question, we observed that, in general, the factors related to the courses and faculty are related to the type of motivation, as shown in Table 56.

Table 56. Factors related to classes and professors

	Grade average		**Approval rate**		**Type of motivation**	
Item	***F value***	***Pr (>F)***	***F value***	***Pr (>F)***	***X^2***	***P value***
6A – Active learning	1.402	0.248	0.807	0.493	13.215	0.3536
6B – Fun	1.377	0.256	3.568	0.018	37.157	0.0002
6C – Academic Challenges	0.566	0.639	1.263	0.293	55.951	0.0000
6D – Peer learning	0.71	0.549	2.455	0.069	23.59	0.0231
6E – Diversity of pedagogical approaches	0.202	0.936	0.452	0.771	13.618	0.6271
6F – Team spirit	0.104	0.981	0.347	0.845	12.337	0.7205
6G – Practice outside the classroom	0.477	0.753	0.787	0.537	20.438	0.2012
6H – Utility and future application of the contents	0.203	0.936	0.113	0.978	68.010	0.0000
6I – Participation of the student in decision taking	0.348	0.844	0.660	0.622	18.812	0.2785
6J – Reward to effort	0.707	0.590	0.676	0.611	27.580	0.0355
6K – Information provided	0.243	0.866	0.475	0.701	15.842	0.1986
6L – Adequate difficulty level	0.372	0.828	0.149	0.963	18.286	0.3074
6M – Clear and defined goals of the course	0.645	0.632	0.916	0.459	31.904	0.0103
6N – Students with difficulty are not exposed	0.515	0.673	0.675	0.570	7.0252	0.9728
6O - Gender distribution of faculty	0.162	0.957	0.362	0.835	19.454	0.2458

Source: Developed by the author

We found no significant relationship between the factors and the grade average. We found significant variance in the type of motivation for the following factors: i) fun (enjoyable activities) – 6B; ii) academic challenges (activities that promote reflection, creation of strategies, etc.) – 6C; iii) peer learning (exchange of knowledge with colleagues) – 6D; iv) usefulness and future application of the content learned – 6H; v) fair reward for students who are most dedicated – 6J; vi) clear goals allowing to realize whether the student is achieving satisfactory performance throughout the course/program – 6M.

In general, students demotivated or partially demotivated have evaluated more negatively (disagree or totally disagree) the items related to classes and faculty, while students intrinsically motivated evaluated them more positively. Considering the approval rate, the only factor that had significant variance was “Fun”. Students who strongly agreed or strongly disagreed with the factor "fun" had a lower approval rate.

4.5.5 Discussion and Conclusion

Regarding AQ1, we confirmed a significant variance in the student's performance (average grade and approval rate) according to the entrance exam position. However, we found no significance in other relations. This shows that the performance in the entrance exams is related to the student's performance at the university. Then, we suggest an effort to increase the number of enrollments of those candidates who passed in the first places.

Regarding AQ2, we realized that the intention to continue is mainly related to aspects of the teaching-learning process in the classroom and which therefore can be managed by the professors. It is important to note that these are not necessarily the factors causing dropout, for example. But there is a different perception about these factors for students who have the intent to drop out and those who do not. We identified that students are more motivated if they perceive better the future utility and application of the subject and think that the goals of the course are clear. Also, the use of some strategies to teach, such as fun activities, challenges, peer learning, and rewards can improve the students' motivation and their success. To do that, we recommend using more active, dynamic, and experiential activities, gamification, and serious games, among others.

In relation to AQ3, we found significant variances in students’ performance. In three university-related factors and the type of motivation according to the perception of the qualification of the faculty. Interestingly, intrinsically motivated students are more critical about professor qualification. The infrastructure is a current problem in the studied university, but it has been evaluated in an equivalent way regardless of the level of motivation or performance of the students.

In relation to AQ4, we realized that the greatest number of performance-related factors is related to the behavior and engagement of the students. This may indicate that the lack of engagement impacts performance or that there is perceived guilt of students with poor performance. However, the motivation of the students does not change because of this perception, and it can indicate that they continue with interest and other aspects could help them succeed.

With respect to AQ5, we verified that there was a significant variance in the type of motivation for most items related to the program and found no significant variance for general grade average and success rate. Regarding the AQ6, we observe that, in general, the classes and faculty factors have a relation with motivation but have no significant relation with grade average and success rate of the students. These results from AQ5 and AQ6 may indicate that students, regardless of performance, are motivated by interaction with the professor and classmates in the classroom, didactics, and other aspects related to the professor and classes, in addition to the characteristics of the program. In addition, these results may indicate that perceptions of teaching-learning aspects are similar among students with lower and higher performance, or that these factors do not influence students' performance. In the context of this research, we think that the first option fits better.

With respect to AQ7, as expected, students with better performance perceptions have a higher overall grade average ($p = 0.00224$). The approval rate had the same behavior, but with lower statistical significance ($p = 0.0596$). With respect to the type of motivation ($p = 0.454$), we found no significant variance.

Although there are several factors that are related to the type of motivation, it is interesting to note that the type of motivation was not related to the grade average and the success rate of students. Partially and totally unmotivated students have higher approval rates. This may indicate that although there are factors that may impact on the reasons that lead the student to graduate (type of motivation), the type of motivation is not necessarily associated with the performance of students, unlike many studies in the literature (SOENENS and VANSTEENKISTE, 2005) (RYAN and DECI, 2000).

In the context studied, these results suggest that performance is not a decisive factor for the student to be motivated. We believe that performance is a consequence of motivation and not vice versa. In summary, we found 15 factors with significant variance in the type of motivation of the student (has adequate learning resources, sense of belonging to the university, qualification of faculty, alignment with the job market, access to facilities of industrial importance and relatively up-to-date, and suitable program and courses type, interaction with students outside the academic environment, adequate curricular matrix and courses contents, and existence of mechanisms to facilitate their insertion into the labor market and prospects for the future). We found only three factors with variance in the approval rate (entrance exam position, fun, and LMS support) and five factors with variance in the overall grade average (entrance exam position, feel prepared for study, do the possible to stand out in the class, and commitment of activities and deadlines).We conclude that, for the sample used, the perception of student behavior (engagement) is associated with performance. The type of motivation is associated mainly with the perception of the program, and the perception about the classes and faculty. The intention of dropout is mainly associated with the perception of the classes and faculty.

These results show that the type of motivation does not impact on performance, but it can impact on the students' retention. The results also show that the factors related to the course and the process of teaching (didactic, professors) are the most relevant to students. Through these results, we assume that the students' retention has a greater relationship with factors perceived throughout the program than with previous factors.

Based on the results obtained, we understand that in order to increase students' motivation, the classes should have more fun, have more practical aspects that promote more interaction and group work among students and find a way to reward the students that are more dedicated.

In this sense, we understand that the concern with the didactic aspects, through the continued training of the professors in pedagogical issues or the incentive to the development of innovative and active teaching strategies are very important to increase the retention of students in computing programs. Especially in countries where basic education is deficient, and students do not have as much discipline and autonomy to study, teaching strategies can make a difference for students to be motivated and be engaging in their activities and thereby increase retention and success in the computing courses and programs.

5 AT-RISK STUDENTS PREDICTION

In this chapter, we describe the research method, application, and results of the implementation of the method to predict in advance students at risk. The chapter is divided into 4 sections. In section 5.1 it is detailed the research methods used. In Section 5.2 the prediction results obtained are presented. In sections 5.3 and 5.4, the results are discussed and conclusions presented, respectively.

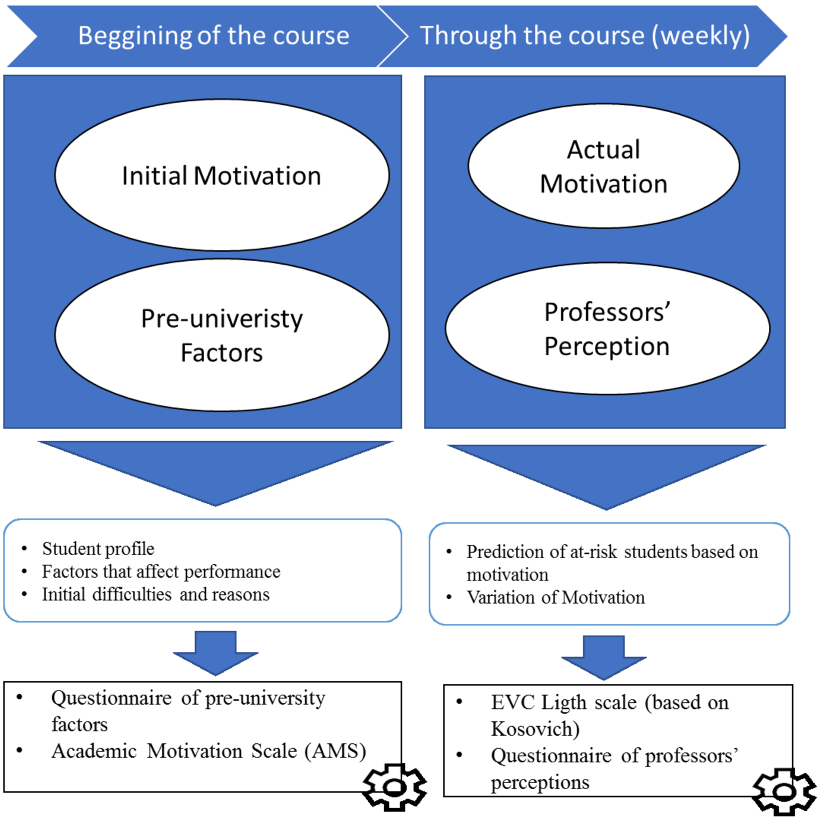
.

5.1 RESEARCH METHOD

This chapter aims to demonstrate a method for prediction of student performance, which uses four main components: initial motivation, pre-university, motivation factors along the course, and professor's perception.

Figure 18 shows an overview of the proposed method, containing the stages in the process, the information gathered, and tools used.

Figure 18: General scheme of the proposed method



Source: Developed by the author

One of the goals of the proposed method is that it is possible to predict the student outcome based on information and motivation factors. To identify these factors, we proposed to use instruments to collect these factors in two different times (detailed in Chapter 4):

1. At the beginning of the course, aiming to identify some demographic data and pre-university factors that can impact on motivation, engagement and student outcome. In addition, it must be identified the initial motivation of students, that is, the reasons why they chose to take the course;
2. Through the course (weekly), aiming to identify the student's motivation, based on their perception of expectation, value, and cost in the course;
    a) Through the course (weekly), aiming to identify the perception of the professor for each student, according to his motivation, interaction, and interest.

To perform these steps, four approaches were used for data collection:

1) Initial factors questionnaire (beginning of the course): pre-university factors (26 items) based on (SCHOEFFEL, WAZLAWICK, *et al.*, 2018b) and initial motivation (Academic Motivation Scale - AMS) (four items) based on (VALLERAND, 1992). Table 57 shows the attributes related to the initial factors. The entire questionnaire can be shown in Appendix E.

Table 57: Pre-university input attributes

Group	Factor
1. Personal and demographic data	Gender Quota Entrance order - entrance exam position Age Entrance way
2. Taste and knowledge of the area	Taste1 – taste for programming Taste2 – taste for informatics Taste3 – taste for games Know1 – knowledge about the undergraduate program goals Know2 – knowledge about the undergraduate program content Know3 – correct perception about computing professionals
3.Computing and programming experience	Know4 – knowledge, and experience in computing Know5 – knowledge, and experience in computer programming Know6 – programming experience in high school
4. Prior school performance	Know7 – general educational performance Know8 – prior math performance
5. Reasons to choose computing	Reason1 – program content Reason2 – interest in learning Reason3 – career/job Reason4 – parents influence Reason5 – friends influence

	Reason6 – challenge to success Reason7 – challenges of computing area Reason8 – lack of other options Reason9 – not having passed another course Reason10 – family pressure
5. Initial motivation	IntrM – intrinsic motivation ExtrM – extrinsic motivation Amot – amotivation or lack of motivation IMIndex – initial motivation index

Source: Developed by the author

2) Questionnaire of the motivation in the course (weekly throughout the semester): EVC Light scale (Expectation-value-cost) (ten items) based on (KOSOVICH, HULLEMAN, *et al.*, 2014) (Table 58). The entire questionnaire can be shown in Appendix D;

Table 58: Longitudinal motivation input attributes

Group	Factor
EVC items (it1 to it6)	It1 – I am confident that I am going to learn the content and have success in the course; It2 - I think that the course lessons are useful for the program; It3 - the course requires a lot of time and I am having to give up doing other activities; It4 - I am understanding and learning the course lessons; It5 - I am understanding the importance of this course in the program; It6 - I am not getting time to dedicate myself to this course.
EVC indexes	Expect - expectancy index Cost - cost index Value - value index EVC - EVC index

Source: Developed by the author

3) Perception - the perception of the professor (weekly during the semester), Likert scale from -3 to +3, indicating positive or negative evolution of the student each week. The entire questionnaire can be shown in Appendix F.

The dependent variable (output) used in predictions was the final student outcome (pass or fail) in the course since the objective of this work is to identify ways to increase student success in introductory computing courses.

The prediction process involved three main phases:

Features selection: originally there were 43 features. In order to deal with this number to optimize the sorting algorithms three different approaches were used:

a) All attributes: keeping all original 43 attributes;
b) Pre-selection: reducing the dataset before the classification (both training and testing data), using a mix of features selection algorithms;

c) Post-selection: applying two features selection algorithms only in the training dataset.

For the pre-selection approach, we use seven features selection algorithms available on Weka: CorrelationAttributeEval, CfsSubsetEval, InfoGainAttributeEvbal, ReliefAttributeEval, OneRAttributeEval, SymmetricalUncertAttributeEval, and GainRatioAttribute:

a) CfsSubsetEval: Values subsets that correlate highly with the class value and low correlation with each other;
b) CorrelationAttributeEval: Evaluates the worth of an attribute by measuring the correlation (Pearson's) between it and the class;
c) GainRatioAttributeEval: Evaluates the worth of an attribute by measuring the gain ratio with respect to the class;
d) InfoGainAttributeEval: Evaluates the worth of an attribute by measuring the information gain with respect to the class;
e) OneRAttributeEval: Evaluates the worth of an attribute by using the OneR classifier;
f) ReliefFAttributeEval: Evaluates the worth of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different class;
g) SymmetricalUncertAttributeEval: Evaluates the worth of an attribute by measuring the symmetrical uncertainty with respect to the class.

Table 59 shows the results with the selected features for each selection algorithm.

Table 59: Features selection results ranked by importance

CfsSub SetEval	**Correlation Attribute Eval**	**Gain ratio**	**Info Gain**	**OneR**	**Relief**	**Symmetrical**
Amot	know8	know1	know8	age	IntrM	reason1
age	taste1	taste2	age	Amot	ExtrM	know8
taste1	reason1	reason1	reason1	reason1	IM Index	age
know5	know7	It3	It4	know5	know8	cost
know7	Amot	cost	Amot	taste1	Entrance way	It3
know8	know5	know8	cost	know6	taste1	It4
reason1	age	age	taste1	know8	know3	Amot
It3	expect	It4	know5	know7	Amot	taste1
It4	It4	know7	It3	expect	know1	know7
cost	cost	Amot	know7	reason7	know5	know5
	It3	taste1	know1	IntrM	age	know1
	It1	know5	Entrance Order	reason2	know2	taste2

	It6	entrada	taste2	IM Index	It1	entrada
	evc	Entrance Order	entrada	It4	expect	Entrance Order
	entrada	gender	gender	gender	reason8	gender
	ExtrM	quota	quota	reason3	reason4	quota
	reason2			know4	reason1	
	know6			taste3	Entrance Order	
	Entrance Order			reason8	gender	
	gender			It2	It4	

Source: Developed by the author

We ranked the attributes according to the frequency of occurrence for each algorithm, using the technique of frequency distribution. So we selected the ten attributes that appeared in most of the seven algorithms (more than 3), as can be seen in Table 60.

Table 60: Frequency distribution of feature selection

Attribute	Frequency	Description
know8	7	Prior math performance
Amot	7	Amotivation index (initial motivation)
know5	6	Knowledge and experience in computing
taste1	6	Taste for programming
age	6	Student age
reason1	6	Program content (reason to choose computing program)
know7	6	General educational performance
it4	5	I am understanding and learning the course lessons (expectancy)
cost	5	Cost index (student perception of course cost)
it3	4	The course requires a lot of time and I am having to give up doing other activities (cost)

Source: Developed by the author

For the post-selection approach, we use two features selection algorithms only on the training data, using the classifier “FilterClassifier” on Weka. We used the CorrelationAttributeEval and InfoGainAttributeEval because they were the algorithms that had more attributes selected among the most frequent.

In order to validate if the selection of attributes is more efficient than the original attributes, we simulated and compared the results with selected attributes and with all attributes, which will be detailed in section 5.2.

Preparation: to train the classification algorithms, we applied the technique of cross-validation with 10 folds. To ensure the best dataset, we performed a step of preprocessing, creating different datasets according to five characteristics: phase of the study, temporal aggregation, features pre-selection, features post-selection, and data balancing (Table 61).

Table 61: Criteria for dataset variations

Filters	Datasets	
Phase of study	Freshman: only students in the first phase/period	Veterans: students in the second, third, and fourth phases/periods
Temporal aggregation	Average values: motivation and professors' perception data are aggregated considering all weeks past (the data was grouped by student, maintaining a single record for each student and averaging of temporal values)	Original values: we consider only the motivation data related to the specific week
Features pre-selection	All attributes: all the 43 original attributes are retained	Selected attributes: only the 10 selected attributes are retained
Features post-selection	Filter all: all attributes from dataset after pre-selection	Filtered: applying a selection algorithm on the training dataset.
Dataset balancing	Imbalanced: keeping the original data	Balanced: using balancing techniques to equalize the dataset

Source: Developed by the author

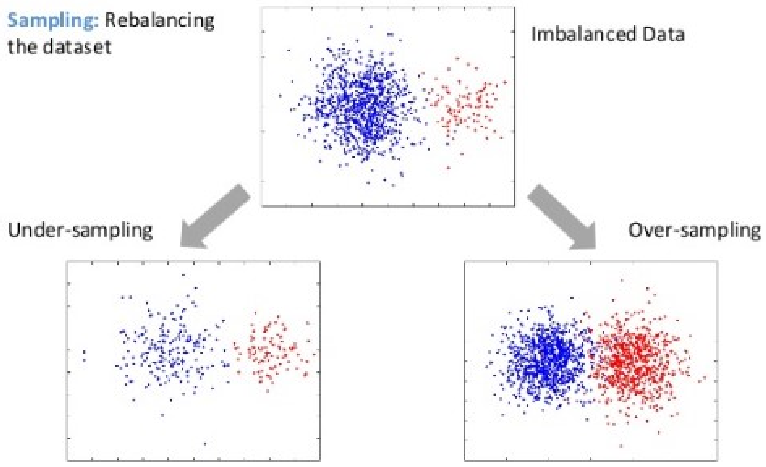
As the dataset used was slightly imbalanced (58% fail and 42% pass), we used some balancing techniques for comparison purposes. We used the oversampling method, applying the technique SMOTE (Synthetic Minority Over-sampling Technique) and resampling, in addition to the undersampling method applying the resampling technique.

The oversampling method, new instances of the minority class are created, in order to balance the dataset (HULSE, KHOSHGOFTAAR and NAPOLITANO, 2007). In our case, new instances of the class "Pass" were created. In the undersampling method, random instances of majority class are eliminated (HULSE, KHOSHGOFTAAR and NAPOLITANO, 2007). In our case instances of the class "Fail" were eliminated.

The resampling technique simply makes copies or removes instances chosen randomly (HULSE, KHOSHGOFTAAR and NAPOLITANO, 2007). With SMOTE, the minority class is over-sampled by creating “synthetic instances” considering the difference between the feature vector (sample) and its nearest neighbor (CHAWLA, BOWYER, *et al.*, 2002).

Figure 19 shows a diagram of the operation of oversampling and undersampling methods.

Figure 19: Balancing data example



Source: (KARAGOD, 2018)

In order that the instances created were not part of the test dataset and did not influence the results, we used the *FilteredClassifier* classifier. This class is a “class for running an arbitrary classifier on data that has been passed through an arbitrary filter. Like the classifier, the structure of the filter is based exclusively on the training data and test instances will be processed by the filter without changing their structure. If unequal instance weights or attribute weights are present, and the filter or the classifier are unable to deal with them, the instances and/or attributes are resampled with replacement based on the weights before they are passed to the filter or the classifier (as appropriate)” [4].

Classification: to perform the prediction we used the following classification algorithms, called as classifiers: NaiveBayes, Multilayer Perceptron, SimpleLogistic, J48, Random Forest, LMT, NetBayes, SMO (SVM), K-Nearest Neighbors (iBk), and AdaBoostM1. These techniques were chosen because they are the most used for similar purposes in related works found in the literature. We use the seed value equal to 2 for random generation of folds. The dependent variable was the student outcome in the course, classified as *pass* or *fail*. Fail students include students that dropped out during the course.

We used the API (Application Programming Interface) of the Weka for applying each of the classifiers. We used the settings as Table 62.

[4] http://weka.sourceforge.net/doc.dev/weka/classifiers/meta/FilteredClassifier.html

Table 62: Classifiers config in Weka

Classifier	Technique	Weka Options
Naive Bayes	Naïve Bayes	
Multilayer Perceptron	Artificial neural nets	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a
Simple Logistic	Linear logistic regression	-I 0 -M 500 -H 50 -W 0.0
J48	C4.5 decision tree	-C 0.25 -M 2
Random Forest	Forest of random trees	numInteractions = 40, numFeatures = 20
LMT	Logistic model trees	-I -1 -M 15 -W 0.0
BayesNet	Bayes Network	-D -Q weka.classifiers.bayes.net. search.local.K2 -- -P 1 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5
SMO	Support vetor machine	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions. supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers. functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
iBk	K Neighboard	-K 1 -W 0 -A "weka.core. neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""
AdaBoostM1	Adaboost M1 method (boosting algorithm)	-W weka.classifiers.trees.DecisionStump numInteractions = 7
AdaBoostM1-SMO	Adaboost M1 method (boosting algorithm)	-W weka.classifiers.functions.SMO numInteractions = 5

Source: Developed by the author

For all the classifiers, except AdaBoostM1 and AdataBoostM1-RandomForest-SMO, initial WEKA standards configuration were maintained. These settings have been changed because they showed superior performances in comparative tests performed for the sample used.

As the main goal of the research is the identification of positive cases (fail), we use as main measure the recall and the f-measure of the fail class. The recall calculates the percentage of true positive in relation to the total of positive results in the sample (1). The f-measure or f-score is defined as the weighted harmonic mean of its precision and recall (2). However, we also calculated and evaluated the accuracy, precision, and ROC Area (AUC).

Recall = (TP / (TP + FN) (1)

where:

TP = true positive;

FN = false negative.

F-measure = (2 * precision * recall) / (precision + recall) (2)

5.2 RESULTS

In total, 245 students participated in the research, as shown in Table 63. The research was conducted in introductory computing courses from four universities, being two public universities and two private universities, belonging to following programs: Software Engineering, Information Systems, Computer Science, Systems for Internet, and Information and Communication Technology. The duration of each course was eighteen weeks in all.

The survey was accomplished during the first half of 2018 and the second half of 2018. In the first half, we have carried out a pilot survey in four classes from one of the universities. In the second half, the research was expanded to other scenarios (programs and universities). Of the students who participated, 213 were from the first period (semester), 38 were from the second period (semester), and 38 were from the third period (semester) of the program.

Table 63: Respondents by University

University	Course [a]	Period	Students
A	Introductory Programming	20181	53
A	Programming I	20181	21
A	Programming II	20181	38
A	Discrete Mathematics	20181	43
A	Introductory Programming	20182	47
A	Programming I	20182	17
B	Algorithms	20182	30
C	Algorithms and Programming	20182	16
D	Algorithms and Programming Techniques	20182	24

a. The courses “Introductory Programming”, “Algorithms”, “Algorithms and Programming” and “Algorithms and Programming Techniques” are distinct names but similar to CS101.

Source: Developed by the author

Of students participating in the study, 127 were approved and 118 failed. The high rate of failure and the high rate of students with poor performance, as shown in Table 57, is a common scenario in introductory computing courses, and it corroborates the reason for this work.

Table 64: Students distribution by grade level

Level [a]	n
A	43
B	72
C	32
D	142

a. A = 8.00 a 10.00, B = 6.00 a 7.99, C = 4 a 5.99, D = 0.00 a 3.99

Source: Developed by the author

In Brazil, most undergraduate programs are divided into semesters and not years. Thus, we will consider the students of the second period (semester) onwards as *veterans*.

We identified a significant difference in the distribution of approved and failed students among freshmen and veterans (Table 65). We compared the results between these two groups of students, in order to assess in what situation the proposed method fits better.

Table 65: Outcome comparison between freshman and veterans

	1º period	2º period	3º period
Pass	42.77%	67.57%	80.00%
Fail	57.23%	32.43%	20.00%

Source: Developed by the author

Despite the method having the objective to identify students at risk in the context of introductory computing courses, we performed some pilot tests with subjects of later periods, for comparison purposes. Table 66 shows the average results considering all classifiers and datasets to freshmen (F) and veterans (V) students. We use only the first eight weeks for comparison because the purpose of the method is to identify students at risk in advance, with priority for the first few weeks.

Table 66: Prediction comparison between freshman (F) and veterans (V)

	Average f-measure		Average recall		Average Accuracy		Best f-measure		Best recall	
Week	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**
1	0.415	0.820	0.442	0.785	59.628	77,148	0.615	0.889	0.750	0.925
2	0.420	0.712	0.509	0.688	64.158	66,746	0.609	0.849	0.790	0.957
3	0.397	0.668	0.501	0.676	62.153	63,456	0.596	0.774	0.842	0.871
4	0.407	0.699	0.509	0.678	62.969	66,124	0.625	0.909	0.842	0.889
5	0.394	0.646	0.499	0.632	61.961	63,264	0.625	0.764	0.842	0.785
6	0.386	0.731	0.479	0.704	62.768	70,640	0.596	0.857	0.737	0.838
7	0.409	0.690	0.508	0.674	63.573	67,628	0.667	0.870	0.842	0.909
8	0.408	0.690	0.503	0.654	63.657	66,532	0.583	0.857	0.790	0.840
mean	0.4045	0.7070	0.4937	0.6863	62.608	67.692	0.6145	0.8461	0.8044	0.8768
t		15.946		10.713		3.0612		11.371		2.9019
p		$3.8x10^{-7}$		$6x10^{-7}$		$3.8x10^{-7}$		0.01208		0.01208

Source: Developed by the author

We performed a hypothesis test to verify whether the differences in the results are significant, using the Welch Two Sample t-test. As shown in Table 66, we realized that the results for freshmen are significantly better than veterans. As the number of veterans in the sample is small, the focus of the method is introductory courses, and we identify the method fits better to the freshmen students, we chose to use only the data of the freshmen students.

So, for the following prediction analysis, we used the 173 instances of the freshmen, being 99 failed students and 74 passed students.

We collected weekly data from students' motivation and perception of professors. So, we got different amounts of data collected each week, due to absences, non-availability of the professor, evaluations, etc. Especially in the first two weeks, the collection was less because of institutional activities are performed with the freshmen at the beginning of the program. Table 67 shows the number of student respondents in each week.

Table 67: Respondents for week

Week	Students	%	Accumulated	%
1	60	34.68%	60	34.68%
2	79	45.66%	98	56.65%
3	122	70.52%	156	90.17%
4	29	16.76%	160	92.49%
5	53	30.64%	166	95.95%
6	65	37.57%	169	97.69%
7	23	13.29%	169	97.69%
8	43	24.86%	172	99.42%
10	6	3.47%	172	99.42%
11	35	20.23%	173	100.00%
12	5	2.89%	173	100.00%
15	19	10.98%	173	100.00%
16	16	9.25%	173	100.00%
17	12	6.94%	173	100.00%

Source: Developed by the author

As the purpose of the method is the early prediction of students at risk, the data collects were higher for the first four weeks. From fifth to eighth week the data collects were held fortnightly. After the ninth week, the data collection was optional and spontaneous.

To evaluate each of the remaining data filtering strategies, we have done simulations of all possible combinations, considering data aggregation (avg), features pre-selection (selected), and temporal variation (var) (Table 68).

Table 68: Results of combinations between datasets and classifiers

Selected	Avg	Var	F-score	Accuracy	Recall	Best F-score	Best accuracy	Best recall
No	No	No	0.752	72.044	0.744	0.788	75.723	0.798
No	No	Yes	0.748	71.571	0.736	0.800	78.035	0.778
No	Yes	No	0.752	72.044	0.744	0.788	75.723	0.798
No	Yes	Yes	0.748	71.571	0.736	0.800	78.035	0.778
Yes	No	No	0.742	70.836	0.738	0.784	74.567	0.869
Yes	No	Yes	0.742	70.836	0.738	0.784	74.567	0.869
Yes	Yes	No	0.742	70.836	0.738	0.784	74.567	0.869
Yes	Yes	Yes	0.742	70.836	0.738	0.784	74.567	0.869

Source: Developed by the author

We can perceive some slight variations, such as, for example, that the values of f1-measure and accuracy are better for all attributes (no selected). However, in general, there is not a significant variation according to each feature and their combinations.

With respect to the other combinations of datasets, we calculated the variance for each variable or combination of variables, using multi-way ANOVA. We found no significant variance, except for the best accuracy, which is influenced by the use of mean values, as shown in Table 69.

Table 69: ANOVA results of dataset combinations related to the best accuracy value

	Sum Sq	Df	F-value	Pr(>F)
AVGValues	174.46	1	62.206	0.0169
AttributeSelection	0.84	1	0.0299	0.8637
WeeklyVariation	4.05	1	0.1445	0.7058
AVGValues: AttributeSelection	0.68	1	0.0241	0.8774
AVGValues: WeeklyVariation		0		
AttributeSelection: WeeklyVariation		0		
AVGValues:selected:variation		0		
Residuals	1121.81	40		

Source: Developed by the author

We also analyzed the question of the balance of the data. Although our sample is not significantly imbalanced (99 failed and 74 passed), we did not find in literature a minimum value or limit to balance or not the data. Although some authors assert that dataset balancing is necessary when data classes are heavily imbalanced (KLEMENT, WILK, *et al.*, 2009) (PRATI, BATISTA and MONARD, 2009), we decided to do some simulations to verify the behavior of classifications.

We utilized three data balancing techniques, being one of oversampling and two of undersampling. For undersampling technique, we used the Resample filter of the Weka, considering the parameters "biastoUniformeClass" = 1.0 and "sampleSizerPercent = 86.25, in order to reduce the majority class (pass) to the same number of minority class instances (fail). For the oversampling, we also used the Resample filter of Weka, however with the parameter "sampleSizerPercent" set to create new instances of the minority interest (fail) until it reaches the number of majority class (pass). We also used the algorithm SMOTE (Synthetic Minority Oversampling Technique), that at a high level, creates synthetic instances of the minority class by finding the k-nearest-neighbors for minority class (finding similar observations), randomly choosing one of the k-nearest-neighbors, and using it to create a similar, but randomly tweaked, new instance (CHAWLA, BOWYER, *et al.*, 2002).

We used the FilterClassifier classifier of Weka, which allows the balancing to be applied only in the training set, keeping the test data preserved. Table 70 shows a comparison of the first eight weeks between datasets without balancing (No balanced) and the datasets with balancing, considering the average recall and better recall value. In order to verify if the results obtained with the balanced dataset were better or worse than the results with the imbalanced dataset, we did a hypotheses test using the non-parametric Wilcoxon test for paired samples

Table 70: Comparison of data balancing strategies (recall)

	Average Recall				Best Recall			
week	**No balanced**	**Resample Over**	**Resample Under**	**SMOTE**	**No balanced**	**Resample Over**	**Resample Under**	**SMOTE**
1	0.820	0.785	0.744	0.791	0.925	0.875	0.800	0.875
2	0.743	0.657	0.657	0.698	0.957	0.746	0.745	0.830
3	0.720	0.637	0.664	0.685	0.855	0.761	0.871	0.839
4	0.729	0.665	0.633	0.688	0.889	0.833	0.778	0.889
5	0.669	0.606	0.628	0.624	0.785	0.720	0.769	0.769
6	0.752	0.676	0.681	0.710	0.838	0.823	0.811	0.811
7	0.700	0.701	0.642	0.644	0.823	0.909	0.818	0.818
8	0.714	0.615	0.631	0.655	0.806	0.786	0.760	0.840
avg	0.731	0.668	0.660	0.687	0.860	0.807	0.794	0.834
sd	0.044	0.057	0.039	0.051	0.060	0.065	0.040	0.037
	V	35	36	36	-	30	33	23
	p-value	0.016	0.008	0.008	-	0.109	0.039	0.016
median difference		0.070	0.073	0.043	-	0.053	0.037	0.151

Source: Developed by the author

We realized that the results with the balanced datasets were worse than results with the imbalanced datasets, mainly in the first weeks. Considering the average recall, only in week 7 the balanced dataset using Resample Over outperformed the unbalanced dataset (Table 70).

We also compared the value of accuracy, to make sure that even having a lower rate of recall, there may have been an improvement in the accuracy. However, we identified that there was no significant improvement. As we can see in Table 70, the accuracy was slightly better for balanced data in a few weeks and on general average, mostly using the technique SMOTE. However, in general, we identified a no statistically significant difference, except for ResampleUnder which underperformed.

Table 71: Comparison of data balancing strategies (accuracy)

week	FilterAll	ResampleOver	ResampleUnder	SMOTE
1	76.631	77.237	77.020	77.929
2	67.610	64.561	67.403	67.676
3	64.248	61.608	63.874	64.347
4	66.458	66.585	64.325	67.533
5	63.795	62.822	63.329	63.051
6	71.941	68.109	70.691	72.293
7	67.837	68.343	67.190	66.949
8	67.724	64.970	66.659	66.871
avg	68.28029	66.779	67.561	68.33
sd	4.205196	4.844	4.507	4.731
	V	30	33	15
	p-value	0.109	0.0391	0.742
median diference		1.806	0.557	-0.083

Source: Developed by the author

Therefore, the balance of the data has not shown to be effective for the sample used. In addition, we performed a test to evaluate if the proportions are significantly different. Applying student's t-test for one proportion (1-sample proportions test without continuity correction), we got X-squared = 3.6127, df = 1, p-value = 0.05734, and thus we have not been able to confirm that the proportions are significantly different. Thus, we decided to include the original data without balancing for the next analysis, in addition to balanced data, for the purposes of comparisons. One possible reason for having worse results with the balanced database is that the original data are slightly imbalanced, and there is no statistically significant difference in the number of instances between classes. Another possible reason is that we have already used stratified cross-validation, which considers the proportionality of the classes. Additionally, a small sample and the creation of synthetic instances can generate noises in the data.

Then we compare the combinations with each of the classifiers used and the different datasets. In total, 75 different combinations of datasets were generated by calculating the average results of all weeks and for each classifier. Table 72 and Table 73 present the most relevant results (best results for each classifier) for the first four and first eight weeks, respectively. In Appendix J are shown the results of all combinations.

Table 72: Comparison between data and classifier combinations considering the first four weeks

Classifier	Dataset	Accuracy	Recall	ROC	f-measure
AdaBoostM1	AllAttributesSMOTEAll	0.690	0.755	0.733	0.744
AdaBoost M1-SMO	AllAttributesFilterAllAll	0.753	0.882	0.772	0.806
BayesNet	SelectedAVGVARSMOTE All	0.701	0.749	0.711	0.748
IBk	SelectedAVGFilterAllAll	0.644	0.732	0.617	0.710
J48	SelectedSMOTE CorrelationAttributeEval	0.680	0.740	0.667	0.731
LMT	SelectedAVGVARFilterAll CorrelationAttributeEval	0.699	0.766	0.746	0.752
Multilayer Perceptron	AllAtributesAVGFilterAll CorrelationAttributeEval	0.719	0.800	0.767	0.772
NaiveBayes	AllAttributesSMOTE InfoGainAttributeEval	0.677	0.711	0.691	0.721
Random Forest	AllAttributesFilterAllAll	0.702	0.793	0.735	0.758
Simple Logistic	SelectedAVGFilterAll InfoGainAttributeEval	0.716	0.782	0.759	0.766
SMO	AllAttributesFilterAll CorrelationAttributeEval	0.776	0.850	0.762	0.816

Source: Developed by the author

As can be seen in Table 72 and Table 73, although there is not a large variation between the results of the different classifiers, the AdaBoostM1-SMO and SMO had the best performances in both settings.

Table 73: Comparison between the best dataset for classifier considering the first eight weeks

Classifier	Dataset	Accuracy	Recall	ROC	f-measure
AdaBoost M1	SelectedAVGFilterAllAll	0.734	0.678	0.771	0.727
AdaBoost M1-SMO	AllAtributesAVGFilterAll All	0.760	0.702	0.819	0.733
BayesNet	SelectedAVGFilterAll CorrelationAttributeEval	0.729	0.681	0.734	0.700
IBk	SelectedAVGVAR FilterAllAll	0.708	0.659	0.712	0.637
J48	SelectedAVGVARFilter AllInfoGainAttributeEval	0.714	0.664	0.724	0.678
LMT	SelectedFreshmenAVG FilterAllAll	0.759	0.714	0.775	0.764
Multilayer Perceptron	AllAtributesAVG FilterAllInfoGainAttribute	0.752	0.706	0.772	0.758
Naive Bayes	AllAttributesFreshmen SMOTEInfoGainAttributeEval	0.739	71.777	0.716	0.732
Random Forest	AllAtributesAVGFilterAll CorrelationAttributeEval	0.748	0.700	0.767	0.767
Simple Logistic	SelectedAVGFilterAll InfoGainAttribututeEval	0.767	0.723	0.784	0.771
SMO	AllAtributesAVGFilterAll CorrelationAttributeEval	0.774	0.720	0.826	0.700

Source: Developed by the author

For comparison purposes, we show the best dataset combination for each classifier as settings in Table 74. We can see that, in general, the best settings of datasets are those that use the average of the weekly values (nine of eleven classifiers), do not consider the weekly variation of motivation (ten of eleven classifiers) and do not use data balancing techniques (eight of eleven classifiers). With respect to pre-selection and post-selection features, we observed different results, depending on the classifier used.

Table 74: Several combinations of datasets

Classifier	Features pre-selection	Average values	Weekly variation	Balancing strategy	Features post-selection [a]
AdaBoost M1	Yes	Yes	No	Filter All	No
AdaBoostM1-SMO	No	Yes	No	Filter All	No
BayesNet	Yes	Yes	Yes	SMOTE	No
IBk	Yes	Yes	No	Filter All	No
J48	Yes	No	No	SMOTE	CAE
LMT	Yes	Yes	No	Filter All	No
Multilayer Perceptron	No	Yes	No	Filter All	IGA
Naive Bayes	No	Yes	No	SMOTE	IGA
Random Forest	No	No	No	Filter All	No
Simple Logistic	Yes	Yes	No	Filter All	IGA
SMO	No	Yes	No	Filter All	CAE

a. IGA = Info Gain Attribute Eval, CAE = Correlation Attribute Eval

Source: Developed by the author

In order to evaluate the performance of the predictions over the weeks, we compared the scenarios with the best average performance of each of the combinations.

In accordance with the comparisons shown in Table 75 and Table 76, the best average results whereas the metric f-measure were obtained using the SMO classifier, with original data (imbalanced), selected attributes by CAE algorithm and grouping the data per student. And the best average results using the metric “recall” were obtained using AdaBoostM1 with SMO classifier.

Table 75: Best f-measure results for each classifier

F-measure	1	2	3	4	5	6	7	8	End	Avg (4w)	Avg (8w)
AdaBoost M1	0.851	0.733	0.677	0.699	0.681	0.733	0.756	0.743	0.733	0.740	0.734
AdaBoost M1 - SMO	0.861	0.760	0.718	0.767	0.724	0.776	0.755	0.723	0.771	0.776	0.760
BayesNet	0.857	0.763	0.685	0.688	0.663	0.698	0.688	0.674	0.732	0.748	0.715
IBk	0.750	0.733	0.681	0.677	0.691	0.663	0.670	0.700	0.758	0.710	0.696
J48	0.857	0.688	0.641	0.737	0.588	0.818	0.500	0.756	0.731	0.731	0.698
LMT	0.800	0.767	0.704	0.745	0.735	0.781	0.776	0.766	0.765	0.754	0.759
Multilayer Perceptron	0.833	0.737	0.750	0.738	0.739	0.744	0.742	0.736	0.735	0.765	0.752

Naive Bayes	0.850	0.736	0.630	0.667	0.612	0.771	0.857	0.792	0.722	0.721	0.739
Random Forest	0.815	0.740	0.687	0.790	0.583	0.816	0.583	0.711	0.742	0.758	0.716
Simple Logistic	0.815	0.783	0.723	0.744	0.749	0.781	0.776	0.766	0.765	0.766	0.767
SMO	0.889	0.756	0.755	0.769	0.737	0.778	0.771	0.735	0.757	0.792	0.774

Source: Developed by the author

Table 76: Best recall results for each classifier

Recall	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**End**	**Avg (4w)**	**Avg (8w)**
AdaBoostM1	0.925	0.746	0.727	0.714	0.699	0.771	0.792	0.796	0.748	0.778	0.771
AdaBoostM1-SMO	0.925	0.831	0.796	0.868	0.774	0.813	0.802	0.745	0.798	0.855	0.819
BayesNet	0.825	0.763	0.705	0.703	0.656	0.688	0.667	0.633	0.677	0.749	0.705
IBk	0.750	0.746	0.727	0.703	0.710	0.656	0.656	0.725	0.808	0.732	0.709
J48	0.825	0.681	0.677	0.778	0.577	0.730	0.455	0.680	0.687	0.740	0.675
LMT	0.800	0.780	0.716	0.769	0.774	0.781	0.792	0.786	0.788	0.766	0.775
MultilayerPerceptron	0.875	0.712	0.784	0.758	0.731	0.771	0.792	0.755	0.758	0.782	0.772
Naive Bayes	0.850	0.681	0.645	0.667	0.577	0.730	0.818	0.760	0.657	0.711	0.716
Random Forest	0.825	0.787	0.726	0.833	0.539	0.838	0.636	0.640	0.727	0.793	0.728
Simple Logistic	0.825	0.797	0.727	0.780	0.785	0.781	0.792	0.786	0.788	0.782	0.784
SMO	0.900	0.814	0.841	0.857	0.785	0.823	0.823	0.765	0.788	0.853	0.826

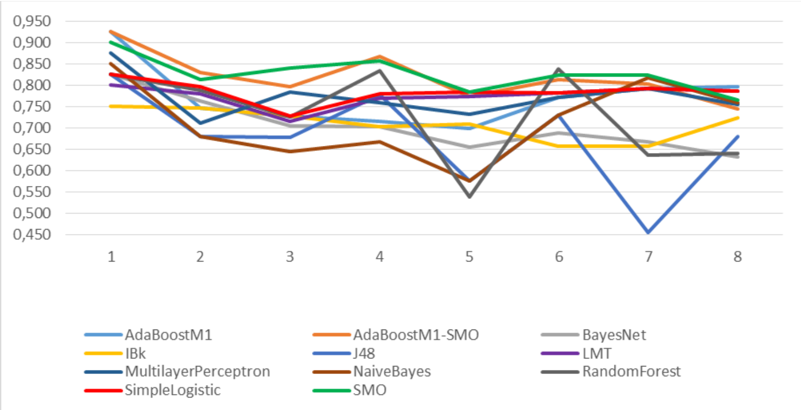
Source: Developed by the author

We can observe in Table 75 and Table 76 that, in general, the results are better in the first week, keeping constant until week eight.

Figure 20 and Figure 21 show the weekly results for each combination (classifier and dataset) whereas the metrics recall and f-measure, respectively. It is interesting to notice that combinations using only the motivational data of the current week (J48 and RandomForest) had a greater variation, despite having the best results in a couple of weeks. That's predictable, given that the results are very sensitive to changes in opinions and behaviors of students weekly. The combinations that use the average of the data until the current week motivation had a more stable behavior. This behavior was expected because when considering the average every week we have reduced the impact of temporary variation of weekly data (motivation and perception of the professor).

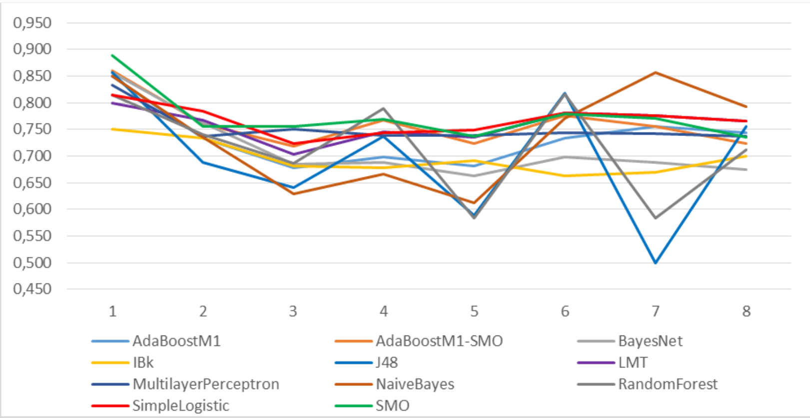
Although simulations have been carried out using the week variation of motivation as an input attribute, just to try to avoid this bias, this strategy proved not to be effective.

Figure 20: Weekly results (recall) of the best combination (classifier and dataset)



Source: Developed by the author

Figure 21: Weekly results (f-measure) of the best combination (classifier and dataset)



Source: Developed by the author

Surprisingly, the first two weeks had high rates of recall, falling a little in the next week, but maintaining a good rate until the end of the period. In general, the datasets with all attributes had better results, except for the classifier Simple Logistic. We can see that the iBk, J48, BayesNet and Naïve Bayes classifiers performed worse, close to 70% on average. The best performing combination of RandomForest had a very wide variation from week to week, as it did not use data aggregation. Of the others, we then selected the top four performers, considering the average recall of the first four and eight weeks.

We resume the best four combinations in Table 77 for all sixteen weeks.

Table 77: Weekly results of the three best sets of classifiers and datasets

	AdaBoostM1		Multilayer		SimpleLogistic		SMO	
Week	**Recall**	**f-score**	**Recall**	**f-score**	**Recall**	**f-score**	**Recall**	**f-score**
1	0.925	0.861	0.875	0.833	0.825	0.815	0.900	0.889
2	0.831	0.760	0.712	0.737	0.797	0.783	0.814	0.756
3	0.796	0.718	0.784	0.750	0.727	0.723	0.841	0.755
4	0.868	0.767	0.758	0.738	0.780	0.744	0.857	0.769

5	0.774	0.724	0.731	0.739	0.785	0.749	0.785	0.737
6	0.813	0.776	0.771	0.744	0.781	0.781	0.823	0.778
7	0.802	0.755	0.792	0.742	0.792	0.776	0.823	0.771
8	0.745	0.723	0.755	0.736	0.786	0.766	0.765	0.735
9	0.765	0.732	0.745	0.730	0.786	0.766	0.776	0.742
10	0.786	0.740	0.745	0.734	0.796	0.772	0.776	0.738
11	0.788	0.757	0.758	0.743	0.788	0.776	0.778	0.744
12	0.798	0.760	0.768	0.742	0.778	0.770	0.778	0.748
13	0.778	0.751	0.768	0.738	0.778	0.770	0.778	0.751
14	0.828	0.770	0.788	0.757	0.778	0.770	0.788	0.746
15	0.758	0.739	0.748	0.726	0.808	0.784	0.788	0.757
16	0.778	0.748	0.748	0.733	0.808	0.781	0.788	0.754
Avg	0.798	0.771	0.765	0.745	0.788	0.765	0.788	0.757
4w [a]	0.855	0.776	0.782	0.765	0.782	0.766	0.853	0.792

a. 4w = average of the first four weeks

Source: Developed by the author

In order to check if any of the four classifiers is significantly better than another, we used the Friedman test rank-sum test. The result (Friedman chi-squared = 19,808, df = 3, p-value = 0.000186) shows that we have at least one classifier with a significant difference. To check which classifier can be considered best, we applied the Wilcoxon test for paired samples, as shown in Table 78.

Table 78: Wilcoxon test to verify the best classifier

	AdaBoostM1-SMO	**Multilayer Perceptron**	**Simple Logistic**
AdaBoostM1-SMO	-		
Multilayer Perceptron	0.0007 (v=133.5) md = 0.03	-	
Simple Logistic	0.3002 (v=69.5) md = 0.005	0.0404 (v=23.5) md = -0.026	-
SMO	0.7758 (v=54.5) md = -0.005	0.0007 (v=120) md = -0.031	0.2933 (v=61) md = 0

Source: Developed by the author

These results demonstrate that the combination that uses the classifier "Multilayer Perceptron" has significantly lower results than the others. The combinations that use the classifiers "AdaBoostM1-SMO, Simple Logistic and SMO" did not have significant differences and can be considered similar.

Summarizing, Table 79 shows the results of three best combinations of classifiers and datasets, excluding Multilayer Perceptron.

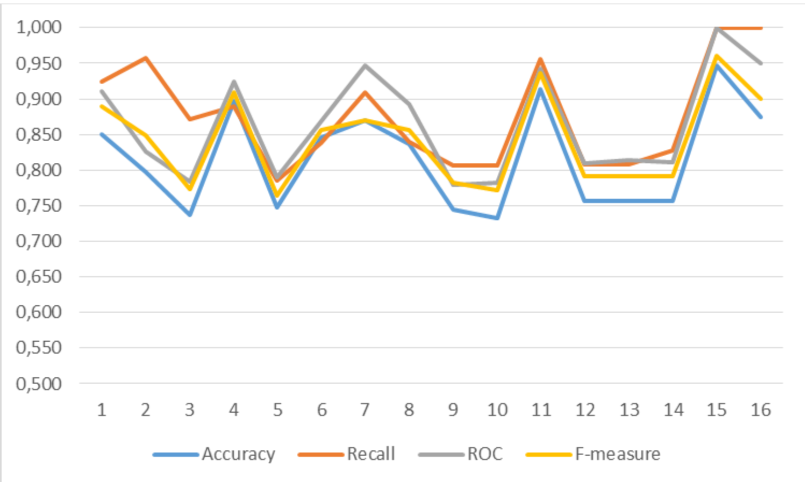
Table 79: Summary of the best classifiers

	AdaBoostM1	SimpleLogistic	SMO
	Recall		
4w	0.855	0.782	0.853
8w	0.819	0.784	0.826
All	0.798	0.788	0.788
	F-measure		
4w	0.776	0.766	0.792
8w	0.760	0.767	0.774
All	0.771	0.765	0.757
	Accuracy		
4w	70.778	71.609	73.306
8w	70.151	72.338	72.008
All	72.832	72.254	71.098
	AUC (ROC Area)		
4w	0.747	0.759	0.707
8w	0.733	0.771	0.700
All	0.749	0.814	0.698

Source: Developed by the author

If we selected the best results of each week, we would get a recall average of 0.874 and an f-measure average of 0.831, as weekly values shown in Figure 22.

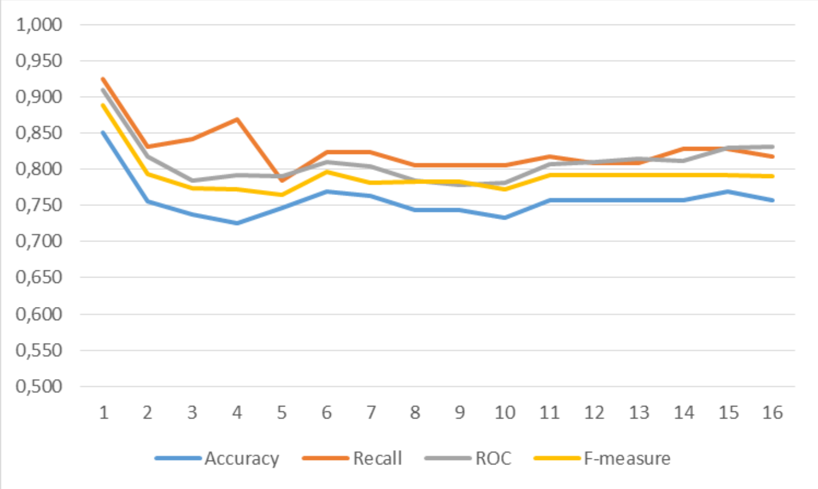
Figure 22: The best weekly results (all combinations)



Source: Developed by the author

As you can see in Figure 22, there is a significant variation considering the best values in each week. This is because datasets that consider only the last week and not the average of all the weeks suffer significant variations in the results. Possibly this is due to the different moments of the student in the context of the course. Isolating the best results only from datasets that consider the average of motivation, we obtained values a bit smaller, but more stable, as shown in Figure 23.

Figure 23: The best weekly results (only datasets with the weekly average)



Source: Developed by the author

As we intend weekly data collection and that an estimate for all students must be generated, regardless of whether they participated in the last evaluation, the methods that use the average of all the weeks are more suitable. In addition, this strategy proved to have more stable results.

5.3 DISCUSSION

Regarding RQ1, we have identified that the model identified more than 90% of the failing students in the first week. Although this index has reduced slightly after that, the value of recall remained near or above 80% in the other weeks.

We have identified that the average accuracy obtained in this work (72.83%) was tightly higher than the average of the related work using only demographic data, pre-university, and psychometrics. The average accuracy of the works reported by Shahiri, Husain and Rashid (2015) was 67.57% and the average accuracy of the works described in systematic mapping described in section 3.2.2 was 72.00%.

If we consider the recall (failed students identified correctly), the results of this work are similar or better to related studies found, even these using attributes based on grades, assessments performance and interaction with LMS, as shown in Table 16. Considering the best combination of classifier and dataset (SMO AllAttributesAVGFilterAll), we have achieved a recall average of 80.03%. If we select the best values every week, we have achieved a recall average of 87.4%.

In addition, a very positive result is that the identification of students at risk had good performance since the first few weeks, reaching 90% in the first week.

It is interesting to note that the performance of the prediction had a little drop between the third and fifth weeks. A supposed reason for this decrease in performance is that the expectancy and value perception of students had a reversal, comparing approved and failing students. From the second week, we identified that approved students reduce their expectations and increase their perception of cost, while the inverse occurs for failing students. Only from the fourth week that failing students tend to reduce their expectations and increase significantly their cost perception. A possible cause for this is the moment of the first evaluations.

With respect to the RQ2, the prediction of failing students (recall) had better results for freshmen, possibly because the failure rate is significantly lower in veteran students. There were differences in the results using selected attributes or all attributes. Depending on the classifier and the time (week) the results presented variation.

Comparing classification algorithms (classifiers), various classifiers had similar performance, varying according to the datasets and metrics used. As the main metric adopted by this study was the recall, and the purpose of the method is to identify students at risk in advance, the SMO and AdaBoostM1 algorithms had the best performances.

5.4 CONCLUSION

This chapter described the application of EMMECS method to identify students at risk in advance, in introductory computing courses. The method is based on questionnaires that measure pre-university factors, initial motivation, motivation along the course and perception of the professor.

The proposed method was satisfactorily efficient for the context applied, identifying correctly around 90% of the students failed in the first week and keeping results around 80% until the end of the course. These findings are similar to results of related works that use input attributes as grades and assessments. However, the advantages of the EMMECS is that its application is simple and fast, it is possible to predict at-risk students since the first few weeks, and it allows replication independent of the course context or specific tools.

6 EMMECS APPLICATION

In this chapter are present the process and tools for the use of EMMECS, as a pilot case study of applying the method in a class of introduction to programming. The chapter is divided into two sections. In Section 6.1 it is shown the structure of the framework, describing its components and instructions for use. In Section 6.2 we describe the implementation and evaluation of a pilot case study.

6.1 FRAMEWORK SPECIFICATION

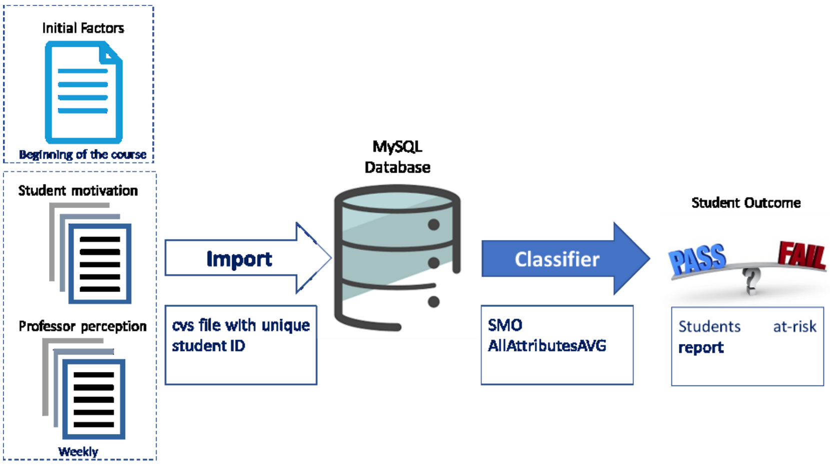
The proposed method was shown to be effective in predicting at-risk students in introductory computing courses. However, it is important that these results and the information collections are useful and are available for professors and coordinators, in order to carry out the necessary interventions that they consider necessary and appropriate.

For this, we present here a proposal for a framework that indicates how to collect, integrate and get the key information, some indicators, and computational tools in order to facilitate the application of professors and educational managers.

6.1.1 Framework Design

The EMMECS framework defines what tools and procedures must be used to collect and integrate data from the EMMECS method, and which data types should be used to generate relevant information for decision making. Figure 24 shows an overview of the EMMECS framework schema.

Figure 24: EMMECS framework schema



Source: Developed by the author

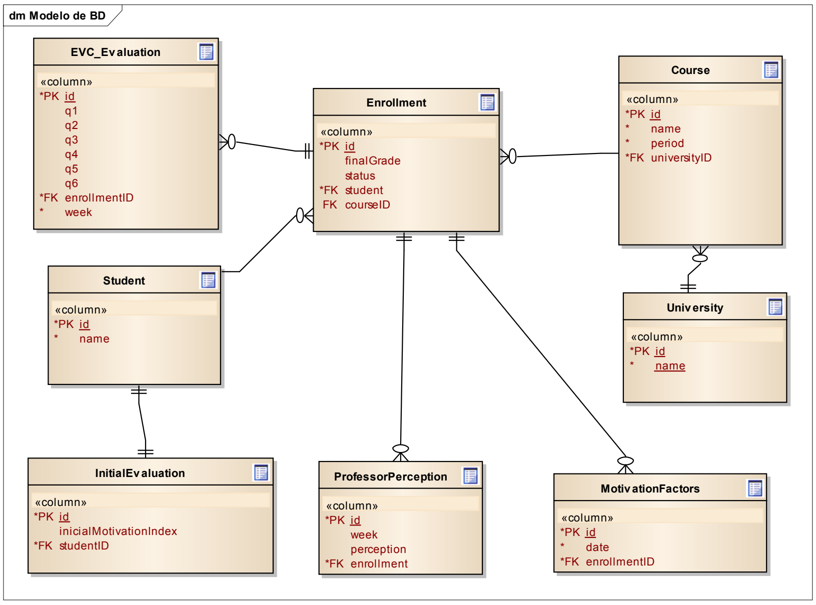
The framework consists of three files of data gathering, which will provide the information of students and to be used as input features to the prediction model. The structure of each of the three files is available on EMMECS website[5]. All files must have a unique identification of the student, which will be used to relate data from different sources.

Each of the files must be imported into the MySQL database or generated a single integrated file in CSV format (Comma Separated Value) or generated directly the ARFF file, for import into Weka tool.

In the case of use of MySQL database structure, we created scripts for importing from CSV files and generating new SQL scripts for import into the database.

The entity-relationship model (ERM) of the database is divided into two main schemes. One of them has the function of saving the data collected by EMMECS instruments (Figure 25). The model consists of tables that store student enrollment data (*Enrollment*, *Student*, *Course*, and *University*) and tables that store data collected by the EMMECS method (*InitialEvaluation*, *EVC_Evaluation*, *ProfessorPerception*, and *MotivationFactors*). The main table that joins the two structures is the Enrollment table.

Figure 25: ER Model of data gathering



a. We hide some fields from some tables to make it easier to visualize

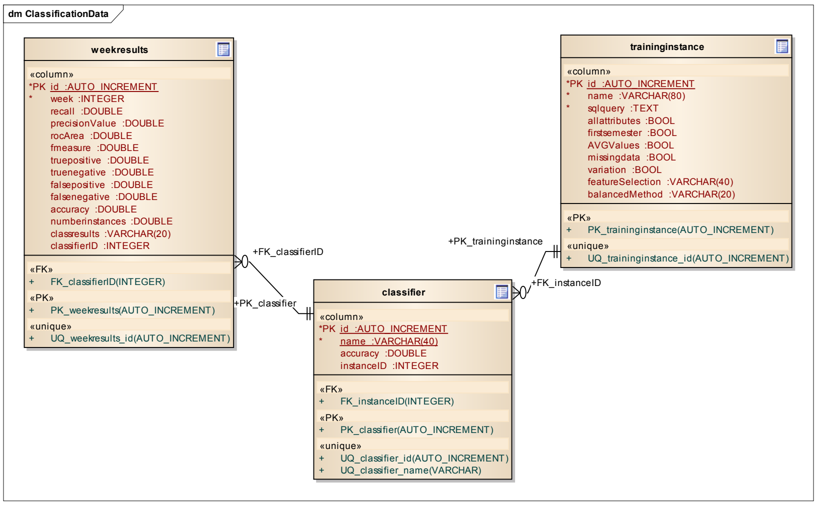
Source: Developed by the author

The other scheme is used to store the results of the predictions (Figure 26), mainly used in validating the method, in order to be able to compare the results of different

[5] http://pablo.pro.br/emmecs

combinations. In this structure are stored each combination of training datasets (*trainninginstance*). For each combination, are stored the classifiers (*classifier*) and their results by week (*weekresults*), for each of the classes (Fail and Pass).

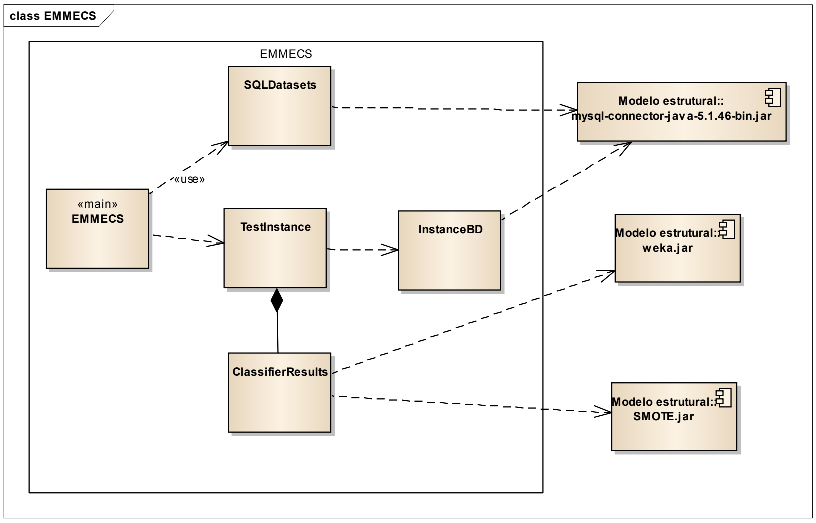
Figure 26: ER Model of prediction results



Source: Developed by the author

In order to automate the classification process to validate the model, because it was generated hundreds of different combinations, we created an application in Java. The application queries the data collected in the database and generates the classification for each different combination of classifiers, dataset, settings, and week. Figure 27 shows a diagram with the architecture and application components.

Figure 27: Components and classes diagram of EMMECS application



Source: Developed by the author

We used the Weka API (*weka.jar*) and also a package specific to balancing technique SMOTE (*SMOTE.jar*), that is not part of the standard package. The *SQLDatasets* class has the scripts to retrieve the training and testing data from the database. *TestInstance* and *InstanceDB* classes are responsible for defining and recording the structure of datasets combinations, the classifiers to be used, and the prediction results. The *ClassifierResults* class is responsible for the classification and getting the results.

The main method for training is the BuildInstance, which should be called for each dataset and distinctive setting. This method gets the settings and raises the whole process until the recording in the database. Table 80 describes the required parameters for this method.

Table 80: Parameters of BuildInstance method

Parameter	Description
java.util.List<TestIntance> pinstancesList	List in which is stored the instance generated
Map<String,String> pfeaturesFilters	Set of features selection filters to be applied
Map<String,String> pbalancedFilters	Set of balancing techniques to be applied
InstanceQuery pquery	SQL script for fetching instances to be trained
String pname	Name of the instance
boolean pFirstSemester	If it considers only freshmen classes
boolean pSelectedAttributes	If it uses a pre-selected set of features
boolean pAVG	If it calculates the average of weekly data
boolean pMissingData	If it considers instances with missing data
boolean pConsiderVariation	If it includes the weekly variation of motivation as an input feature
boolean pTest	If it generates results for instances of tests (false if training)

Source: Developed by the author

The source code of the application, a database backup used (without identifying students, courses, and universities), the file ARFF with all data used for training and testing and all the models of the collection instruments are available for download at http://pablo.pro.br/emmecs.

6.1.2 Operations Manual

We describe, in this section, a step by step implementation of the method, considering the process, instruments, and tools, in order to assist researchers and professors who want to use the method.

Step 1. Elect the class where the method will be applied. It is important to mention that the proposed method has been validated only for introductory computing courses.

Step 2. On the first class, apply the instrument to assess the initial motivation and pre-university factors. The questionnaire can be seen in Appendix E.

Step 3. Weekly, including the first week and with priority for the first four weeks, two instruments must be applied: i) questionnaire EVC Light for evaluation of the students' motivation; ii) questionnaire for evaluation of professor's perception. The EVC Light questionnaire can be seen in Appendix D and the professor's perception questionnaire can be seen in Appendix F.

Step 4. Import the data into the database or single file. The ARFF file structure for use in the Weka is available in Appendix G.

Step 5. Weekly or as required, perform the classification, taking as input the features collected. Use the EMMECS Java application or the tool Weka (in case of using the ARFF file) and then the prediction is calculated for each student, including the predicted class (pass or fail) and the probability. If it is used the EMMECS tool, it will generate several other relevant information, according to the report presented in Appendix H. So, this process can be performed in two ways:

a. Use of EMMECS framework: import in database MySQL (using import worksheets available) and run EMMECS application for training and testing;
b. Creation of the ARFF file: run training and tests in Weka tool, using the classification model available.

Step 6. When the professor or researcher think to be necessary, the instrument to identify the motivation factors must be applied. This instrument serves to improve the diagnosis and to direct interventions. It may be carried out exclusively with students at risk or with all students. The questionnaire for the identification of the motivation factors is available in Appendix C.

Steps 3, 4, and 5 are iterative and performed weekly. Step 5 can be performed when the professor believes it to be necessary.

For reporting the results, we offer a semi-automated worksheet formatted for the reports. In case of not using the EMMECS framework, the data need to be reported manually. Otherwise, we have a script to extract the relevant data and import into the worksheet provided.

6.2 APPLICATION AND EVALUATION OF THE PREDICTION TOOL

In order to validate the method and the proposed framework, a pilot case study was conducted in a class with 34 students in an introductory programming course of Bachelor of Software Engineering program, in the first half of 2019. The objective of this pilot was to evaluate the framework, the viability, and the utility of the method. To do that, we collected opinions and conducted interviews with the professor.

The pre-university and initial motivation questionnaire was applied in the first week of the course. Motivation questionnaires (EVC Light) and perception of the professor were applied in the first four weeks because the goal was to identify students at risk in advance so that the professor can perform interventions.

In the second and fourth week, the professor received reports with key indicators collected and the prediction of the student's outcome generated by the method. Figure 28 shows the general report of the class, composed of key indexes for each student.

Figure 28: Report of EMMECS results by class

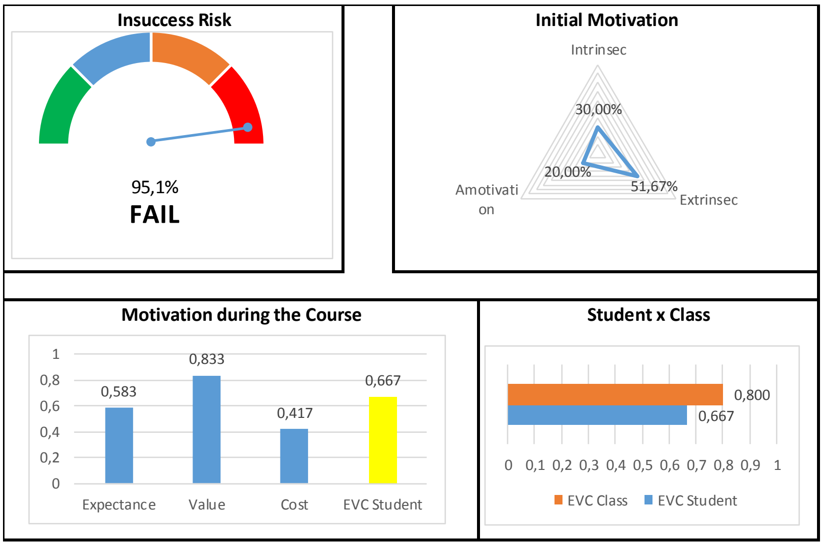
Aluno	Prediction	Insucess Risk	SCALE index	SCALE intrinsic	% Intrinsic	SCALE extrinsec	% Extrinsic	SCALE amotivation	% Amotivation	Expect.	Value	Cost	EVC Index
Aluno 1	FAIL	95,14%	12.333	18	0,300	31	0,517	4	0,200	0,583	0,833	0,417	0,667
Aluno 2	FAIL	99,22%	23.333	38	0,633	47	0,783	5	0,250	0,583	0,333	0,500	0,472
Aluno 3	**PASS**	37,82%	24.000	42	0,700	42	0,700	4	0,200	1,000	1,000	0,917	0,694
Aluno 4	FAIL	94,69%	23.000	35	0,583	46	0,767	4	0,200	0,500	0,917	0,250	0,722
Aluno 5	FAIL	95,62%	25.333	42	0,700	49	0,817	5	0,250	0,750	0,750	0,667	0,611
Aluno 6	FAIL	50,81%	26.667	44	0,733	48	0,800	4	0,200	0,833	1,000	0,583	0,750
Aluno 7	PASS	29,14%	32.000	60	1,000	60	1,000	8	0,400	1,000	1,000	0,083	0,972
Aluno 8	PASS	4,60%	23.667	52	0,867	31	0,517	4	0,200	0,833	1,000	0,167	0,889
Aluno 9	PASS	12,54%	31.667	57	0,950	50	0,833	4	0,200	1,000	1,000	0,167	0,944
Aluno 10	FAIL	99,36%	22.000	45	0,750	36	0,600	5	0,250	0,667	1,000	0,417	0,750
Aluno 11	PASS	12,27%	33.333	52	0,867	60	1,000	4	0,200	1,000	1,000	0,250	0,917
Aluno 12	PASS	4,60%	35.667	59	0,983	60	1,000	4	0,200	1,000	1,000	0,250	0,917
Aluno 13	FAIL	62,12%	21.667	32	0,533	45	0,750	4	0,200	0,917	0,917	0,667	0,722
Aluno 14	PASS	12,54%	17.667	29	0,483	36	0,600	4	0,200	0,917	0,917	0,000	0,944
Aluno 15	PASS	7,88%	27.333	41	0,683	53	0,883	4	0,200	1,000	1,000	0,250	0,917
Aluno 16	FAIL	50,81%	**28.000**	49	0,817	47	0,783	4	0,200	0,917	1,000	0,667	0,750
Aluno 17	**FAIL**	62,12%	**26.667**	41	0,683	51	0,850	4	0,200	0,833	0,917	0,333	0,806
Aluno 18	FAIL	98,68%	21.667	39	0,650	47	0,783	7	0,350	0,750	0,833	0,417	0,722
Aluno 19	PASS	18,06%	26.000	42	0,700	48	0,800	4	0,200	1,000	1,000	0,083	0,972
Aluno 20	FAIL	62,12%	**21.000**	30	0,500	45	0,750	4	0,200	0,833	0,833	0,500	0,722

Source: Developed by the author

The information in the report includes the prediction of the result (pass or fail), the probability of failure, and the main indexes of students' motivation. For example, in Figure 28, the method predicted the *Aluno1* as an at-risk student, with high probability of failure (95%). Analyzing the indexes, we can realize that the student has low initial motivation, especially intrinsic motivation. This may indicate a lack of knowledge of the field and the program or even the incompatibility of the student with the field. *Aluno4* has a medium initial motivation index, but low expectancy, which may indicate learning difficult.

For each student was raised a dashboard with the main information, in order to facilitate the analysis of the professor, as Figure 29.

Figure 29: Student dashboard generated using EMMECS



Source: Developed by the author

New information available in the dashboard is the comparison of the student motivation with the class motivation. We believe this is important for the teacher can analyze if there are problems related to individual students or factors that might be affecting the whole class. As the purpose of the case study was to evaluate the implementation of the method, and the results were not available until the writing of this work, we do not show here the results of the students' outcome prediction. In addition, the method allowed the professor to do interventions based on the information available, altering the student's outcome.

In order to evaluate the ease of use and usefulness of the method and tools developed, we applied a questionnaire with the course professor adapted of the Technology Acceptance Model (TAM) (DAVIS, 1989).

Table 81 shows the questions and answers obtained, whereas the answer options are a Likert scale of 7 points from -3 to +3 (extremely positive, quite positive, slightly positive, indifferent, slightly negative, quite negative and extremely negative).

Table 81: Evaluation of EMMECS application

	Item	Answer
Ease of use	It was easy for me to interpret the information from EMMECS	+3
	I was able to find easily the information	+3
	I think the information and features available are sufficient for my need	+3
	I think the results are trustworthy	+2
	I think EMMECS can help the professor improve his teaching process	+3
	I thought it was fun and interesting to receive data from EMMECS	+3
Int en	I think the EMMECS method is useful for the computing education community	+2

	I would use EMMECS in my courses to identify students at risk	+2
	I think the other professors of the area would use the EMMECS	+2
	I would suggest the EMMECs for other professors and coordinators	+2

Source: Developed by the author adapted from Davis (1989)

Also, in a descriptive question, the professor mentioned as positive points of his experience with EMMECS "the ease of crossing the information between the professor's perception and the prediction provided by the system" and that "the reports are very good" (our translation[6]).

We also conducted an interview with the professor at the time of delivery of the first report. He reported his first perception of the information received. Follows the transcription of the relevant part of his report (our translation):

“I understood that here still has a variation, it's a model, it's a forecast, but some things like that are cool already see, because these guys like that, like those who are there 60%, 50%, we pull them up easy. And those with a higher rate, we also go after them, pull, send to the monitoring, send to other activities, I think it’s possible. It's a precious tool because we can get these people, right. So cool, mostly I really liked the value information, that they really understood the course's purpose…this to me is amazing because it means they are understanding the importance. Man, what a cool, cool, very cool!”[7]

So, even though it was only a pilot of the application with a single class, we can realize that the acceptance was very positive, having evidence that both the model, as the tool brings relevant information and are useful for the professor.

---

[6] “a facilidade de cruzar as informações entre a percepção do professor com a predição do sistema” “os relatórios são muito bons”

[7] “Entendi que isso aqui ainda tem uma variação, é um modelo, é uma previsão, mas algumas coisas assim é legal já ver, por que esses caras assim, tipo aqueles que estão ali 60%, 50%, a gente puxa eles para cima fácil. E aqueles que estão com uma taxa mais alta, a gente também vai atrás deles, puxa, manda para a monitoria, manda para outras atividades, acho que rola. É uma ferramenta preciosa, porque a gente consegue buscar esse pessoal, né. Então bem legal, principalmente gostei bastante do valor, que eles realmente entenderam a função da disciplina.....isso para mim é sensacional, pois significa que eles estão entendendo a importância. Cara, que show, show, muito legal!”

7 RESULTS

This chapter presents the main results of the work, evaluating the research questions. The chapter is divided into four sections. In Section 7.1 the research questions are answered. In Section 7.2 we show a comparison of the results obtained in this work with the state of the art. In Section 7.3 the publications resulting of this work are described. In Section 7.4 we analyze the threats of the validity of this work.

7.1 RESEARCH QUESTIONS ANALYSIS

In this section we evaluate and answer each of the research questions, analyzing their hypotheses.

RQ01 – Is motivation related to the students' outcome in introductory computing courses?

Corroborating with other studies already carried out we have identified a significant relationship between the motivation and the result of students. In our study, students with positive results reported a higher expectation index ($p = 0.00001$) and lower cost index ($p = 0.0025$), generating a greater motivation EVC index for passed students ($p = 0.0007$).

With that, we rejected the null hypothesis H_{A0} (Motivation have no significant relationship with the outcome of students in introductory computing courses) and we confirm the alternative hypothesis H_{A1} (Motivation have a significant relationship with the outcome of students in introductory computing courses).

RQ02 - Is it possible to detect in advance students at risk according to their motivation?

Using the EMMECS method proposed in this work, we identified three combinations that were more effective: AdaBoostM1 classifier with dataset AllAtributesAVGFilterAllAll, SimpleLogistic classifier with the dataset SelectedAVGFilterAllInfoGainAttributeEval, and SMO classifier with the dataset AllAtributesAVGFilterAllCorrelationAttributeEval. We have identified, in the best scenario, an average of 85.5% of the students failing (using the metric recall) before the fourth week of study, 82.6% before the eighth week and 80.3% in the overall average. The best accuracy average was 72.8% (AdaBoostM1), f-measure 77.07% (AdaBoostM1) and AUC 81.35% (SimpleLogistic).

With that, we rejected the null hypothesis H_{B0} (No prediction methods can identify at least 80% of failure students until de fourth week in introductory computing courses in different institutions) and we confirmed the alternative hypothesis H_{B1} (One or more prediction methods can identify at least 80% of failing students until the fourth week in introductory computing courses in different institutions).

RQ03 – Is there a significant variation in students' motivation during the introductory computing courses?

We analyzed the difference between the subscales index of each week. We found significant variation only between week three and week two ($p = 0.046$) and between the end of the course and week four ($p = 0.003$). Both variations were positive. To analyze the variance in student outcome, we compared the results of passed and failed students. We did not find significant differences or any relevant pattern in the variation. We also analyzed the correlation between the weekly indexes of motivation and the final grade of the students. We identified a medium correlation only between subscale "expectancy" and the final grade, during all periods. In addition, we identified that the change in motivation has not had a significant impact on the prediction of students' outcome. With this, results were controversial and cannot reject the hypothesis H_{CO} (The students' motivation does not change significantly through the course).

RQ04 – Is there motivation factors related to the educational context that impact the students' outcome?

In the context studied, the results suggest that performance is not a decisive factor for the student to be motivated. We believe that performance is a consequence of motivation and not vice versa. In summary, we found 15 factors with significant variance in the type of motivation of the student (has adequate learning resources, sense of belonging to the university, qualification of faculty, alignment with the job market, access to facilities of industrial importance and relatively up-to-date, and suitable program and courses type, interaction with students outside the academic environment, adequate curricular matrix and courses contents, and existence of mechanisms to facilitate their insertion into the labor market and prospects for the future). We found only three factors with variance in the approval rate (entrance exam position, fun, and LMS support) and five factors with variance in the overall grade average (entrance exam position, feel prepared for study, do the possible to stand out in the class, and commitment of activities and deadlines).

With that, we rejected the null hypothesis H_{D0} (No motivation factors related to educational context impact significantly the students' outcome) and we confirmed the alternative hypothesis H_{D1} (One or more motivation factors related to educational context impact significantly the students' outcome).

7.2 COMPARING EMMECS TO STATE OF THE ART

In order to evaluate the results of the EMMECS application for identification of students at risk in introductory programming courses, this section shows a comparison of the results with the studies found in the literature.

We compared our study with studies that demonstrate results of temporal prediction through the course and also with studies using only demographic or psychometric data such as proposed for this work. Table 82 shows a comparison of the prediction of the results per week between this work and similar studies found in the literature, even these studies do not use the motivation or psychometric data as input attribute.

Table 82: Weekly comparison between EMMECS and state of the art

Week	**1ª**	**2ª**	**3º**	**4ª**	**5ª**	**6ª**	**Final**	**Features input type**
(VIHAVAINEN, 2013)		64.0%					78.0%	Interaction, programming behavior
(UMER, SUSNJAK, *et al.*, 2017) (AUC)	82.9%	86.1%	87.2%	87.1%	87.8%	88.0%	NI	Grade, activities perforamnce
(WATSON and LI, 2014)			only variance is explained				75.6%	Interaction. programming behavior
(DETONI, ARAUJO and CECHINEL, 2014)	35.0%	48.0%	52.0%	57.0%	64.0%	68.0%		LMS interaction
(HUNG, WANG, *et al.*, 2017)	84.1%	84.1%	84.1%	84.1%	84.1%	84.1%	89.8%	Grade, LMS interaction
(PARDO, HAN and ELLIS, 2016) (MSE)		15.8	15.5	15.1	14.5	29.2	32.0	LMS interaction, pre-university
(SANTOS, PITANGUI, *et al.*, 2016)	67.7%	68.1%	70.6%	74.2%	75%	74.6%	84.7%	LMS interaction
(KLOFT, STIEHLER, *et al.*, 2014)	72.0%	74.0%	83.0%	83.0%	84.0%	85.0%	85.0%	LMS interaction
(MACHADO, CECHINEL and RAMOS, 2018)	10.0%	50.0%	60.0%	75.0%	85.0%	75.0%	95.0%	LMS interaction (recall) [a]
This work [a]	90.0%	81.4%	84.1%	85.7%	78.5%	82.3%	78.8%	Pre-university and motivation

a. Considering the best dataset for the classifier SMO (SVM)

We realize that the results are similar or superior to existing studies with the best performances, even if they use input attributes related to interaction or performance on assessments. The most interesting is that the experiment using EMMECS had the best performance in the first week and remained stable over the weeks, while most of the studies have worse results in the first few weeks and gradually improve until the end of the period. Of the seven works that detail the prediction performance of students at risk from first to the sixth week, this work had 83.67% average, trailing behind the work of Umer et al. (2017) had 86,52%, and Hung et al. (2017) had 84,1%.

Comparing with similar studies that use only demographic or psychometric data, the experiment described in this work proved to be significantly better. Table 83 shows the works found in the mapping study of the literature described in section 3.2.2, in addition to the work described in the review of Shahiri et al. (2015).

Table 83: Comparison of results to the state of the art

Author(s)	Year	Sample	Best Results
U. Ninrutsirikun (2016)	2016	85	78.33% (accuracy) 0.709 (f-measure)
Ayub (2017)	2017	41	70.73% (accuracy)
Aziz (2014)	2014	399	68.80% (accuracy)
Quille (2018)	2018	692	71.00% (accuracy) 75.00% (recall) 66.00% (specificity)
Gray et al. (2014)	2014	914	73.33% (accuracy)
Ramesh et al. (2013) [a]	2013	500	72.38% (accuracy)
Sembiring et al. (2011) [a]	2011	300	78.33% (accuracy) [b]
This work	2019	173	72.83% (accuracy) 0.803 (recall) 0.771 (f-measure) 0.814 (AUC)

a. Not specific to introductory computing courses

b. Estimated by an average of classes accuracy

Source: Developed by the author

Therefore, the present work has generated a result similar to studies with better performances. For example, Sembiring et al. (2011) used a questionnaire with 50 items for interest, belief, behavior, family support and time engaged. Despite the result, the work described by Sembiring and colleagues performs the collection in the third semester of studies, refers to three different majors in a college of Computer System and Software Engineering, and the dependent variable is the grade CGPA. Still, the average accuracy is demonstrated only by the class of performance. As the work does not mention how many instances exist in each class, we did an estimate by calculating the arithmetic mean. However, this value may be higher or lower depending on this sample variation.

Ninrutsirikun et al. (2016) did a study with 85 students of the programming course, using as attributes: academic record, namely, the national level test scores in: English language, Mathematics, and Science, high school GPA, and the measure of Gardner's seven intelligences (Multiple Intelligences model). For the sample to be small and imbalanced, the authors made the balancing of the dataset using the technique SMOTE. The sample contained only eight instances of the class with worse performance (Poor).

Quille and Bergin (2018) replicated a previous study using the Predict Student Success (PRESS) model. The authors suggested also updates the model, and the best results were obtained including four new attributes: the students' age, students' self-reported expected end of module result, students' mathematical ability and students programming self-efficacy. With this updated model, the study reached 71% of accuracy, 75% of sensibility, and 66% of specificity.

Therefore, we realize that the proposed method was shown to be efficient, achieving similar results in terms of accuracy, but best in terms of sensibility/recall. Although some related studies have not used metrics such as recall or f-measure, we believe that as the goal is to identify students at risk, these should be the main metrics used.

7.3 PUBLICATIONS

During the development of this research, we published partial results as several conferences and journal papers, website, and technical reports (Table 84).

Table 84: Publications related to this thesis

Id	Reference	Results	Computer Science Qualis
1	(SCHOEFFEL, ROSA and WAZLAWICK, 2016)	Indirect: motivation to thesis theme and identification of factors that affect student's motivation.	B3
2	(SCHOEFFEL, WAZLAWICK and RAMOS, 2018c)	Analysis of state-of-art (Section 3.1)	B1
3	(SCHOEFFEL, WAZLAWICK, *et al.*, 2018b)	Validation of part of the method proposed (Section 4.2)	B1
4	(SCHOEFFEL, WAZLAWICK and RAMOS, 2017)	Validation of part of the method proposed (Section 4.2)	B1
5	(SCHOEFFEL, RAMOS, *et al.*, 2018)	Validation of part of the method proposed (Section 4.5)	B1
6	(SCHOEFFEL, WAZLAWICK and RAMOS, 2018b)	Indirect: motivation to thesis theme and identification of factors that affect student's motivation.	B1
7	(SCHOEFFEL, WAZLAWICK, *et al.*, 2018a)	Validation of part of the method proposed (Section 4.5)	B1
8	(SCHOEFFEL, 2017)	First validation of method proposed (Section 4)	B4
9	(SCHOEFFEL, WAZLAWICK and RAMOS, 2019)	Validation of part of the method proposed (Section 4.3)	A2 (in review)
10	pablo.pro.br/emmecs	A website with the EMMECS method description and available tools	-
Other publications indirectly related to the research topic of this thesis			

11	(VAHLDICK, MENDES, *et al.*, 2015)		B3
12	(VAHLDICK, SCHOEFFEL, *et al.*, 2018)		-

Source: Developed by the author

7.4 THREATS TO VALIDITY

Like any kind of research, this study has limitations and it is subject to threats to validity. In order to minimize the impact of potential threats, we applied mitigation strategies.

With respect to internal threats, Campbell and Stanley (1963) identified and discussed eight extraneous variables types that can, if not controlled, compromising the internal validity of an experiment. We identified three of those types that can apply to our research: history, selection, and experimental mortality.

History: “the specific events occurring between the first and second measurement in addition to the experimental” (CAMPBELL and STANLEY, 1963). With this, it is likely that the experiences vary between subjects and have a differential effect on the responses of the subject Studies that take repeated measurements in subjects over time are more likely to be affected by variables of history than those that collect data in shorter periods of time or who do not use repeated measures. In our case, to mitigate the impact of these experiences between measurements, we used the average of all the longitudinal measures in order to correct biases of specific measures.

Selection: individuals in comparison (for example, control group and experimental) should be functionally equivalent at the beginning of a study. If this is the case, the observed differences between groups, as measured by the (s) (s) dependent variable (s) of performance, at the end of the study, are more likely to be caused only by the independent variable, rather than spurious variables. If the comparison groups differ from each other at the beginning of the study, the study results are biased. In our case, despite not having a control group, we compared students who passed or failed. In order to mitigate the impact of the difference in the number of individuals in each group, we used data balancing techniques.

Experimental mortality: occurs when subjects abandon the research. If a comparison group to present a higher level of withdrawal/mortality of subject than other groups, the differences observed between the groups become questionable. Are the observed differences produced by independent variable or by different rates of abandonment? (Mortality is also a threat when the drop-out rates are similar in comparison groups but are high). To avoid the bias of individuals that drop the course during the semester, we used always the last measure or the average of the measurements, keeping your historical data and avoiding unbalancing the data set even more.

With respect to the external validity of the research, Campbell and Stanley (1963) identified four possible factors that can affect the validity of a study. Of these, two are inherent in this research and were mitigated.

An interaction between how subjects were selected and the treatment can occur. If individuals are not selected randomly from a population, its demographic characteristics an individual may promote your performance and results of the study may not apply to the population or to another group that represents more accurately the characteristics of the population. This is a limitation of the survey since the validations were carried out in the context of computing undergraduation programs in southern Brazil. To minimize this threat, we tried to vary the sample, applying the method in four different universities, two public and two private, since this is a scenario that impacts on the characteristics of the students and universities in Brazil.

The performance of subjects in some studies is a reaction to the experimental setup (for example, the situation where the study is conducted). For example, the guys who know are participants of a study may react differently to treatment than a guy who tried the treatment, but was not aware of being observed. To mitigate this threat, we followed procedures approved on the Ethics Committee, ensuring total confidentiality, the participants were aware that their individual responses would not be disclosed and that they could answer honestly, with no type of prejudice.

8 CONCLUSION AND FUTURE WORKS

The main contribution of this work is to develop a method that makes it possible to identify in advance at-risk students in introductory computing courses. The EMMECS method is based on three main instruments, in the form of questionnaires. The three instruments have been validated as for their reliability and validity using statistical methods such as Cronbach's alpha coefficient, omega coefficient, intercorrelation, and factor analysis. All results were satisfactory.

Each of the instruments allowed to find factors related to the outcome and motivation of students. In several case studies, we found evidence of the relationship of the students' outcome to four aspects of initial motivation, fifteen educational factors, the professor's perception and motivation through the course.

With respect to the validation of the method, the EMMECS method was applied with 173 students in nine classes of four different universities in Brazil. The results of the assertiveness of identification of students fail (recall) was, at best, about 80% on average, but reaching more than 90%. An interesting result is that the method had a good and consistent performance since the first few weeks. The application of the method in a real context showed that the EMMECS is a tool that can assist significantly professors in driving their courses and help in interventions with specific students.

Since the method is independent of content, or tools, other researchers and professors can use the EMMECS as a tool to assist in improving introductory computing courses.

For the continuation of the work, we have planned the continuity of use of the EMMECS method, including different contexts (courses and universities). We want to upgrade the tool in order to automate the data collection process and reports delivery.

We also intend to conduct a more detailed analysis and identification of atypical situations, as students with a profile for failure but with good results, or vice versa. This analysis aims to improve the method, trying to identify factors not covered or reducing the number of unnecessary attributes.

Despite the method validation have taken into account the context of introductory computing courses, we believe that the EMMECS can be adapted and used in other contexts, including other courses and programs. For this, further studies are needed to validate the methods in these contexts.

REFERENCES

ABDULLAH, A.; YIH, Y. Implementing Learning Contracts in a Computer Science Course as a Tool to Develop and Sustain Student Motivation to Learn. **Procedia - Social and Behavioral Sciences**, 123, 2014. 256 – 265.

AGGARWAL, C. C. **Data Classification:** Algorithms and Applications. Boca Raton, FL: Taylor & Francis Group, 2015.

ALHAZBI, S. **ARCS-Based Tactics to Improve Students' Motivation in Computer Programming Course**. 10th International Conference on Computer Science & Education (ICCSE 2015). Cambridge, UK: IEEE. 2015. p. 317-321.

ALLEN, et al. Third-year college retention and transfer: Effects of academic performance, motivation, and social connectedness. **Research in Higher Education**, v. 49, p. 647- 664, 2008.

ANOOPKUMAR, M.; RAHMAN, Z. **A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration**. 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE). Ernakulam, India: [s.n.]. 2016.

APPLETON, J. J. et al. Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. **Journal of School Psychology**, 44, n. 5, 2006. 427-445.

ASTIN, A. W. Student Involvement: A Development Theory for Higher Education. **Journal of College Student Development**, v. 40, n. 5, p. 518-524, July 1984.

ATES, A. Self-efficacy Beliefes, Achievement Motivation and Gender as Related to Educational Software Development. **Turkish Online Journal of Distance Education**, 12, n. 3, July 2011. 11-22.

ATKINSON, JOHN W.. Motivational determinants of risk-taking behavior. **Psychological Review**, 64, n. 6, nov. 1957. 359-372.

AYNSLEY, B. Shortage of IT and communications tech talent costly to economy. **The Australian**, 17 Junho 2015. Disponivel em: <http://www.theaustralian.com.au/business/business-spectator/shortage-of-it-and-

communications-tech-talent-costly-to-economy/news-story/833473b28c80ee505d4f2fcf1fa75097>. Acesso em: 13 Outubro 2016.

AYUB, M.; KARNALIM, O. Predicting outcomes in Introductory Programming using J48 classification. **World Transactions on Engineering and Technology Education**, v. 15, n. 2, p. 132-136, 2017.

AZIZ, A. A.; AHMAD, N. H. I. **First Semester Computer Science Students' Academic Performances Analysis by Using Data Mining Classification Algorithms**. Proceeding of the International Conference on Artificial Intelligence and Computer Science(AICS. Bandung, Indonesia: [s.n.]. 2014. p. 100-109.

BAKER, R. S.; INVENTADO, P. S. Educational data mining and learning analytics. **Learning analytics**, New York, p. 61-75, 2014.

BAYER , J. et al. **Predicting drop-out from social behaviour of students**. Proceedings of the 5th International Conference on Educational Data Mining - EDM 2012. Greece: [s.n.]. 2012. p. 103 - 109.

BEECHAM, S. et al. Motivation in Software Engineering: A systematic literature review. **Information and Software Technology**, 17 Oktober 2007. 860–878.

BEGENY, J. C. et al. Teachers' Perceptions of Students' Reading Abilities: An Examination of the Relationship Between Teachers' Judgments and Students' Performance Across a Continuum of Rating Methods. **School Psychology Quarterly**, 23, n. 1, 2008. 43-55.

BENNEDSEN, J.; CASPERSEN, E. Failure Rates in Introductory Programming. **ACM SIGCSE Bulletin**, 39, n. 2, June 2007. 32-36.

BERGIN, S.; REILLY, R. **The influence of motivation and comfort-level on**. Proceedings of the 17th Workshop of the Psychology of Programming Interest Group, PPIG 05. Brighton, UK: University of Sussex. 2005. p. 293 - 304.

BIGGS, J.; TANG, C. **Teaching for Quality Learning at University**. 3. ed. [S.l.]: RefineCatch Limited, 2007.

BOMIA, L. et al. **The Impact of Teaching Strategies on Intrinsic Motivation**. ERIC. [S.l.], p. 28p. 1997.

BORBA, S. F. P. **Ações realizadas em um curso superior de tecnologia para reduzir sua evasão**. XXXV Congresso da Sociedade Brasileira de Computação. Recife: [s.n.]. 2015.

BORNAT, R.; DEHNADI, S.; SIMON. **Mental models,consistency and programming aptitude**. Proc. Australasian Computing Education. [S.l.]: [s.n.]. 2008. p. 53-61.

BOSSE, Y.; GEROSA, M. A. **As Disciplinas de Introdução à Programação na USP:** um Estudo Preliminar. Proceedings of Workshops of CBIE 2015. [S.l.]: SBC. 2015. p. 1-10.

BOUSBIA, N.; BELAMRI, I. Which Contribution Does EDM Provideto Computer-Based Learning Environments? In: PEÑA-AYALA, A. **Educational Data Mining:** application and trends. 1. ed. [S.l.]: Springer International Publishing, v. 524, 2013. Cap. 1, p. 3-28.

BRASSCOM. **Relatório Setorial de TIC**. [S.l.]. 2019.

BROPHY, J. **Conceptualizing Student Motivation**. Institute for Research on Teaching. Michigan, USA, p. 34. 1983.

BROWN, T. A. **Confirmatory Factor Analysis for Applied Research**. 2. ed. New York, NY: The Guilford Press, 2015.

BRUINSMA, M. Motivation, cognitive processing and achievement in higher education. **Learning and Instruction**, 14, n. 6, December 2004. 549–568. Disponivel em: <http://www.sciencedirect.com/science/article/pii/S0959475204000702>. Acesso em: 13 outubro 2016.

BUGHIN, J. et al. **Skill Shift Automation and the Future of the Wordforce**. McKinsey Global Institute. [S.l.], p. 23. 2018.

CAMPBELL, D. T.; FISKE, D. W. CONVERGENT AND DISCRIMINANT VALIDATION BY THE MULTITRAIT-MULTIMETHOD MATRIX. **Psychological Bulletin**, v. 56, n. 2, p. 81-105, March 1959.

CAMPBELL, D. T.; STANLEY, J. C. **Experimental and Quasi-experimental Designs for Research**. Boston: Houghton Mifflin Company, 1963.

CARMINES, E. G.; ZELLER, R. A. **Reliability and Validity Assessment Sage University Papers Series. Quantitative Applications in the Social Sciences**. Serie No. 07-017. ed. Iowa City, IA: Sage Publications, Inc., 1979.

CARTER, L. **Why Students with an Apparent Aptitude for Computer Science Don't Choose to Major in Computer Science**. Proceedings of SIGCSE '06. Houston, Texas, USA: [s.n.]. 2006.

CHAWLA, N. V. et al. SMOTE: Synthetic Minority Over-sampling Technique. **Journal of Artificial Intelligence Research**, v. 16, n. 2002, p. 321–357, 2002.

CODE.ORG. Code.org. **Code.org**, 2019. Disponivel em: <https://code.org/>. Acesso em: 19 June 2019.

COHEN, J. **Statistical Power Analysis for the Behavioral Sciences**. 2. ed. New York: LAWRENCE ERLBAUM ASSOCIATES, 1988.

COMREY, A. L. . &. L. H. B. **A first course in factor analysis**. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc, 1992.

COOPER, S. et al. A Future for Computing Education Research. **Communication of the ACM**, v. 57, n. 11, p. 34-36, 2014.

COOPER, S. et al. **The Importance of Computing Education Research: A white paper prepared for the Computing Community Consortium committee of the Computing Research Association**. Computing Research Association. [S.l.], p. 12. 2016.

CRESWELL, J. W. **Research Design:** qualitative, quantitative, and mixed methods approaches. 2. ed. London, New Delhi: SAGE Publications, v. 1, 2003.

CRISP, G.; TAGGART, A.; NORA, A. Undergraduate Latina/o Students: A Systematic Review of Research Identifying Factors Contributing to Academic Success Outcomes. **Review of Educational Research**, v. 85, n. 2, p. 249-274, June 2015.

CRONBACH, L. J. Coefficient alpha and the internal structure of tests. **PSYCHOMETRIKA**, v. 16, n. 3, p. 297-334, September 1951.

DAMASCENO, I. L.; CARNEIRO, M. G. **Panorama da Evasão no Curso de Sistemas de Informação da Universidade Federal de Uberlândia:** Um Estudo Preliminar. Proceedings of the XXIX Simpósio Brasileiro de Informática na Educação (SBIE 2018). Natal: SBC. 2018. p. 1766-1770.

DAVIS, F. D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. **MIS Quarterly**, v. 3, n. 3, p. 319-340, 1989.

DEBDI, ; PAREDES-VELASCO, ; VELÁZQUEZ-ITURBIDE, J. Á. **Relationship between learning styles, motivation and educational efficiency in students of computer science**. 2014 International Symposium on Computers in Education (SIIE). Logrono: IEEE. 2014. p. 13-16.

DEJAEGER, K. et al. Gaining insight into student satisfaction using comprehensible data mining techniques. **European Journal of Operational Research**, n. 218, 2012. 548–562.

DETONI, D.; ARAUJO, R. M.; CECHINEL, C. **Predição de Reprovação de Alunos de Educação a Distância Utilizando Contagem de Interações**. XXV Simpósio Brasileiro de Informática na Educação (SBIE 2014). [S.l.]: SBC. 2014. p. 896-905.

DEVILLES, R. F. **Scale development:** theory and applications. 2ª. ed. Thousand Oaks, California: Sage Publications, Inc., 2003.

DREW, C. Why Science Majors Change Their Mind. **New York Times**, New York, p. 6, Novembro 2011. Disponivel em: <http://www.nytimes.com/2011/11/06/education/edlife/why-science-majors-change-their-mind-its-just-so-darn-hard.html?_r=4&pagewanted=all>. Acesso em: 13 Outubro 2016.

DUNN, T. J.; BAGULEY, T.; BRUNSDEN, V. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. **British Journal of Psychology**, v. 105, p. 399-412, 2014.

EASTERBROOK, S. et al. Selecting Empirical Methods for Software Engineering Research. In: SHULL, F.; SINGER, J.; SJøBERG, D. I. K. **Guide to Advanced Empirical Software Engineering.** 1. ed. [S.l.]: Springer, v. 1, 2008. Cap. 11, p. 285-311.

ECCLES, J. S. et al. Expectancies, values, and academic behaviors. In: ______ **Achievement and achievement motivation**. San Francisco, CA: W. H. Freeman., v. 1, 1983. p. 75–146.

ELATIA, S.; IPPERCIEL, D.; ZAÏANE, O. R. **Introduction:** educational at computational crossroads. 1. ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2016.

ENTWISTLE, N. **Concepts and conceptual frameworks underpinning the ETL project**. ETL Project. Edinburgh, Scotland. 2003.

EUBANKS, D.; EVERS JR., W.; SMITH, N. Finding Predictors in Higher Education. In: ELATIA, S.; IPPERCIEL, D.; ZAÏNE, O. R. **Data Mining and Learning Analytics:**

applications in educational research. Hoboken, New Jersey: John Wiley & Sons Inc., 2016. Cap. 3, p. 41-53.

EUROPEAN POLITICAL STRATEGY CENTRE (EPSC). **Global Trends to 2030: the future of work and workspaces**. European Strategy and Policy Analysis System. [S.l.], p. 14. 2018.

FERREIRA, C. E. **Evasão:** alguns dados e muitas dúvidas. XXXV Congresso da Sociedade Brasileira de Computação. Recife: [s.n.]. 2015.

FIGAS, P.; HAGEL, G.; BARTEL, A. **The Furtherance of Motivation in the Context of Teaching Software Engineering**. IEEE Global Engineering Education Conference (EDUCON). Berlin, Germany: IEEE. 2013. p. 1299-1304.

FINCHER, S.; PETRE, M. **Computer Science Education Research**. London, UK: Taylor & Francis Group, 2004. 231 p.

FRANÇA, A. C. C. et al. **Motivation in Software Engineering:** A Systematic Review Update. Proceedings of EASE 2011. [S.l.]: [s.n.]. 2011. p. 10.

GAREIS, K. et al. **E-skills for Jobs in Europe: measuring progress and moving ahead**. European Comission. [S.l.]. 2014.

GIESBERS, B. et al. Investigating the relations between motivation, tool use, participation, and performance in an e-learning course using web-videoconferencing. **Computers in Human Behavior**, n. 29, 2013. 285–292.

GOMES, A. D. J. **Dificuldades de Aprendizagem de Programação de Computadores: contributos para sua compreensão e resolução**. Universidade de Coimbra. Coimbra, p. 468. 2000.

GOMES, A.; MENDES, A. J. N. **Learning to program-difficulties and solutions**. International Conference on Engineering Education. Coimbra, Portugal: [s.n.]. 2007. p. 1-5.

GRAY, G.; MCGUINNESS, C.; OWENDE, P. **An application of classification models to predict learner progression in tertiary education**. IEEE International Advance Computing Conference (IACC). [S.l.]: IEEE. 2014. p. 549-554.

GUAY, F.; VALLERAND, R. J.; BLANCHARD, C. On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). **Motivation and Emotion**, 24, n. 3, 2000. 175-213.

GUO, W. W. Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction. **Expert Systems with Applications**, n. 37, 2010. 3358–3365.

GURULER, H.; ISTANBULLU, A.; KARAHASAN, M. A new student performance analyzing system using knowledge discovery in higher educational databases. **Computers & Education**, 2010. 247–254.

HAIR , J. et al. **Multivariate data analysis**. 7th. ed. New Jersey: Prentice-Hall Inc;, 2010.

HÄMÄLÄINEN, W.; VINNI, M. Classifiers for Educational Data Mining. In: ROMERO, C., et al. **Handbook of Educational Data Mining**. Boca Raton, FL: CRC Press, 2011. Cap. 5, p. 57-72.

HARDRÉ, P. L.; DAVIS, K. A.; SULLIVAN, D. W. Measuring teacher perceptions of the “how” and “why” of student motivation. **Educational Research and Evaluation**, 14, n. 2, 2008. 155–179.

HELLAS, A. et al. **Predicting Academic Performance:** A Systematic Literature Review. Proceedings of ITiCSE ’18. Larnaca, Cyprus: ACM. 2018. p. 175-199.

HERZBERG, F. One More Time: How do you motivate employees? **Harvard Business Review**, p. 53-62, January-February 1968.

HIDAYAH, I.; PERMANASARI, A. E.; RATWASTUTI, N. **Student classification for academic performance prediction using neuro fuzzy in a conventional**. International Conference on Information Technology and Electrical Engineering (ICITEE). [S.l.]: IEEE. 2013. p. 221–225.

HOWLES, T. **Work In Progress - Intrinsic Motivation as a Predictor of Success in Computer Science**. 35th ASEE/IEEE Frontiers in Education Conference. Indianapolis, USA: [s.n.]. 2005.

HULSE, V.; KHOSHGOFTAAR, T. M.; NAPOLITANO, A. **Experimental Perspectives on Learning from Imbalanced Data**. Proceedings of the 24 th International Conference on Machine Learning. Corvallis, OR: [s.n.]. 2007. p. 35.

HUNG, J.-L. et al. Identifying At-Risk Students for Early Interventions—A Time-Series Clustering Approach. **IEEE Transactions on Emerging Topics in Computing**, 5, n. 1, 2017. 45-55.

HÜSING, T. et al. **e-Leadership: e-Skills for Competitiveness and Innovation Vision, Roadmap and Foresight Scenarios**. European Comission. [S.l.]. 2013.

HÜSING, T.; DASHJA, E. **Innovation Leadership Skills for the High-Tech Economy - Demand, Supply and Forecasting**. Brussels. 2017.

IHANTOLA, P. et al. **Educational Data Mining and Learning Analytics in Programming:** Literature Review and Case Studies. Proceedings of ITiCSE WGR'16. Vilnius, Lithuania: [s.n.]. 2016.

JENKINS, T. **The Motivation of Students of Programming**. The University of Kent. Canterbury, p. 253. 2001.

JENKINS, T. **On the difficulty of learning to program**. Proceedings of the 3rd Annual Conference of the LTSN Centre for Information and Computer Sciences. [S.l.]: [s.n.]. 2002. p. 53-58.

KANAPARAN, G.; CULLEN, R.; MASON, D. **Effect of Self-efficacy and Emotional Engagement on Introductory Programming Students**. Australasian Conference on Information Systems. Hobart, Australia : [s.n.]. 2017. p. 1-11.

KARAGOD, V. How to Handle Imbalanced Data: An Overview. **DataScience.com**, 2018. Disponivel em: <https://www.datascience.com/blog/imbalanced-data>. Acesso em: 10 jun. 2019.

KASUNIC, M. **Designing an Effective Survey**. Software Engineering Institute - Carnegie Mellon University. Pittsburgh, PA, p. 143. 2005. (CMU/SEI-2005-HB-004).

KELLER, J. M. Development and use of the ARCS model of instructional design. **Journal of Instructional Development**, v. 10, n. 3, p. 2 - 10, 1987.

KITCHENHAM, B. A. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. EBSE Technical Report. [S.l.], p. 65. 2007.

KITCHENHAM, B. A.; BUDGEN, D.; BRERETON, O. P. **The value of mapping studies-A participant-observer case study**. [S.l.]: [s.n.]. April 2010. p. 25-33.

KLEMENT, et al. **Dealing with Severely Imbalanced Data**. Proc. of the PAKDD Conference. [S.l.]: [s.n.]. 2009. p. 14.

KLOFT, M. et al. **Predicting MOOC Dropout over Weeks Using Machine Learning Methods**. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics. 2014. p. 60-65.

KORI, et al. **First-year dropout in ICT studieS**. 2015 IEEE Global Engineering Education Conference (EDUCON). Tallinn, Estonia: IEEE. 2015. p. 437-445.

KORI, et al. The Role of Programming Experience in ICT Students' Learning Motivation and Academic Achievement. **International Journal of Information and Education Technology**, 6, n. 5, Maio 2016.

KOSOVICH, J. J. et al. A Practical Measure of Student Motivation: Establishing Validity Evidence for the Expectancy-Value-Cost Scale in Middle School. **Journal of Early Adolescence**, v. 35, n. 5-6, p. 790-816, 2014. ISSN 10.1177/0272431614556890.

KUEHN, M. et al. **An expert system for the prediction of student performance in an initial computer science course**. 2017 IEEE International Conference on Electro Information Technology (EIT). Lincoln, NE, USA: IEEE. 2017.

LAK. Learning Analytics. **1st International Conference on Learning Analytics and Knowledge**, Banff, Alberta, 2011. Disponivel em: <https://tekri.athabascau.ca/analytics/>.

LARUSSON, J. A.; WHITE, B. **Learning Analytics:** From Research to Practice. New York: Springer Science, 2014.

LAW, K. M. Y. et al. comparative study of learning motivation among engineering students in South East Asia and beyond. **International Journal of Engineering Education**, v. 25, n. 1, p. 144-151, 2009.

LAW, K. M. Y.; LEE, V. C. S.; YU, Y. T. Learning motivation in e-learning facilitated computer programming courses. **Computers & Education**, 55, 2010. 218–228.

LISHINSKI, A. et al. **Learning to Program:** Gender Differences and Interactive Effects of Students' Motivation, Goals, and Self-Efficacy on Performance. Proceedings of ICER '16. Melbourne, VIC, Australia: ACM. 2016. p. 211-220.

LISTON, M. et al. **A Study of Progression in Irish Higher Education: 2014/15 to 2015/16**. [S.l.]. 2018.

MACHADO, M.; CECHINEL, C.; RAMOS, V. F. C. **Comparação de diferentes configurações de bases de dados para a identificação precoce do risco de reprovação:** o caso de uma disciplina semipresencial de Algoritmos e Programação. Proceedings of XXIX Simpósio Brasileiro de Informática na Educação (SBIE 2018). Fortaleza, Brazil: SBC. 2018. p. 1503-1512.

MAMONE, S. Empirical Study of Motivation in a entry Level Programming Course. **ACM SIGPLAN Notices**, v. 27, n. 3, p. 7, March 1992.

MARTIN, A. J. Examining a multidimensional model of student motivation and engagement using a construct validation approach. **British Journal of Educational Psychology**, 77, 2007. 413–440.

MASLOW, A. H. A Theory of Human Motivation, p. 370-396, 1954.

MCCARTNEY, et al. Why Computing Students Learn on Their Own: Motivation for Self-Directed Learning of Computing. **ACM Transactions on Computing Education (TOCE)**, v. 16, n. 1, p. 18, February 2016.

MOISSA, B.; GASPARINI, I.; KEMCZINSKI, A. **Educational Data Mining versus Learning Analytics:** estamos reinventando a roda? Um mapeamento sistemático. Anais do XXVI Simpósio Brasileiro de Informática na Educação (SBIE 2015). [S.l.]: [s.n.]. 2015.

MOONEY, O. et al. **A Study of Progression in Irish Higher Education**. HEA – High Education Authority. [S.l.]. 2010.

NATIONAL SCIENCE FOUNDATION. CS for All. **CS for All**, 2019. Disponivel em: <https://www.nsf.gov/news/special_reports/csed/csforall.jsp>. Acesso em: 19 june 2019.

NAVARRO, E. O.; VAN DER HOEK, A. On the Role of Learning Theories in Furthering Software Engineering Education. **Software Engineering: Effective Teaching and Learning Approaches and Practices**, 2009. 38-59.

NGAN, S.-C.; LAW, K. M. Y. Exploratory Network Analysis of Learning Motivation Factors in e-Learning Facilitated Computer Programming Courses. **The Asia-Pacific Education Researcher**, December 2015. 705–717.

NIELSEN, J. **Usability Inspection Methods**. Conference Companion on Human Factors in Computing Systems. Boston, USA: ACM. 1994. p. 413-414.

NIITSOO, M. et al.**. Predictors of Informatics Students Progress and Graduation in University Studies**. Proceedings of INTED2014 Conference. Valencia, Spain: [s.n.]. 2014.

NIKULA, U.; GOTEL, O.; KASURINEN, J. A motivation guided holistic rehabilitation of the first programming course. **ACM Trans. Comput. Educ**, 11, n. 4, November 2011. 38.

NINRUTSIRIKUN, U. et al. **Effect of the Multiple Intelligences in Multiclass Predictive Model of Computer Programming Course Achievement**. Proceedings of 2016 IEEE Region 10 Conference (TENCON). [S.l.]: IEEE. 2016. p. 297-300.

NUNES, D. **Educação Superior em Computação: estatísticas 2017**. SBC - Sociedade Brasileira de Computação. [S.l.], p. 65. 2018.

PALMEIRA, L. B.; SANTOS, M. P. **Evasão no Bacharelado em Ciência da Computação da Universidade de Brasília: análise e mineração de dados**. Universidade de Brasília. Brasília, p. 287. 2014.

PAPAMITSIOU, Z.; ECONOMIDES, A. A. Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. **Educational Technology & Society**, v. 17, n. 4, p. 49-64, 2014.

PARDO, A. Designing Learning Analytics Experiences. In: ______ **Learning Analytics:** From Research to Practice. New York: Springer Science, 2014. p. 15-38.

PARDO, A.; HAN, F.; ELLIS, R. A. Combining University Student Self-regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance. **IEEE Transactions on Learning Technologies**, 2016.

PASINI, M. et al. **The Challenge of Learning to Program:** motivation and achievement emotions in an eXtreme Apprenticeship experience. PPIG 2016 - 27th Annual Conference. [S.l.]: [s.n.]. 2016.

PAYNE, A. L. L. **Why do students choose computing?: influences, perceptions and engagement**. Coventry University. [S.l.], p. 448. 2013.

PEÑA-AYALA, A. **Educational Data Mining:** Applications and Trends. 1. ed. [S.l.]: Springer, 2014.

PETERS, A.-K.; PEARS, A. **Engagement in Computer Science and IT – What! A Matter of Identity?** Learning and Teaching in Computing and Engineering (LaTiCE). Washington, DC, USA: IEEE Computer Society. 2013. p. 114-121.

PETERSEN, K. et al. **Systematic mapping studies in Software Engineering**. [S.l.]: [s.n.]. 2008.

PINTRICH, P. R. **A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)**. National Center for Research to Improve Postsecondary. [S.l.]. 1991. (NCRIPTAL-91-B-004).

PRATI, R.; BATISTA, G.; MONARD, M. C. **Data mining with imbalanced class distributions:** concepts and methods. Indian International Conference on Artificial Intelligence (IICAI). [S.l.]: [s.n.]. 2009. p. 359-376.

QAA. **UK Quality Code for Higher Education, Part B: Ensuring and Enhancing Academic Quality, Chapter B5 Student Engagement**. Quality Assurance Agency for Higher Education. Gloucester, UK, p. 21. 2012.

QUILLE, K.; BERGIN, S. **Programming:** Predicting Student Success Early in CS1. A Re-validation and Replication Study. Proceedings of ITiCSE'18. Larnaca, Cyprus: ACM. 2018. p. 15-20.

RAMAL, A. Ensino superior: entre o abandono e o baixo desempenho, poucos se destacam. **G1 - globo.com**, 06 jun. 2017. Disponivel em: <http://g1.globo.com/educacao/blog/andrea-ramal/post/ensino-superior-entre-o-abandono-e-o-baixo-desempenho-poucos-se-destacam.html>.

RAMOS, V. et al. **A Comparac¸ao da Realidade Mundial do Ensino de Programac¸ao para Iniciantes com a Realidade Nacional:** Revisao sistemática da literatura em eventos brasileiros. Procedings of XXVI Simpósio Brasileiro de Informática na Educação (SBIE 2015). Maceió, AL, BRA: SBC. 2015. p. 318-327.

RAO, J.; WANG, F. **Research on Motivation of Computer Science Students in Financial College**. 3rd International Conference on Science and Social Research (ICSSR 2014). Tianjin, China: Atlantis Press. 2014. p. 555-559.

ROBEY, M. et al. **Student Self-Motivation:** Lessons Learned from Teaching First Year Computing. 36th ASEE/IEEE Frontiers in Education Conference. San Diego, CA, USA: IEEE. 2006. p. M4E-6 - M4E-11.

ROBINS, A.; ROUNTREE, J.; ROUNTREE, N. Learning and Teaching Programming: A Review and Discussion. **Computer Science Education**, v. 13, n. 2, p. 137–172, 2003.

RODRIGUES, F. S. **Estudo Sobre A Evasão No Curso De Ciência Da Computação Da UFRGS**. Universidade Federal do Rio Grande do Sul. Porto Alegre. 2013.

ROMERO, C. et al. **Handbook of Educational Data Mining**. 1. ed. Boca Raton, FL: CRC Press, 2011.

ROUNTREE, N.; ROUNTREE, J.; ROBINS, A. Predictors of Success and Failure in a CS1 Course. **ACM Bulletin**, 2002. 121-124.

RYAN, R. M.; DECI, E. L. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. **Contemporary Educational Psychology**, 2000. 54-67.

SACHIN, B. R.; VIJAY, S. M. **A Survey and Future Vision of Data mining in Educational Field**. Second International Conference on Advanced Computing & Communication Technologies. [S.l.]: IEEE. 2012. p. 96-100.

SANTOS, R. M. M. D. et al. **Uso de Séries Temporais e Seleção de Atributos em Mineração de Dados Educacionais para Previsão de Desempenho Acadêmico**. Proceedings of XXVII Simpósio Brasileiro de Informática na Educação (SBIE 2016). [S.l.]: SBC. 2016. p. 1146-1155.

SAPNAS, K.; ZELLER, R. A. Minimizing sample size when using exploratory factor analysis for measurement. **J Nurs Meas**, 10, n. 2, 2002. 135-154.

SAYERS, H. M.; NICELL, M. A.; HINDS, A. **Transiction, Engagement and Retention of First Year Computing Students**. 11th Annual Conference of the Subject Centre for Information and Computer Sciences. Durham University: [s.n.]. 2010. p. 90-95.

SCHELLER, ; SCARAMUZZO,. Em um país com desemprego de 13%, sobram vagas na área de tecnologia. **Uol Notícias**, 05 may 2019. Disponivel em: <https://economia.uol.com.br/noticias/estadao-conteudo/2019/05/05/em-um-pais-com-desemprego-de-13-sobram-vagas-na-area-de-tecnologia.htm>.

SCHOEFFEL, P. **Modelo para Predição de Evasão de Estudantes de Computação Baseado na Motivação**. Proceedings of XX Ibero-American Conference on Software Enginerring (CIBSE 2017). Buenos Aires, ARG: [s.n.]. 2017.

SCHOEFFEL, P. et al. **Mapping Factors that Impact on Motivation of Undergraduate Students in Software Engineering:** A case study. Proceedings of XXIX Simpósio Brasileiro de Informática na Educação (SBIE 2018). Fortaleza, BRA: SBC. 2018. p. 963-972.

SCHOEFFEL, P. et al. **Evaluation of an Instrument to Identify Factors that Impact the Motivation and Engagement of Undergraduate Students in Computing**. Proceedings of 2018 IEEE Frontiers in Education Conference (FIE). San Jose, CA, USA: IEEE. 2018a. p. 1-9.

SCHOEFFEL, P. et al. **Identification of Pre-University Factors that Affect the Initial Motivation of Students in Computing Programs:** A multi-institutional case study. 2018 IEEE Frontiers in Education Conference (FIE). San Jose, CA, USA: IEEE. 2018b. p. 1-8.

SCHOEFFEL, P.; ROSA, D.; WAZLAWICK, R. S. Um Experimento do Uso de Coding Dojo na Aprendizagem de Programação Orientada a Objetos. **iSys: Revista Brasileira de Sistemas de Informação**, 9, n. 2, 2016. 71-105.

SCHOEFFEL, P.; WAZLAWICK, R. S.; RAMOS, V. C. **Impact of Pre-University Factors on the Motivation and Performance of Undergraduate Students in Software Engineering**. Proceedings of The 30th IEEE Conference on Software Engineering Education and Training. Savanah, USA: IEEE. 2017. p. 266-275.

SCHOEFFEL, P.; WAZLAWICK, R. S.; RAMOS, V. F. C. **Motivation and Engagement Factors of Undergraduate Students in Computing:** A systematic mapping study. Proceedings of 2018 IEEE Frontiers in Education Conference (FIE). San Jose, USA: IEEE. 2018a. p. 1-5.

SCHOEFFEL, P.; WAZLAWICK, R. S.; RAMOS, V. F. C. **Using Multiple Active Teaching-Learning Approaches in Software Project Management:** A longitudinal analysis of students' motivation and learning. Proceedings of 2018 IEEE Frontiers in Education Conference (FIE). San Jose, CA, USA: IEEE. 2018b. p. 1-9.

SCHOEFFEL, P.; WAZLAWICK, R. S.; RAMOS, V. F. C. **Motivation and Engagement Factors of Undergraduate Students in Computing:** A systematic mapping study.

Proceeding of IEEE Frontiers in Education Conference (FIE). San Jose, USA: [s.n.]. 2018c. p. (in press).

SCHOEFFEL, P.; WAZLAWICK, R. S.; RAMOS, V. F. C. Validation of the Expectancy-Value-Cost Light Scale to Measure Motivation of Students in Computing Courses. **Transactions on Education**, 2019.

SEMBIRING, S. et al. **Prediction os Student Academic Performance by an Application of Data Mining Techniques**. International Conference on Management and Artificial Intelligence. Bali, Indonesia : IACSIT Press. 2011. p. 110-114.

SERRANO-CÁMARA, L. M. et al. An evaluation of students' motivation in computer-supported collaborative learning of programming concepts. **Computers in Human Behavior**, 31, 2013. 499–508.

SHAHIRI, A. M.; HUSAIN, W.; RASHID, N. A. A Review on Predicting Student's Performance using Data Mining Techniques. **Procedia Computer Science**, v. 72, p. 414 – 422, 2015. ISSN 1877-0509.

SHAHIRI, A. M.; HUSAIN, W.; RASHID, N. A. A Review on Predicting Student's Performance using Data Mining Techniques. **Procedia Computer Science**, 72, n. 1, 2015. 414-422.

SHUM, S. B. **Learning Analytics**. UNESCO. Moscow, p. 12. 2012.

SIEMENS, G.; BAKER, R. S. J. Learning analytics and educational data mining: towards communication and collaboration. **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge**, Vancouver, British Columbia, Canada, April 2012. 252-254.

SINCLAIR, J. et al. **Measures of Student Engagement in Computer Science**. Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education (ITICSE 2015). New York, NY, USA: ACM. 2015. p. 242-247.

SOBRAL, D. T. Motivação do Aprendiz de Medicina: Uso da Escala de Motivação Acadêmica. **Psicologia: Teoria e Pesquisa**, 19, n. 1, 2003. 25-31.

SOENENS, B.; VANSTEENKISTE, M. Antecedents and outcomes of self-determination in 3 life domains: The role of parents' and teachers' autonomy support. **Journal of Youth and Adolescence**, 34, n. 6, 2005. 589–604.

SOERJANINGSIH, W. **Student outcomes, learning environment, logical thinking and motivation among computing students in an Indonesian university**. Curtin University of Technology, Science and Mathematics Education. [S.l.]. 2001.

SOFTEX. **Mercado de Trabalho e Formação de Mão de Obra em TI**. Campinas. 2013.

STEPHENSON, C. . D. M. A. . A. C. . B. L. . B. V. . C. T. . F. C. . L. C. . C. M. E. . L. L. . R. D. . S. M. . V. E. . W. H. . A. Z. S. **Retention in Computer Science Undergraduate Programs in the U.S.: Data Challenges and Promising Interventions**. New York. 2018.

TAVAKOL, M.; DENNICK, R. Making sense of Cronbach’s alpha. **International Journal of Medical Education**, 2011, n. 2, 2011. 53-55.

THE CHARTERED INSTITUTE FOR IT. Computing At School. **Computing At School**, 2019. Disponivel em: <https://www.computingatschool.org.uk/about>. Acesso em: 19 June 2019.

TINTO, V. Dropout from Higher Education: A Theoretical Synthesis of Recent Research. **Review of Educational Research**, v. 45, p. 89-125, 1975.

TROCHIM , W. M.; DONNELLY, J. P. **Research methods knowledge base**. 3ª. ed. Mason, OH: Atomic Dog Publishing, 2008. 369 p.

TUAN, H.-L.; CHIN, C.-C.; SHIEH, S.-H. The development of a questionnaire to measure students’ motivation towards science learning. **International Journal of Science Education**, 27, n. 6, 2005. 639-654.

UMER, R. et al. On predicting academic performance with process mining in learning analytics. **Journal of Research in Innovative Teaching & Learning**, 10, n. 2, 2017. 160-176.

UPADYAYA, K.; ECCLES, J. Do teachers’ perceptions of children’s math and reading related ability and effort predict children’s self-concept of ability in math and reading? **Educational Psychology**, 35, n. 1, 19 May 2015. 110-127.

URHAHNE, D. et al. Academic self-concept, learning motivation, and test anxiety of the underestimated student. **British Journal of Educational Psychology**, n. 81, 2011. 161-177.

VAHLDICK, A. et al. **Testando a Diversão em um Jogo Sério para o Aprendizado Introdutório de Programação**. Proceedings of 23o Workshop de Educação em Computação (WEI). Recife, BRA: SBC. 2015. p. 1-10.

VAHLDICK, A. et al. **Maior Frequência na Aplicação de Instrumentos de Avaliação em uma Disciplina Introdutória de Programação:** Impactos no Desempenho e Motivação. Proceedings of Workshops do VII Congresso Brasileiro de Informática na Educação (WCBIE 2018). Natal, BRa: SBC. 2018. p. 739-748.

VALLERAND, R. The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education.. **Educational and psychological measurement** , v. 52 , n. 4, p. 1003-1017, 1992.

VALLERAND, R. J. . E. A. The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. **Educational and psychological measurement**, 52, n. 4, 1992. 1003-1017.

VALLERAND, R. J. et al. Construction et validation de l'echelle de motivation en education (EME). **Canadian Journal of Behavioral Sciences**, v. 21, p. 323-349, 1989.

VERONESE, T. B.; LEMOS, R. M. **SAPERE - Sistema de Apoio ao Ensino para Reducão da Evasão**. XXXV Congresso da Sociedade Brasileira de Computação. Recife: [s.n.]. 2015.

VIHAVAINEN, A. **Predicting Students' Performance in an Introductory Programming Course using Data from Students' own Programming Process**. IEEE 13th International Conference on Advanced Learning Technologies. [S.l.]: IEEE. 2013. p. 498-499.

WATSON, C.; LI, F. W. B. **Failure rates in introductory programming revisited**. Proceedings of the 2014 conference on Innovation technology in computer science education (ITiCSE '14). New York: Association for Computing Machinery (ACM). 2014. p. 39-44.

WHITELEY, ; IAROSSI, ; BRICKNALL, R. **2015 Student Experience Survey Methodology Report**. The Social Research Centre. Melbourne, Australia. 2015.

WIGFIELD,. Expectancy-Value Theory of Achievement Motivation: A Developmental Perspective. **Educational Psychology Review**, 6, n. 1, march 1994. 49-78.

WILLIAMS, B.; ONSMAN, A.; BROWN, T. Exploratory factor analysis: A five-step guide for novices. **Journal of Emergency Primary Health Care (JEPHC)**, v. 8, n. 3, p. 1-13, 2010.

WINSLOW, L. E. Programming pedagogy - a psychological overview. **ACM SIGCSE**, v. 28, n. 3, p. 17-22, 1996.

WOHLIN, C. et al. **Experimentation in Software Engineering**. 1. ed. Berlim: Springer, 2012.

ZAINAL, F. A. et al. Students' perception and motivation towards programming. **Procedia - Social and Behavioral Sciences**, v. 59, p. 277 – 286, 2012.

ZHOU, J.; URHAHNE, D. Teacher judgment, student motivation, and the mediating effect of attributions. **Eur J Psychol Educ**, n. 28, 2013. 275–295.

APPENDIX A – List of Selected Studies in the Systematic Mapping Study about Student Motivation in Computing

Author(s)	Title	Year
Jenkins, T.	The motivation of students of programming	2001
Law, Kris M. Y. ; Lee, Victor C. S. ; Yu, Y. T.	Learning Motivation in E-Learning Facilitated Computer Programming Courses	2010
Jenkins, T. ; Davy, J.	Diversity and motivation in introductory programming	2002
Nikula, U. ; Gotel, O. ; Kasurinen, J.	A Motivation Guided Holistic Rehabilitation of the First Programming Course	2011
Mamone, S.	Empirical study of motivation in an entry-level programming course	1992
Bosnić, I; Čavrak, I; Orlić, M.; Žagar, M.	Student motivation in distributed software development projects	2011
Takemura, Y; Nagumo, H. ; Huang, K. ; Matsumoto, K.	Analysis of the relation between the teaching materials and motivation in programming education	2007
Peters, A. K.; Pears, A.	Engagement in Computer Science and IT--What! A Matter of Identity?	2013
Figas, P. Hagel, G. Bartel, A.	The furtherance of motivation in the context of teaching software engineering	2013
Byrne, P.	Intrinsic motivation in the computer science classroom	1999
Soerjaningsih, W.	Student Outcomes, Learning Environment, Logical Thinking and Motivation AmongComputing Students in an Indonesian University	2001
Robey, M.;Von Konsky, B. R.; Ivins, J.; Gribble, S. J. ; Loh, A.; Cooper, D.	Student self-motivation: lessons learned from teaching first year computing	2006
Magana, A. J.; Mathur, J. I.	Motivation, Awareness, and Perceptions of Computational Science	2012
Zainal, N. F. A.; Shahrani, S; Yatim, N. F. M.; Rahman, R. A.; Rahmat, M.; Latih. R.	Students' perception and motivation towards programming	2012
J Sinclair, M Butler, M Morgan, S Kalvala	Measures of student engagement in computer science	2015
Kori Kuelli ; Pedaste Margus ; Leijen Aeli ; Tonisson Eno	The Role of Programming Experience in ICT Students' learning motivation and Academic Achievement	2016
G Kanaparan, R Cullen, D Mason	Self-Efficacy and Engagement as Predictors of Student Programming Performance	2013
L Payne	Why do students choose computing?: influences, perceptions, and engagement	2013
H Tsukamoto, Y Takemura, H Nagumo, Akito Monden ; Ken-ichi Matsumoto	Prediction of the change of learners' motivation in programming education for non-computing majors	2014

O Debdi, M Paredes-Velasco	Relationship between learning styles, motivation and educational efficiency in students of computer science	2014
S Alhazbi	ARCS-based tactics to improve students' motivation in a computer programming course	2015
DF Shell, LK Soh, AE Flanigan, Markeya S. Peteranetz	Students' Initial Course Motivation and Their Achievement and Retention in College CS1Courses	2016
HM Sayers, MA Nicell, A Hinds	TRANSITION, ENGAGEMENT, AND RETENTION OF FIRST YEAR COMPUTING STUDENTS	2010
Alev Ates	SELF-EFFICACY BELIEFS, ACHIEVEMENT MOTIVATION, AND GENDER AS RELATED TO EDUCATIONAL SOFTWARE DEVELOPMENT	2011
A Abdullah, TY Yih	Implementing Learning Contracts in a Computer Science Course as a Tool to Develop and Sustain Student Motivation to Learn	2014
J Rao, F Wang	Research on Motivation of Computer Science Students in Financial College	2014
N Eltegani, L Butgereit	Attributes of students engagement in fundamental programming learning	2015
SC Ngan, KMY Law	Exploratory Network Analysis of Learning Motivation Factors in e-Learning FacilitatedComputer Programming Courses	2015
Mccartney R. ; Boustedt J. ; Eckerdal A. ; Sanders K. ; Thomas L. ; Zander C.	Why computing students learn on their own: Motivation for self-directed learning of computing	2016

APPENDIX B – List of Selected Studies in the Systematic Mapping Study about Prediction of Students Outcome in Introductory Computing Courses

Document Title	Authors	Year	Sample	Results
A Case Study of Applying the Classification Task for Students' Performance Prediction	M. Silva Guerra	2018	99	75,80%
An expert system for the prediction of student performance in an initial computer science course	M. Kuehn	2017	NI	48,00%
Applying advanced linear models in the task of predicting student success	M. Glavas	2018	77	
Classification and prediction based data mining algorithms to predict students' introductory programming performance	M. Sivasakthi	2017	300	93,00%
Correlation of Grade Prediction Performance with Characteristics of Lesson Subject	S. E. Sorour	2015	123	70,20%
DiCS-Index: Predicting Student Performance in Computer Science by Analyzing Learning Behaviors	D. Capovilla	2016	274	
Effect of the Multiple Intelligences in multiclass predictive model of computer programming course achievement	U. Ninrutsirikun	2016	85	78,33%
Efficiency of data mining models to predict academic performance and a cooperative learning model	P. Amornsinlap hachai	2016	474	75,00%
Implicit Theories and Self-Efficacy in an Introductory Programming Course	F. B. Tek	2018	193	
Learning styles of Computer Science I students	A. N. Kumar	2017	184	
Personalizing Computer Science Education by Leveraging Multimodal Learning Analytics	D. Azcona	2018	325	75,39%
Predicting Performance in an Introductory Programming Course by Logging and Analyzing Student Programming Behavior	C. Watson	2013	45	75,56%
Predicting students' final exam scores from their course activities	M. M. Ashenafi	2015	206	
Predicting Students' Performance in an Introductory Programming Course Using Data from Students' Own Programming Process	A. Vihavainen	2013	152	
Tracking Student Performance in Introductory Programming by Means of Machine Learning	I. Khan	2019	50	88,00%
ASIf_Analyzing undergraduate students' performance using	Asif	2017	210	68,27%
AYUB_Predicting outcomes in Introductory Programming using J48 classification	AYUB	2017	41	70,73%
AZIZ_First Semester Computer Science Students' Academic Performances	AZIZ	2014	399	61,10%
BRITO_Predição de desempenho de alunos do primeiro período	BRITO	2014	300	75,00%
HADEN_Student Affect in CS1 Insights from an Easy Data Collection	HADEN	2017	72	
Iqbal_Machine Learning Based Student Grade	IQBAL	2017	225	67,00%

SANTANA_Um estudo comparativo das técnicas de predição na identificação de insucesso acadêmico dos estudantes durante cursos de programação introdutória	SANTANA	2015	161	83,00%
UMER_On predicting academic	UMEr	2018	167	89,00%
WAtSON_Predicting Performance in an Introductory Programming Course by Logging and Analyzing Student Programming Behavior	WATSON	2013	45	75,56%
COstA_Evaluating the effectiveness of educational data mining techniques	COSTA	2017	161	79,00%
BRITO_Predição de desempenho de alunos do primeiro período	BRITO	2015	406	86,90%
DETONI_Predição de Reprovação de Alunos de Educação	DETONI	2014	329	68,00%
HUNG_Identifying At-Risk Students for Early	HUNG	2015		78,57%
PARDO_Generating Actionable Predictive Models	PARDO	2016	272	
Santos2016_Uso de Séries Temporais e Seleção de Atributos em Mineração de Dados Educacionais para Previsão de Desempenho Acadêmico	Pitangui	2016	248	
ZIMMERMANN_A Model-Based Approach to Predicting	ZIMMERMANN	2014	148	
KLOFT_Predicting MOOC Dropout over Weeks Using Machine Learning Methods	kloft	2014	11606	85,00%
LAMEIRA_Aplicação de Rede Bayesiana na identificação de fatores no curso de SI UFFRJ	LAMEIRA	2014	105	93,30%
YUKCSELTURK_Predicting Dropout Student An Application of Data Mining Methods in an Online Education Program	Yukselturk	2016	189	87,00%
Predicting Success in University First Year Computing Science Courses: The Role of Student Participation in Reflective Learning Activities and in I-clicker Activities	COOKIErMAN	2015	343	
Can Interaction Patterns with Supplemental Study Tools Predict Outcomes in CS1?	ESTEY	2016	652	81,00%
Transfer-Learning Methods in Programming Course Outcome Prediction	LAGUS	2018	358	89,00%
A Robust Machine Learning Technique to Predict Low-performing Students	LIAO	2019	>1000	79,00%
Towards automatic prediction of student performance in STEM undergraduate degree programs	MANHAES	2015	402	85,57%
Lightweight, Early Identification of At-Risk CS1 Students	LIAO	2016	313	70,00%
Programming: predicting student success early in CS1. a re-validation and replication stud	QUILLE	2018	692	71,00%
Student performance prediction model for early-identification of at-risk students in traditional classroom settings	CHANTEKHA	2018	More than 12.000	62,50%
Gender, confidence, and mark prediction in CS examinations	HARRINGTOn	2018	542	73,20%
Predicting Academic Success Based on Learning Material Usage	LEPPANNEN	2018	271	

Comparacão de diferentes configuracões de bases de dados para a identificação precoce do risco de reprovação: o caso de uma disciplina semipresencial de Algoritmos e Programação	Machodo et. Al.	2018	121	+- 75%

APPENDIX C – Questionnaire of Factors that impact the Student Motivation (in Portuguese)

1. Qual seu nível geral de satisfação com o curso? (1 a 5)

Muito Insatisfeito	1	2	3	4	5	Muito Satisfeito

2. Quanto cada razão abaixo lhe motiva a continuar o curso?

Razão	Impacta muito	Impacta pouco	Não impacta
Buscar conhecimento e desenvolver novas habilidades			
Gosto pela área			
Ser uma área desafiadora e interessante			
Pressão da família ou amigos			
Mostrar a todos que consigo / medo de falhar			
O ambiente universitário / amigos / socialização			
Aplicar conhecimento no meu trabalho			
Conseguir um bom posicionamento no mercado de trabalho			
Completar e receber certificação			
Satisfação pessoal em alcançar o êxito na graduação			
Não tenho uma motivação em específico para continuar			

3. Qual seu nível de intenção de continuar e finalizar o curso (1 a 5)?

Penso seriamente em desistir	1	2	3	4	5	**Nunca pensei em desistir e não tenho dúvidas de que quero continuar**

4. Quais podem ser eventuais razões para sua desistência do curso? (pode selecionar mais de uma)
   a. [ ] Nenhuma razão, pois não penso em desistir
   b. [ ] Dificuldade ou desgosto com programação
   c. [ ] Dificuldade com matemática
   d. [ ] Falta de afinidade com o curso
   e. [ ] Dificuldade em conciliar emprego e estudo
   f. [ ] Outras Razões:

   ____

5. Com relação ao seu desempenho no curso em geral: (1 a 5)

Ruim. Longe de atingir minhas expectativas	1	2	3	4	5	Ótimo. Tem superado minhas expectativas

6. Qual sua percepção da universidade?

Fator	PÉSSIMO / DISCORDO fortemente	Ruim / Discordo	Regular/ Neutro	Bom / Concordo	ÓTIMO / CONCORD

		(precisa melhorar bastante)		(poderia melhorar)	O fortemente
Fornece apoio ao estudante de forma adequada (atendimento, secretaria, etc.)					
Possui recursos de aprendizagem adequados (laboratórios, equipamentos, internet)					
Possui ambiente virtual de apoio adequado					
Os professores aparentam estar satisfeitos com sua carreira					
Os professores possuem formação e qualificação adequadas					
Comentários:					

7. Qual seu comportamento nos estudos?

Fator	PÉSSIMO/ DISCORDO fortemente	Ruim / Discordo (precisa melhorar bastante)	Regular/ Neutro	Bom / Concordo (poderia melhorar)	ÓTIMO / CONCORDO fortemente
Sente-se preparado para o estudo					
Interajo com estudantes fora do ambiente acadêmico					
Possuo senso de pertencimento à universidade					
Interajo frequentemente com diferentes estudantes					
Participo frequentemente de discussões online ou presenciais com alunos e professores					
Trabalho frequentemente em grupo com outros estudantes					
Sou assíduo nas aulas					
Cumpro as atividades e prazos					
Estudo de maneira correta e gerencio bem o tempo para minhas atividades					
Procuro fazer mais que solicitado					
Faço o possível para me destacar na turma					

Comentários:	

8. Qual sua avaliação de cada um dos seguintes fatores no seu curso?

Fator	PÉSSIMO / DISCORDO fortemente	Ruim / Discordo (precisa melhorar bastante)	Regular / Neutro	Bom / Concordo (poderia melhorar)	ÓTIMO / CONCORDO fortemente
Possui acesso a instalações de importância industrial e relativamente atualizadas					
Alinhamento com o mercado de trabalho					
Tipo do curso e disciplinas adequado (foco do curso na computação, horários, etc.)					
Matriz curricular e programa das disciplinas (conteúdo) adequados					
Possui mecanismos para facilitar a inserção no mercado de trabalho e perspectivas de futuro					
Equilíbrio entre áreas do conhecimento, o que permite uma visão sistêmica (multidisciplinar)					
A qualidade de ensino oferecida é adequada					
Comentários:					

9. Qual sua avaliação do ensino e professor da disciplina de ________________________________ ?

(   ) Não estou matriculado nessa disciplina – nesse caso não precisa responder as questões abaixo.

Fator	PÉSSIMO/ DISCORDO fortemente	Ruim / Discordo (precisa melhorar bastante)	Regular / Neutro	Bom / Concordo (poderia melhorar)	ÓTIMO / CONCORDO fortemente
Permite participação ativa do estudante (aprendizagem fazendo, prática, atividades desenvolvidas pelos alunos)					

Promove diversão (atividades prazerosas)					
Promove desafios acadêmicos (atividades que promovem a reflexão, criação de estratégias, etc.)					
Permite aprendizagem com pares (troca de conhecimento com colegas)					
Aulas com diversas abordagens pedagógicas e não são baseadas unicamente ou em sua maioria em livro texto					
Promove o espírito de equipe (união da classe, atividades em grupo, colaboração)					
Desenvolve habilidades para prática fora da sala de aula					
Disciplina possui utilidade e aplicação futura do conteúdo aprendidos					
Permite que estudantes participem da escolha de solução de problemas e tomada de decisões durante a aula					
Professor e avaliações recompensam de forma justa os estudantes que mais se dedicam					
As informações detalhadas do conteúdo do curso são disponibilizadas					
Nível de dificuldade adequado					
Professor deixa claro os objetivos e se o estudante está obtendo desempenho satisfatório ao longo do curso/disciplina					
Professor não expõe alunos com dificuldades para a turma					
O professor aparenta estar satisfeito com sua carreira					
Professor possui formação e qualificação adequadas					
Comentários:					

APPENDIX D – Questionnaire EVC Light (in Portuguese)

Pergunta	1 = Discordo totalmente 7 – Concordo totalmente						
Estou confiante que aprenderei o conteúdo e terei sucesso na disciplina.	1	2	3	4	5	6	7
Estou achando as aulas da disciplina úteis para o curso	1	2	3	4	5	6	7
A disciplina exige muito tempo e estou tendo de me abdicar de outras coisas	1	2	3	4	5	6	7
Estou entendendo e conseguindo aprender o conteúdo	1	2	3	4	5	6	7
Estou entendendo a importância dessa disciplina no curso	1	2	3	4	5	6	7
Não estou conseguindo tempo necessário para me dedicar a essa disciplina	1	2	3	4	5	6	7
Comentário:							

APPENDIX E – Questionnaire of Pre-university Factors (in Portuguese)

Sobre seu gosto por computação ANTES de fazer o curso, responda as 3 próximas questões.

	TD	VD	N	VA	TA
1. Gosto de programação e tudo que envolve aspectos de criação de programas					
2. Gosto de hardware e equipamentos de informática, redes, etc					
3. Gosto de jogos e utilizar dispositivos e aplicações envolvendo tecnologia					

TD = totally disagree, VD – very disagree, N – neutral, VA – very agree, TA – totally agree

Sobre sua experiência e conhecimento do curso e da área ANTES de ingressar no curso.

	TD	VD	N	VA	TA
4. Conheço e entendo os objetivos do seu curso e a diferença para os demais cursos da área					
5. Conheço o conteúdo do seu curso (áreas, disciplinas) e que você terá de estudar					
6. Conheço o mercado de trabalho e entende o perfil e atribuições de um profissional da área					
7. Trabalho e tenho experiência na área de informática em geral					
8. Já conheço programação e desenvolvo softwares					
9. Tive computação no ensino fundamental e/ou médio, aprendendo desde informática básica até programação					
10. Em geral, tenho um bom desempenho escolar					
11. Nas disciplinas de matemática, tenho um bom desempenho					

Responda qual o nível de influência de cada fator na sua opção pela computação.

	TD	VD	N	VA	TA
12. Conteúdo do curso (Gostar e achar a área de computação interessante)					
13. Busca por novos conhecimentos (Interesse em aprender mais sobre computação)					
14. Perspectiva de carreira e emprego (Ser uma área com futuro promissor)					
15. Influência dos seus pais					
16. Influência de amigos, professores ou outras pessoas					
17. O desafio em conseguir passar e obter sucesso no curso (desafio pessoal)					
18. Desafios que as atividades de computação promovem (ser uma área difícil e desafiadora)					
19. Ser a única ou uma das poucas opções de curso para sua realidade (região, disponibilidade financeira)					
20. Não ter passado em outro curso que desejaria					

21. Pressão da família/sociedade em fazer algum curso superior					

APPENDIX F – Questionnaire of Professor Perception of Students (in Portuguese)

Qual sua percepção sobre a motivação/desempenho/interesse de cada estudante? (caso não possua uma percepção clara, pode ignorar ou responder "0").

Aluno	SEMANA 1							SEMANA 2							SEMANA 3							SEMANA 4						
	-3	-2	-1	0	+1	+2	+3	-3	-2	-1	0	+1	+2	+3	-3	-2	-1	0	+1	+2	+3	-3	-2	-1	0	+1	+2	+3

Legenda:

Variação negativa			Sem variação / não percebida	Variação positiva		
-3	-2	-1	0	+1	+2	+3

APPENDIX G - ARFF file structure

```
@relation QueryResult
@attribute week numeric
@attribute SCALE_Index numeric
@attribute SCALE_intrinsic numeric
@attribute SCALE_extrinsec numeric
@attribute SCALE_amotivation numeric
@attribute age numeric
@attribute gender {Feminino,Masculino}
@attribute entranceOrder {Primeira,Segunda,'Demais cha'}
@attribute cota {SIM}
@attribute taste1 numeric
@attribute taste2 numeric
@attribute taste3 numeric
@attribute know1 numeric
@attribute know2 numeric
@attribute know3 numeric
@attribute know4 numeric
@attribute know5 numeric
@attribute know6 numeric
@attribute know7 numeric
@attribute know8 numeric
@attribute reason1 numeric
@attribute reason2 numeric
@attribute reason3 numeric
@attribute reason4 numeric
@attribute reason5 numeric
@attribute reason6 numeric
@attribute reason7 numeric
@attribute reason8 numeric
@attribute reason9 numeric
@attribute reason10 numeric
@attribute entrada {VEST,ENEM}
@attribute status {REPR,APR}
@attribute acessos numeric
@attribute q1 numeric
@attribute q2 numeric
@attribute q3 numeric
@attribute q4 numeric
@attribute q5 numeric
@attribute q6 numeric
@attribute expect numeric
```

```
@attribute valuel numeric
@attribute cost numeric
@attribute evc numeric
@attribute perception numeric
```

APPENDIX H – Prediction report per class

Aluno	Predict	Insuccess Risk	Initial motivat. index	Intrinsec	Intrinsic index	Extrinsec	% Extrinsic index	Amotiv.	% Amotiv. index	Expect Index	Value Index	Cost Index	EVC Index	EVC Class
Aluno 1	REPR	95,14%	12.333	18	0,300	31	0,517	4	0,200	0,583	0,833	0,417	0,667	0,800
Aluno 2	REPR	99,22%	23.333	38	0,633	47	0,783	5	0,250	0,583	0,333	0,500	0,472	0,800
Aluno 3	REPR	37,82%	24.000	42	0,700	42	0,700	4	0,200	1,000	1,000	0,917	0,694	0,800
Aluno 4	REPR	94,69%	23.000	35	0,583	46	0,767	4	0,200	0,500	0,917	0,250	0,722	0,800
Aluno 5	REPR	95,62%	25.333	42	0,700	49	0,817	5	0,250	0,750	0,750	0,667	0,611	0,8002
Aluno 6	REPR	50,81%	26.667	44	0,733	48	0,800	4	0,200	0,833	1,000	0,583	0,750	0,8002
Aluno 7	APR	29,14%	32.000	60	1,000	60	1,000	8	0,400	1,000	1,000	0,083	0,972	0,8002
Aluno 8	APR	4,60%	23.667	52	0,867	31	0,517	4	0,200	0,833	1,000	0,167	0,889	0,8002
Aluno 9	APR	12,54%	31.667	57	0,950	50	0,833	4	0,200	1,000	1,000	0,167	0,944	0,8002
Aluno 10	REPR	99,36%	22.000	45	0,750	36	0,600	5	0,250	0,667	1,000	0,417	0,750	0,8002
Aluno 11	APR	12,27%	33.333	52	0,867	60	1,000	4	0,200	1,000	1,000	0,250	0,917	0,8002
Aluno 12	APR	4,60%	35.667	59	0,983	60	1,000	4	0,200	1,000	1,000	0,250	0,917	0,8002

APPENDIX I – TCLE (Termo de Consentimento Livre e Esclarecido)

Você está sendo convidado(a) a participar da pesquisa: *Impacto da motivação no sucesso de estudantes de graduação em cursos de computação,* visando a realização de um estudo piloto com o objetivo de mapear quais as principais motivações para alunos fazerem curso de computação e também quais fatores impactam sua motivação ao longo do curso. O projeto está sendo coordenado pelo Prof. Raul Sidnei Wazlawick – Laboratório de Sistemas do Conhecimento do INE – Departamento de Informática e Estatística da UFSC – Universidade Federal de Santa Catarina e faz parte da tese do doutorando Pablo Schoeffel, do Programa de Pós Graduação em Computação (PPGCC) da UFSC.

A sua participação nessa pesquisa consistirá em preencher um questionário com a sua percepção dos principais fatores que lhe motivaram a cursar um curso na área da computação. Todos os dados coletados serão confidenciais de forma a assegurar a sua privacidade.

Os resultados divulgados em congressos ou revistas científicas serão apresentados de forma a não o identificar. Fotos e vídeos poderão ser produzidos com o objetivo de evidenciar a realização da pesquisa em publicações científicas.

Os pesquisadores estão também disponíveis antes, durante e depois da pesquisa para esclarecimentos e acompanhamento.

A participação nesta pesquisa não traz complicações, eventualmente apenas um pequeno cansaço ou aborrecimento ao responder os questionários. Com o objetivo de minimizar qualquer risco, serão apresentados claramente o objetivo e execução da pesquisa seguindo a ética em pesquisa como também foram criados questionários o menos extenso possível e com respostas rápidas e diretas. Os pesquisadores serão os únicos a ter acesso aos dados e tomarão todas as providências necessárias para manter o sigilo, mas sempre existe a remota possibilidade da quebra do sigilo, mesmo que involuntário e não intencional, cujas consequências serão tratadas nos termos da lei.

Caso você tenha qualquer despesa inerente da pesquisa, como transporte ou alimentação, as mesmas serão ressarcidas pelos pesquisadores.

A pesquisa não trará benefícios imediatos a você, mas poderá ajudar a gestores, coordenadores de curso e professores a melhorarem o curso e a interação com os alunos, visando reduzir o índice de falhas e reprovações em cursos de computação.

A participação é gratuita e voluntária. Caso você tenha algum prejuízo material ou imaterial em decorrência da pesquisa poderá solicitar indenização, de acordo com a legislação vigente e amplamente consubstanciada. A legislação brasileira não permite que você tenha qualquer compensação financeira pela sua participação em pesquisa, mas você não terá nenhuma despesa advinda da sua participação na pesquisa. Caso alguma despesa extraordinária associada à pesquisa venha a ocorrer, você será ressarcido nos termos da lei.

Desta forma, considerando que os riscos implicados nesta pesquisa são mínimos aos participantes, a equipe de pesquisadores juntamente com a equipe multidisciplinar da UFSC oferecerá apoio e suporte para eventuais problemas e acontecimentos que venham ocorrer durante a pesquisa.

A qualquer momento você pode desistir da sua participação desse projeto e retirar o seu consentimento sem qualquer tipo de prejuízo em sua relação à pesquisa.

Caso você aceite participar da pesquisa, o TCLE precisa ser assinado por você e pelo pesquisador responsável em duas vias, sendo que uma das vias ficará com você e a outra será arquivada pelos pesquisadores.

O pesquisador responsável explicitamente declara que o Termo de Consentimento Livre e Esclarecido está em conformidade com as exigências contidas no item IV.3 da Resolução 466/12. Em caso de dúvidas ou notificação de acontecimentos não previstos entrar em contato com CEPSH - Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Santa Catarina, Prédio Reitoria II, R: Desembargador Vitor Lima, nº 222, sala 401, Trindade, Florianópolis/SC, CEP 88.040-400, Contato: (48) 3721-6094, cep.propesq@contato.ufsc.br, pelo qual o projeto de pesquisa foi aprovado.

O CEPSH é um órgão colegiado interdisciplinar, deliberativo, consultivo e educativo, vinculado à Universidade Federal de Santa Catarina, mas independente na tomada de decisões, criado para defender os interesses dos participantes da pesquisa em sua integridade e dignidade e para contribuir no desenvolvimento da pesquisa dentro de padrões éticos.

____

Pesquisador responsável

Prof. Dr. Raul Sidnei Wazlawick

INE – Departamento de Informática e Estatística/UFSC-Universidade Federal de Santa Catarina, *Campus Universitário - Trindade - Florianópolis/SC, CEP 88040-900* email: raul.wazlawick@ufsc.br

CONSENTIMENTO DO PARTICIPANTE

Eu, ____, RG__________________________, abaixo assinado, concordo em participar do estudo: “Impacto da motivação no sucesso de estudantes de graduação em cursos de computação”.

Fui devidamente informado(a) e esclarecido(a) sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes da minha participação. Foi-me garantido que posso retirar meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade.

Florianópolis, ___ de ___________de .

Assinatura do participante:

____

APPENDIX J – Prediction results for all datasets and classifiers combinations

Week	Classifier	Dataset	Accuracy	Recall	AUC/ROC	F-measure
0	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	68.786	0.748	0.746	0.733
1	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	78.333	0.925	0.884	0.851
2	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	67.347	0.746	0.736	0.733
3	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	60.897	0.727	0.656	0.677
4	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	65.000	0.714	0.662	0.699
5	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	63.253	0.699	0.654	0.681
6	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	68.047	0.771	0.750	0.733
7	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	71.006	0.792	0.775	0.756
8	AdaBoostM1	SelectedFreshmenAVG_3FilterAllAll	68.605	0.796	0.701	0.743
0	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	72.832	0.798	0.749	0.771
1	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	80.000	0.925	0.876	0.861
2	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	68.367	0.831	0.756	0.760
3	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	64.744	0.796	0.668	0.718
4	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	70.000	0.868	0.688	0.767
5	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	66.868	0.774	0.702	0.724
6	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	73.373	0.813	0.739	0.776
7	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	70.414	0.802	0.728	0.755
8	AdaBoostM1-SMO	AllAtributesFreshmenAVGFilterAllAll	67.442	0.745	0.708	0.723
0	BayesNet	SelectedFreshmenAVGVARSMOTEAll	71.676	0.677	0.769	0.732
1	BayesNet	SelectedFreshmenAVGVARSMOTEAll	81.667	0.825	0.831	0.857
2	BayesNet	SelectedFreshmenAVGVARSMOTEAll	71.429	0.763	0.768	0.763
3	BayesNet	SelectedFreshmenAVGVARSMOTEAll	63.462	0.705	0.602	0.685
4	BayesNet	SelectedFreshmenAVGVARSMOTEAll	63.750	0.703	0.644	0.688
5	BayesNet	SelectedFreshmenAVGVARSMOTEAll	62.651	0.656	0.670	0.663
6	BayesNet	SelectedFreshmenAVGVARSMOTEAll	66.272	0.688	0.717	0.698
7	BayesNet	SelectedFreshmenAVGVARSMOTEAll	65.681	0.667	0.725	0.688
8	BayesNet	SelectedFreshmenAVGVARSMOTEAll	65.116	0.633	0.722	0.674
0	IBk	SelectedFreshmenAVG_3FilterAllAll	70.520	0.808	0.654	0.758
1	IBk	SelectedFreshmenAVG_3FilterAllAll	66.667	0.750	0.625	0.750

2	IBk	SelectedFreshmenAVG_3FilterAllAll	67.347	0.746	0.650	0.733
3	IBk	SelectedFreshmenAVG_3FilterAllAll	61.539	0.727	0.599	0.681
4	IBk	SelectedFreshmenAVG_3FilterAllAll	61.875	0.703	0.595	0.677
5	IBk	SelectedFreshmenAVG_3FilterAllAll	64.458	0.710	0.614	0.691
6	IBk	SelectedFreshmenAVG_3FilterAllAll	62.130	0.656	0.618	0.663
7	IBk	SelectedFreshmenAVG_3FilterAllAll	63.314	0.656	0.630	0.670
8	IBk	SelectedFreshmenAVG_3FilterAllAll	64.535	0.725	0.635	0.700
0	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	71.098	0.687	0.723	0.731
1	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	81.667	0.825	0.804	0.857
2	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	63.291	0.681	0.631	0.688
3	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	61.475	0.677	0.605	0.641
4	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	65.517	0.778	0.626	0.737
5	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	60.377	0.577	0.549	0.588
6	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	81.539	0.730	0.816	0.818
7	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	56.522	0.455	0.530	0.500
8	J48	V2_SelectedFreshmenSMOTECorrelationAttributeEval	74.419	0.680	0.732	0.756
0	LMT	SelectedFreshmenAVG_3FilterAllAll	72.254	0.788	0.814	0.765
1	LMT	SelectedFreshmenAVG_3FilterAllAll	73.333	0.800	0.774	0.800
2	LMT	SelectedFreshmenAVG_3FilterAllAll	71.429	0.780	0.767	0.767
3	LMT	SelectedFreshmenAVG_3FilterAllAll	66.026	0.716	0.685	0.704
4	LMT	SelectedFreshmenAVG_3FilterAllAll	70.000	0.769	0.752	0.745
5	LMT	SelectedFreshmenAVG_3FilterAllAll	68.675	0.774	0.767	0.735
6	LMT	SelectedFreshmenAVG_3FilterAllAll	75.148	0.781	0.793	0.781
7	LMT	SelectedFreshmenAVG_3FilterAllAll	73.965	0.792	0.804	0.776
8	LMT	SelectedFreshmenAVG_3FilterAllAll	72.674	0.786	0.767	0.766
0	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	68.786	0.758	0.743	0.735
1	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	76.667	0.875	0.841	0.833
2	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	69.388	0.712	0.683	0.737
3	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	70.513	0.784	0.761	0.750
4	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	69.375	0.758	0.751	0.738
5	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	71.084	0.731	0.779	0.739

6	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	69.823	0.771	0.774	0.744
7	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	68.639	0.792	0.739	0.742
8	MultilayerPerceptron	AllAtributesFreshmenAVG_innerFilterAllInfoGainAttr...	69.186	0.755	0.739	0.736
0	RandomForest	AllAttributesFreshmenFilterAllAll	71.098	0.727	0.818	0.742
1	RandomForest	AllAttributesFreshmenFilterAllAll	75.000	0.825	0.837	0.815
2	RandomForest	AllAttributesFreshmenFilterAllAll	67.089	0.787	0.694	0.740
3	RandomForest	AllAttributesFreshmenFilterAllAll	66.393	0.726	0.722	0.687
4	RandomForest	AllAttributesFreshmenFilterAllAll	72.414	0.833	0.687	0.790
5	RandomForest	AllAttributesFreshmenFilterAllAll	62.264	0.539	0.662	0.583
6	RandomForest	AllAttributesFreshmenFilterAllAll	78.462	0.838	0.842	0.816
7	RandomForest	AllAttributesFreshmenFilterAllAll	56.522	0.636	0.735	0.583
8	RandomForest	AllAttributesFreshmenFilterAllAll	69.767	0.640	0.691	0.711
0	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	71.098	0.788	0.698	0.757
1	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	85.000	0.900	0.825	0.889
2	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	68.367	0.814	0.650	0.756
3	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	69.231	0.841	0.671	0.755
4	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	70.625	0.857	0.682	0.769
5	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	68.675	0.785	0.673	0.737
6	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	73.373	0.823	0.720	0.778
7	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	72.189	0.823	0.706	0.771
8	SMO	AllAtributesFreshmenAVGFilterAllCorrelationAttribu...	68.605	0.765	0.673	0.735
0	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttributeEva...	72.254	0.788	0.814	0.765
1	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttributeEva...	75.000	0.825	0.771	0.815
2	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttributeEva...	73.469	0.797	0.775	0.783
3	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttributeEva...	68.590	0.727	0.729	0.723
4	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttributeEva...	69.375	0.780	0.760	0.744
5	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttributeEva...	70.482	0.785	0.766	0.749

6	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttribute Eva...	75.148	0.781	0.799	0.781
7	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttribute Eva...	73.965	0.792	0.804	0.776
8	SimpleLogistic	SelectedFreshmenAVG_3FilterAllInfoGainAttribute Eva...	72.674	0.786	0.767	0.766
0	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	71.098	0.657	0.769	0.722
1	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	80.000	0.850	0.799	0.850
2	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	70.886	0.681	0.717	0.736
3	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	61.475	0.645	0.636	0.630
4	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	58.621	0.667	0.611	0.667
5	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	64.151	0.577	0.684	0.612
6	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	75.385	0.730	0.765	0.771
7	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	86.957	0.818	0.830	0.857
8	NaiveBayes	V2_AllAttributesFreshmenSMOTEInfoGainAttribute Eval	76.744	0.760	0.813	0.792

APPENDIX K – Quality evaluation of selected work of the mapping study of motivation factors

Author(s)	Title	Year	Problem relevance	Contribution	Results	Average
S Mamone	Empirical study of motivation in an entry level programming course	1992	3	1	3	2,33
P Byrne		1999	4	1	2	2,33
W Soerjaningsih	Student Outcomes, Learning Environment, Logical Thinking and Motivation AmongComputing Students in an Indonesian University	2001	4	3	3	3,33
Jenkins Tony	The motivation of students of programming	2001	4	3	3	3,33
Tony Jenkins ; John Davy	Diversity and motivation in introductory programming	2002	2	2	1	1,67
M Robey, Brian R. Von Konsky ; Jim Ivins ; Susan J. Gribble ; Allan Loh ; David Cooper	Student self-motivation: lessons learned from teaching first year computing	2006	4	3	2	3,00
Y Takemura, Hideo Nagumo; Kuo-Li Huang; Kenichi Matsumoto	Analysis of the relation between the teaching materials and motivation in programming education	2007	3	3	3	3,00
Law Kris M. Y. ; Lee Victor C. S. ; Yu Y. T.	Learning Motivation in E-Learning Facilitated ComputerProgramming Courses	2010	4	4	3	3,67
HM Sayers, MA Nicell, A Hinds	TRANSITION, ENGAGEMENT AND RETENTION OF FIRST YEAR COMPUTINGSTUDENTS	2010	2	3	2	2,33
Nikula Uolevi; Gotel Orlena; Kasurinen Jussi	A Motivation Guided Holistic Rehabilitation of the FirstProgramming Course	2011	4	3	3	3,33
Alev Ates	SELF-EFFICACY BELIEFS, ACHIEVEMENT MOTIVATION AND GENDER AS RELATED TO EDUCATIONAL SOFTWAREDEVELOPMENT	2011	3	3	3	3,00
Magana A. J.; Mathur J. I.	Motivation, Awareness, and Perceptions of Computational Science	2012	3	3	2	2,67
NFA Zainal, S Shahrani, NFM Yatim, Rohizah Abd	Students' perception and motivation towards programming	2012	3	3	3	3,00

Rahman, Masura Rahmat, Rodziah Latih						
AK Peters, A Pears	Engagement in Computer Science and IT--What! A Matter of Identity?	2013	3	2	2	2,33
G Kanaparan, R Cullen, D Mason	Self-Efficacy and Engagement as Predictors of Student Programming Performance	2013	4	3	0	2,33
P Figas, G Hagel, A Bartel	The furtherance of motivation in the context of teaching software engineering	2013	4	2	1	2,33
L Payne	Why do students choose computing?: influences, perceptions and engagement	2013	4	4	1	3,00
A Abdullah, TY Yih	Implementing Learning Contracts in a Computer Science Course as a Tool to Develop and Sustain Student Motivation to Learn	2014	2	2	3	2,33
H Tsukamoto, Y Takemura, H Nagumo, Akito Monden ; Ken-ichi Matsumoto	Prediction of the change of learners' motivation in programming education for non-computing majors	2014	4	3	2	3,00
O Debdi, M Paredes-Velasco	Relationship between learning styles, motivation and educational efficiency in students ofcomputer science	2014	4	3	4	3,67
J Rao, F Wang	Research on Motivation of Computer Science Students in Financial College	2014	3	3	4	3,33
S Alhazbi	ARCS-based tactics to improve students' motivation in computer programming course	2015	4	3	2	3,00
N Eltegani, L Butgereit	Attributes of students engagement in fundamental programming learning	2015	4	3	1	2,67
SC Ngan, KMY Law	Exploratory Network Analysis of Learning Motivation Factors in e-Learning FacilitatedComputer Programming Courses	2015	4	4	4	4,00
J Sinclair, M Butler, M Morgan, S Kalvala	Measures of student engagement in computer science	2015	4	3	2	3,00
DF Shell, LK Soh, AE Flanigan, Markeya S. Peteranetz	Students' Initial Course Motivation and Their Achievement and Retention in College CS1Courses	2016	4	3	4	3,67
Kori Kuelli ; Pedaste Margus ; Leijen Aeli ; Tonisson Eno	The Role of Programming Experience in ICT Students' LearningMotivation and Academic Achievement	2016	4	3	4	3,67

Mccartney R. ; Boustedt J. ; Eckerdal A. ; Sanders K. ; Thomas L. ; Zander C.	Why computing students learn on their own: Motivation for self-directed learning of computing	2016	4	3	3	3,33
I Bosnić, I Čavrak, M Orlić, M Žagar…	Student motivation in distributed software development projects	2011	4	4	3	3,67
N Pratheesh, T Devi	Assessment of student's learning style and engagement in traditional based software engineering education	2013	3	3	2	2,67
A Steele	First Year Programming: Engagement vs. Success Measurements from UCOL	2010	4	3	3	3,33
Settle Amber; Sedlak Brian	Computing Educator Attitudes about Motivation	2016	4	4	3	3,67

ANNEX A – AMS Questionnaire (in Portuguese)

Por que venho à universidade?

Participante: Matrícula:

Usando a escala abaixo, indique – por favor – em que extensão cada um dos itens corresponde, atualmente, a uma das razões porque você vem à Universidade.

1 - Nenhuma correspondência
2- Pouca correspondência
3- Moderada correspondência
4- Muita correspondência
5- Total correspondência

1. Porque preciso do diploma, ao menos, a fim de conseguir uma ocupação bem remunerada, no futuro
2. Porque sinto satisfação e prazer enquanto aprendo coisas novas
3. Porque acho que a formação universitária ajuda a me preparar melhor para a carreira que escolhi
4. Porque gosto muito de vir à universidade
5. Honestamente, não sei; acho que estou perdendo meu tempo na universidade
6. Pelo prazer que sinto quando supero a mim mesmo nos estudos
7. Para provar a mim mesmo que sou capaz de completar o curso
8. A fim de obter um emprego de prestígio, no futuro
9. Pelo prazer que sinto quando descubro coisas novas que nunca tinha visto ou conhecido antes
10. Porque o curso me capacitará, no final, a entrar no mercado de trabalho de uma área que eu gosto
11. Porque, para mim, a universidade é um prazer
12. Já tive boas razões para isso; agora, entretanto, eu me pergunto se devo continuar
13. Pelo prazer que sinto quando supero a mim mesmo em alguma de minhas realizações pessoais
14. Por causa do fato que me sinto importante quando sou bem sucedido na universidade
15. Porque quero levar uma boa vida no futuro
16. Pelo prazer que tenho em ampliar meu conhecimento sobre assuntos que me atraem
17. Porque isso me ajudará a escolher melhor minha orientação profissional

18. Pelo prazer que tenho quando me envolvo em debates com professores interessantes
19. Não atino (percebo) porque venho à universidade e, francamente, não me preocupo com isso
20. Pela satisfação que sinto quando estou no processo de realização de atividades acadêmicas difíceis
21. Para mostrar a mim mesmo que sou uma pessoa inteligente
22. A fim de ter uma boa remuneração no futuro
23. Porque meus estudos permitem que continue a aprender sobre muitas coisas que me interessam
24. Porque eu creio que a formação universitária aumentará minha competência como profissional
25. Pela euforia que sinto quando leio sobre vários assuntos interessantes
26. Não sei; não entendo o que estou fazendo na universidade
27. Porque a universidade me permite sentir uma satisfação pessoal na minha busca por excelência na formação
28. Porque quero mostrar a mim mesmo que posso ter sucesso nos meus estudos

Diante do exposto, assinalo com um traço na linha abaixo o grau de minha motivação global para prosseguir nos estudos

0 ____ 100

2, 9, 16, 23: Intrinsic motivation - to know
6, 13, 20, 27: Intrinsic motivation - toward accomplishment
4, 11, 18, 25: Intrinsic motivation - to experience stimulation
3, 10, 17, 24: Extrinsic motivation - identified
7, 14, 21, 28: Extrinsic motivation - introjected
1, 8, 15, 22: Extrinsic motivation - external regulation
5, 12, 19, 26: Amotivation