FEDERAL UNIVERSITY OF SANTA CATARINA
TECHNOLOGY CENTER
AUTOMATION AND SYSTEMS DEPARTMENT
UNDERGRADUATE COURSE IN CONTROL AND AUTOMATION ENGINEERING

Luis Felipe Araújo Both

Feature selection for predicting milling tool condition by using machine learning methods.

Aachen
2022

Luis Felipe Araújo Both

Feature selection for predicting milling tool condition by using machine learning methods.

Final report of the subject DAS5511 (Course Final Project) as a Concluding Dissertation of the Undergraduate Course in Control and Automation Engineering of the Federal University of Santa Catarina.
Supervisor: Prof. Marcelo Ricardo Stemmer, Dr.
Supervisora Zhen Zhen, Eng.

Aachen
2022

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

Both, Luis Felipe Araújo
Feature selection for predicting milling tool condition by using machine learning methods. / Luis Felipe Araújo Both ; orientadora, Zhen Zhen, 2022.
85 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Centro Tecnológico, Graduação em Engenharia de Controle e Automação, Florianópolis, 2022.

Inclui referências.

1. Engenharia de Controle e Automação. I. Zhen, Zhen. II. Universidade Federal de Santa Catarina. Graduação em Engenharia de Controle e Automação. III. Título.

Luis Felipe Araújo Both

Feature selection for predicting milling tool condition by using machine learning methods.

This dissertation was evaluated in the context of the subject DAS5511 (Course Final Project) and approved in its final form by the Undergraduate Course in Control and Automation Engineering

Florianópolis, December 12, 2022.

Prof. Hector Bessa Silveira, Dr.
Course Coordinator

Examining Board:

Prof. Marcelo Ricardo Stemmer, Dr.
Advisor
UFSC/CTC/DAS

Zhen Zhen, Eng.
Supervisor
Institute Fraunhofer IPT

Thiago Raulino Dal Pont, Eng.
Evaluator
UFSC/CTC/DAS

Prof. Eduardo Camponogara, Dr.
Board President
UFSC/CTC/DAS

This work is dedicated to my parents, my sister and my girlfriend, who accompanied me throughout my undergraduate course.

ACKNOWLEDGEMENTS

I would like to express here my gratitude to all the people who were present with me throughout the entire duration of my undergraduate course, especially my colleagues and course friends who always were united in search of the best ways to solve problems that we faced on our professional path, for the uncountable times we helped each other and for offering and sharing opportunities and different visions that we were not able to reach by ourselves.

I would also like to thank the institutions that provided me with the opportunity for personal and professional growth. First, to the Federal University of Santa Catarina (UFSC), which provided me and many other students with an academic education full of ethics, respect and quality despite financial and political difficulties. I sincerely wish all young students have an education as good as the one offered by UFSC.

Secondly to Institute Fraunhofer for Production Technology for these valuable years of professional apprenticeship, which helped me realize this work and many other activities. It was only possible thanks to my supervisor Zhen Zhen and my colleague and department head, Tae Hun Lee. My sincere thanks to you both.

My special thanks to my closest friends Marcio, Busiquia, Ivan, Makino, Jean, Édina, Karini, Carol (A.K.A Titi) and Gabriela for the companionship and beers shared since my first week at UFSC, stimulating discussions and experiences. Also, to Automação 13-2 classmates for all the moments we had together. To my girlfriend, Joyce for all the support and encouragement, for being my confidant and standing by my side all the time.

Last but not least, my family, Felipe, Zenilda and Michelly. You are my world! Everything I do is because of you and for you. I have no words to describe how much I am grateful for your encouragement, support, both emotional and financial, and for standing by my side in every life situation even though we were a few thousand kilometers apart.

My most profound appreciation to all of you.

“Nossa maior fraqueza é a desistência.
O caminho mais certeiro para o sucesso
é sempre tentar apenas uma vez mais.”
(Thomas Edison)

ABSTRACT

The project aims to study feature selection techniques to improve a tool wear prediction system in high-precision milling processes. The chosen approach combines the use of sensor data collected through practical tests performed on a high-precision CNC milling machine with tool flank wear measurements obtained under a microscope. In order to select the best feature selection technique for the development of the project, thorough state-of-the-art research was carried out, raising the possible best techniques to be later implemented in the project. In this phase of the project, both feature selection techniques and machine learning techniques were studied; in addition, a study on the phenomenon of tool wear was conducted. The implemented data acquisition system records data from several different sensors, so the document addresses the hardware and software setup for the used acquisition system, parameter planning and data analysis. The next section of the document discusses the pre-processing strategies of the collected data, such as eliminating noise and sensors that were acquired but are not part of the project. At this stage of the project, all datasets containing vital information for the project are produced. In the end, the project carried out discusses the feature selection techniques, points out the best ones and why they were chosen. The results are compared with the tool wear prediction without any feature selection technique. Finally, it is possible to conclude that, for the specific case of study, we can considerably reduce the number of features and, consequently, the number of sensors used in data collection, reducing the cost of the operation, and providing high accuracy of tool wear prediction.

Keywords: Feature Selection. Prediction of Tool Wear. Machine Learning. High Precision Milling. Data Analysis.

RESUMO

O projeto visa o estudo de técnicas de seleção de features com o intuito de melhorar um sistema de predição de desgaste de ferramenta em processos de fresamento de alta precisão. A abordagem escolhida combina a utilização de dados de sensores coletados através de testes práticos realizados em uma fresadora CNC de alta precisão com medições de desgaste de flanco da ferramenta obtidos em um microscópio. Com o intuito de selecionar a melhor técnica de seleção de features para o desenvolvimento do projeto, uma minuciosa pesquisa do estado da arte foi realizada, levantando as possíveis melhores técnicas a serem depois implementadas no projeto. Nesta fase do projeto, tanto técnicas de seleção de features quanto técnicas de machine learning foram estudadas, além disso, um estudo sobre o fenômeno de desgaste de ferramenta foi conduzido. O sistema de aquisição de dados implementado, grava dados de diversos sensores diferentes, sendo assim, o documento aborda o setup de hardware e software para o sistema de aquisição utilizado, planejamento dos parâmetros e análise dos dados. A próxima seção do documento faz uma abordagem das estratégias de pré-processamento dos dados coletados, como eliminação de ruídos e de sensores que foram adiquiridos mas não fazem parte do projeto. Nesta etapa do projeto, todos os datasets contendo as informações vitais para o projeto são produzidos. Ao final, o projeto realizado discute sobre as técnicas de seleção de features, aponta as melhores e o porque das mesmas serem escolhidas. Os resultados são comparados com a predição do desgaste da ferramenta sem nenhuma técnica de seleção de features. Consegue-se então por fim concluir que, para o caso de estudo especifico, podemos reduzir consideravelmente o número de features e consequentemente o número de sensores empregados na coleta de dados reduzindo o custo da operação, e garantir uma precisão alta de predição de desgaste de ferramenta.

Palavras-chave: Seleção de Features. Predição de Desgaste de Ferramenta. Machine Learning. Fresamento de Alta Precisão. Análise de Dados.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AE	Acoustic Emission
Ai	Analog Input
AI	Artificial Intelligence
AIHS	High Speed Analog Input
ANN	Artificial Neural Network
CNC	Computer Numerical Control
DAQ	Data Acquisition Device
FFT	Fast Fourier Transform
HMM	Hidden Markov Model
IPT	Fraunhofer Institute for Production Technology
KNN	K-Nearest Neighbour
LR	Logistic Regression
ML	Machine Learning
PER	Perceptron
PLC	Programmable Logic Computer
RMS	Root Mean Square
Skew	Skewness
SVM	Support Vector Machine
TCM	Tool Condition Monitoring
VB	Flank Wear

LIST OF SYMBOLS

$cov(X, Y)$	Covariance
$E[X]$	Expected value of X
$E[Y]$	Expected value of Y
x_i	X variables samples
y_i	Y variables samples
$\overline{x}$	Mean of values in X variable
$\overline{y}$	Mean of values in Y variable
r_{xy}	Pearson Correlation
ρ	Spearman Correlation
$cov(R(x), R(y))$	Covariance of the ranks
$\sigma_{R(x)}$	Standard deviation of rank of X variable
$\sigma_{R(y)}$	Standard deviation of rank of Y variable
$R(x)$	Rank of X variables
$\overline{R(x)}$	Mean rank of X variables
$R(y)$	Rank of Y variables
$\overline{R(y)}$	Mean rank of Y variables
X_i	Sensor Signal
σ	Standard deviation
$S(f)_i$	Power at a specific frequency

CONTENTS

1 INTRODUCTION

1.1 MOTIVATION

The metal-cutting industry, besides being one of the oldest manufacturing processes, is also one of the most important and diversified manufacturing processes in our society's history. Metal cutting makes its presence in almost all products present in the industrial market nowadays. Inside the metal cutting industry, due to its flexibility of manufacturing multiple types of goods with the same equipment and power to make such complex work-pieces profiles, milling has become one of the most relevant machining processes (ZHOU; XUE, 2018).

The demand for customized and flexible production and, at the same time, high precision has led the machine industry to new challenges (CAO; ZHANG, X.; CHEN, X., 2017). Metal cutting, being one of the primary manufacturing processes, is required to achieve very high dimensional accuracy and almost the desired surface finish (SINGH, 1996). Even though metal cutting has existed for decades, it still seeks methods to decrease production costs and downtime of the machine.

Until the mid-1950s, the metal-cutting process relied on highly skilled labor. Since then, automated machining began to replace human operators transforming the metal-cutting process into something more efficient and less costly for industries. Since then, industries have shown great interest in automated machining due to the capability of these systems to increase product quality with a reduction of production costs (O'DONNELL et al., 2001).

More demands for customized and flexible production are being made for the metal-cutting industry. In order to attend to those demands, the industry is moving gradually towards the complete integration of the shop floor (CAO; ZHANG, X.; CHEN, X., 2017). Since the introduction of the Internet of Things, intelligent components have been added to the shop floor in order to make the machining processes more automatic and efficient.

Allied with these intelligent components, the Big-Data age has become one of the handy tools to provide new solutions for the machining processes. With the integration of the shop floor and hence the integration of the manufacturing processes with intelligent components, acquired data can now be stored and furthermore be used as elements to produce new and powerful tools to monitor and optimize operations. The combination of Big-Data with Machine Learning (ML) became a powerful solution and is causing a revolution in the metal industry (YAN et al., 2017).

To avoid any interruptions in any machining process, cutting tool wear is a critical issue that has to be monitored. An effective monitoring system should detect the tool wear level and predict failures beforehand. This way, corrective actions can be done to maintain the process parameters or prompt effective tool change strategies (LEE, J.; KIM, D.; LEE, S., 1998). If the mentioned monitoring system fails to detect the abnormalities mentioned above, it could result in poor surface quality, defects in the workpiece and even machine breakage or defects (LI; LAU; ZHANG, Y., 1992).

A new and active research field in the machining processes to control the surface finish, integrity of the workpiece and even the tool's vibration was opened with the monitoring of tool wear. Additionally, with the development of the mentioned monitoring system, industries could reduce machine downtime, resulting in higher efficiency of time and resources.

Even though there are already studies in the monitoring system mentioned above, work is still required to transport these technologies to the industrial environment. Aligned with that, there is still room for improvements and new ideas for these technologies, being one of these improvements the main focus of this work.

1.2 INSTITUTE FRAUNHOFER IPT

The main goal of Fraunhofer Institut für Produktionstechnologie (Fraunhofer Institute for Production Technology (IPT)) is to align the upcoming equipment and system produced in the market with new technologies like ML and Big-Data. The Institute is well known for providing technology systems for production worldwide.

Figure 1 – Fraunhofer Institut für Produktionstechnologie.

Source – Fraunhofer Institute (2021).

The Fraunhofer Institute makes its presence all around the world with more than 80 working units, most of them located in Germany. The Institute has a great partnership with big companies in diverse fields such as machining, automobilists and the aero-space industry.

The task of transporting technology from the academic environment directly into industrial practice is held by the IPT Institute, as itself defines (IPT, 2021). The main focus is to be able to provide reliable tools for tasks required by the customers. In other words, the Fraunhofer IPT Institute makes use of high-technology equipment aligned with cutting-edge state-of-the-art research in order to develop innovative systems for the industrial environment.

The Precision Technology and Automation Department inside IPT, being the department where this project was developed, maintain the goal of providing solutions for high-precision machining. As the precision in its name suggests, small tolerance and better control of the positioning system of the machine are required to keep the accuracy to its smaller levels. Errors like chattering, thermal deflection, positioning offset and tool wear are more harmful to the milling process than conventional milling operations.

In order to achieve this higher precision, it requires the acquisition of information from sensors and the Programmable Logic Computer (PLC) machine system at a faster speed than normal machining processes. Therefore, also in this department, work regarding high-frequency data acquisition is carried out.

1.3 PROBLEM DESCRIPTION

The present work aims to study and develop a system using ML and statistical techniques to improve ML models for tool flank wear prediction.

The definition of flank wear is the loss of material on the cutting tool's edge, which is caused by the tool's frequent contact with the workpiece's surface. There are many types known of tool wear and, as stated by Palanisamy, Rajendran, and Shanmugasundaram (2008), the reasons behind tool wear can be imperfections in the tool composition, physical shock between tool and workpiece, abfraction on the cutting edge of the tool and heat provoked by the friction of the tool and workpiece. Usually, the tool wear increases gradually with the use of the tool, by other means, the more time you use the tool, the more it is going to wear.

Because of the small tolerance required in higher precision milling processes, in the order of micrometers, the effects caused by tool wear are largely intensified. With the tool's shape and sharpening affected by the tool wear, the impact of small errors is carried out to the workpiece. Such deterioration of the tool can then provoke undesired vibrations in the tool and, consequently, the poor surface quality of the workpiece, resulting in rework or even the rejection of the final manufactured product.

Besides the problems caused in the workpiece by the tool wear, the tool cutting power decreases with the wear, increasing the load on the machine necessary to produce the same results, reducing then the lifetime of the machine and even provoking the failure of the equipment.

The damages caused by tool wear can be avoided by monitoring the tool wear phenomena, making it possible to determine the precise moment for the tool change. Despite that, an accurate measurement of the tool wear is made directly in the tool and, in the industrial environment, it turns out to be a very costly procedure and not commonly used due to the fact of the interruption of the process, removing the tool from the milling machine and usage of special equipment to check its status.

Nowadays, the experience of the machine operator is commonly used to determine the time to change the tool. This method has low accuracy and can produce undesired results. On one side, if the machine operator changes the tool too early, causing the underuse of the tool, it will create extra monetary and time costs. On the other side, if the machine operator takes too long to change the tool, causing the overuse of the tool, it can damage the workpiece and machine, as mentioned above (ZHOU; XUE, 2018).

As stated in Zhou and Xue (2018) and Rehorn, Jiang, and Orban (2005), 6-20% of the downtime in milling is caused by tool wear and tool breakage. Some authors even ranked those factors as the most significant barrier in the actual scenery of the manufacturing industry. In Cao, Xingwu Zhang, and Xuefeng Chen (2017), when the author approaches the critical technologies required for smart spindles, as he stated, he mentioned the importance of TCM, highlighting it as one of the essential factors in order to achieve high accuracy and efficiency in milling, once it is directly connected with the performance of the machine.

According to Rehorn, Jiang, and Orban (2005), a good TCM system can shorten the downtime and cutting costs in the order of 10-40% by increasing the lifetime of the cutting tool on 10-50%. As the direct measurement is out of hand in the machining industry, the indirect measurement, determining the tool's wear by outside sensors, is recommended for a TCM system.

Having such importance and impact in the milling industry, finding ways to improve already existing TCM with reduced costs of implementing the system and increasing prediction accuracy is highly encouraged.

1.4 OBJECTIVE

In order to build the improvements in the TCM system, many different ML techniques and Feature Selection methods will be studied and employed in this project. However, it is mandatory that data will be needed in order to train and test our models to confirm that the improvements are being made. Therefore, a set of experiments will be conducted to gather the data necessary to implement the system. The resulting data of the experiments will hold information of multiple synchronous sensors.

The goal of the proposed work is to implement prediction models using different ML architectures and different Feature Selection methods to test their combinations. Finally, the performance of each combination will be analyzed and the most suitable combination can be identified. Furthermore, the presented results will further explain the tool wear phenomena and signal behavior during the milling process.

As a result, this project will present the following elements:

- A sensor setup installation for the studying of the tool wear phenomena during the milling operation;
- A study in the different Feature Selection methods and their pros and cons;
- A dataset containing all the information of sensors gathered during the experiments performed on the milling machine;
- A compiled of information regarding which sensor and feature is relevant to predict tool wear;
- And finally, but not least importantly, a system capable of predicting and evaluating the combinations of Feature Selection methods with ML algorithms;

1.5 MONOGRAPHY STRUCTURE

This monography will be structured as follows.

In the Chapter 2 will be presented the literature review of the relevant knowledge needed for this work, including a brief explanation of the milling process, ML, tool wear, tool condition monitoring and feature selection.

In Chapter 3, the experimental setup will be presented, giving the essential details about the hardware and software used to make and collect the data used in the ML models. Also will be presented a brief component description of the sensors used to collect the data.

In Chapter 4, the treatment of the data will be presented, from its acquisition until after its pre-processing, where the data will be ready to be used in the feature selection methods and, afterward, ML models.

In Chapter 5, the Feature Selection and ML tuning will be explained by describing the parameters changed in the tuning and why they were changed. Also, this Chapter will present the results obtained after the feature selection and ML models were applied.

Lastly, in Chapter 6, this document will be concluded, presenting a summary of the monography and suggestions for future works.

2 LITERATURE REVIEW

Several studies were proposed in the past involving various methodologies and technologies for the development of a Tool Condition Monitoring system. In this chapter, a brief review of all the subjects that this work will contain, including the milling process, tool wear mechanism, tool condition monitoring, machine learning and feature selection, will be covered.

2.1 MACHINING PROCESSES

Studying machine processes is like studying human history itself. Since the beginning of civilizations, humans produced machine tools, mostly from stone and wood, to help them in manufacturing. After long years of evolution, nowadays our tools, and machines, are produced by metal thus, increasing their durability and complexity of them.

Machining processes are the techniques used in the industry for removal, in the form of chips using a cutting tool with single or multiple wedge-shaped, or the addition of material from a workpiece. Machining processes can be grouped into distinguished categories of processes by means of how the process removes or adds material from the workpiece. Some of those categories are Abrasive processes, Cutting processes and Additive manufacturing processes.

This study is focused on cutting processes, more specifically on the milling process which involves the machining of an external surface of a workpiece with a rotating cutting tool that can produce flat or curved surfaces and prismatic shapes.

In past studies, the author Diei and D. A. Dornfeld (1987) indicates that the cutting phenomenon which is involved in the milling process is very complex. Some of the distinctions are mentioned below.

- Milling process in metal cutting is a discontinuous operation, which introduces some complications in the processing of the data acquired.
- The cutting forces and operating conditions at the tool flank edge surface may change throughout the cutting cycle, thereby causing variation in the characteristics of the process.
- There can be random variations in the characteristics of the signal due to the carrying of chip build-up at the tool exit of the previous cut onto the next cut.

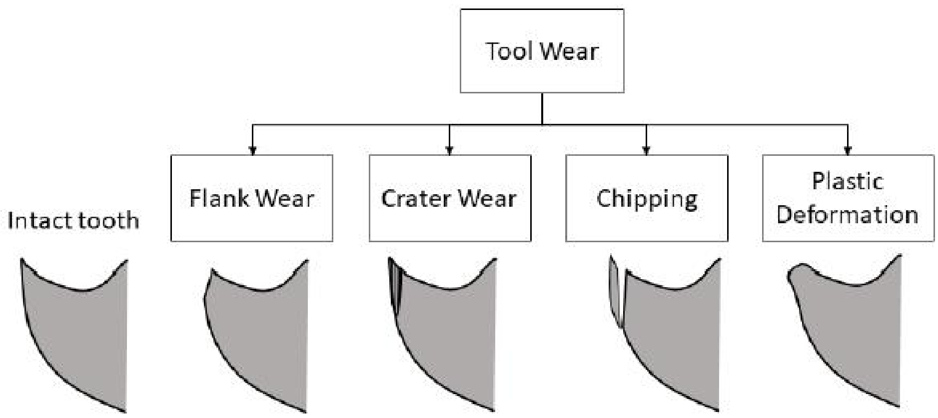
2.2 TOOL WEAR

Determining the state of a cutting tool in the milling process is a factor of great importance to define the efficiency of the machining process and represents a direct correlation with the cost of the manufacturing process as the tool failure will provoke unscheduled machine downtime (DIMLA, D. E., 2000).

The wear of a tool is the critical indicator of the cutting tool's life, as with the increase of the wear, the tool leads to failure which is caused by a combination of thermos-mechanical factors. Kurada and Bradley (1997) has stated that flank wear, crater wear, chipping built-up edge and breakage are the main modes of tool wear, which are identified by their geometry and locations in the tool.

When approaching the tool wear, Dolinšek and Kopač (2006) described types of tool wear during the lifetime of a cutting tool as seen in Figure 2. The author also pointed out that flank wear and central wear (in spherical tools) are the most influencing types of wear in the progress of a tool deterioration in High-Speed Milling.

Figure 2 – Types of tool wear in milling tools.



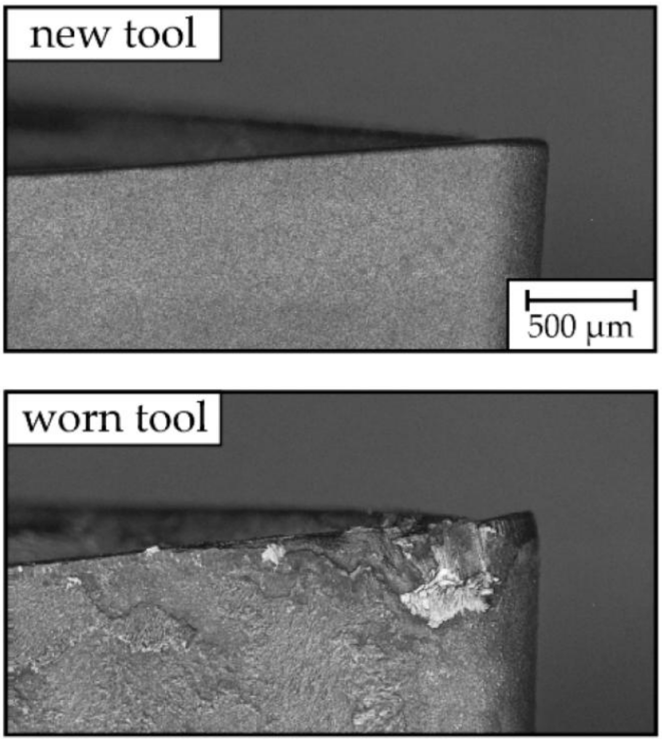
Source – Dolinšek and Kopač (2006)

The author of Palanisamy, Rajendran, and Shanmugasundaram (2008) stated some effects of the tool wear in the milling process. According to the author, flank wear is the prevalent wear type observed in cutting tools. It is defined as the loss of material in the relief face of the tool, which is caused mainly by the contact of the cutting face of the tool with the workpiece material.

The author Dolinšek and Kopač (2006) also described the tool wear progress characteristics during the tool's lifetime. The document explained that in the beginning, the tool develops first wear on the cutting edge by deformation on the cutter edge material, which will be generated in the first 30 minutes of operation. After this first deformation, the wear stabilizes, improving the surface quality and developing the wear slower than before.

The flank wear continues to develop until the coating that protects the tool is worn out. As consequence, the wear of the tool increases rapidly at this point with heavy material losses on the tool edge and a great increase in the cutter load. This effect provokes increasing friction between the tool and workpiece and also the temperature of the tool. The result of this process is a fast development of tool wear and, by means of the high temperature, other types of wear start to grow quicker like crater wear and chipping. The degradation of the tool grows at a fast pace until it finally needs replacement or, in some cases, even the breakage of the tool due to the high load on the cutter. Those effects cause the loss of accuracy and surface quality of the finished product. The differences between a new tool and a worn one can be seen in Figure 3.

Figure 3 – Comparison between new and worn out tools.



Source – Denkena, Bergmann, and Witt (2020)

Scheffer and Heyns (2004) stated that a tool could last from a few hundred to a few thousand of components until it gets worn out, meaning that the tool wear phenomenon is unpredictable based upon the observations from the experiences of the operators of machining. These statements suggest a need for a tool wear monitoring system to avoid the possibility of damaging the machine or any component of it and to maximize the utilization of the tool, reducing the cost of the operation.

The automation of a tool monitoring system allows relocation of capital and human operators to perform higher level tasks than to monitor the machine tools (BURKE, 1992). Therefore, an intelligent TCM to detect the tool wear avoiding machine down time is required.

2.3 TOOL CONDITION MONITORING

Since the late 1980s, TCM has been studied by researchers to detect tool wear and tool failure in advance of the failure event by the usage of different sensors as seen in Rangwala and D. Dornfeld (1990), Byrne et al. (1995) and Wilcox, Reuben, and Souquet (1997). The study of TCM system has earned high attention in the metal cutting field, due to the increasing need for properties like the surface quality of the workpiece and process efficiency, which are strongly related to tool condition and, by means, strictly related to tool wear progress. As pointed out before in this paper, determining the right time for tool replacement has great importance in order to minimize production costs and machine downtime, meaning an increase of the properties before mentioned.

According to the literature, three stages such as signal selection and acquisition, signal processing and features selection, and decision making, are required to develop a generic methodology for TCM system for machining, therefore, a TCM system is a combination of information flow and processing system. A TCM system is also a combination of hardware and software. Signal acquisition is present in the process representing the hardware part of the system, the remaining part of it, including signal analysis, interpretation, and tool state prediction, compound the software part of the system.

The direct measurement of the flank wear in tools, using a microscope or laser beams to measure it, has a high accuracy, due to the used tool being measured directly. On the other hand, direct measuring is usually expensive and time inefficient, due to the fact that this type of measurement not being possible to be executed when the machine is running, in other words, increasing the machining downtime. Direct measuring is also confined to laboratory techniques and the process is complex enough to be used in the industrial environment. Researchers moved then towards indirect measurement techniques. However, the use of indirect measurement brings up new difficulties to the system, like the problem of dealing with noisy acquired data.

The author of the review (REHORN; JIANG; ORBAN, 2005) approached the main methods for indirect measurement of tool wear. The first one is the use of cutting force as source signal, usually done by the use of table dynamometers and it has been found to be the most reliable approach to monitor the tool condition (CHEN, J. et al., 1999). The author raises up the problem of the inconstancy of the signal related to the position of the spindle, due to the signal having peaks and lows in the exit and entry of each tool tooth on the workpiece material. When using conventional methods for tool condition estimation, like threshold criteria, this behavior of inconstancy of the signal can easily lead to misunderstanding decisions.

The author on Cao, Xingwu Zhang, and Xuefeng Chen (2017) raised another question when approaching TCM system using threshold criteria. The document stated that the threshold criteria is highly dependent on the CNC process parameters. In other words, in order to have a robust TCM system, the threshold must have a high variety of parameters and complex programmed rules, otherwise, the accuracy is heavily damaged.

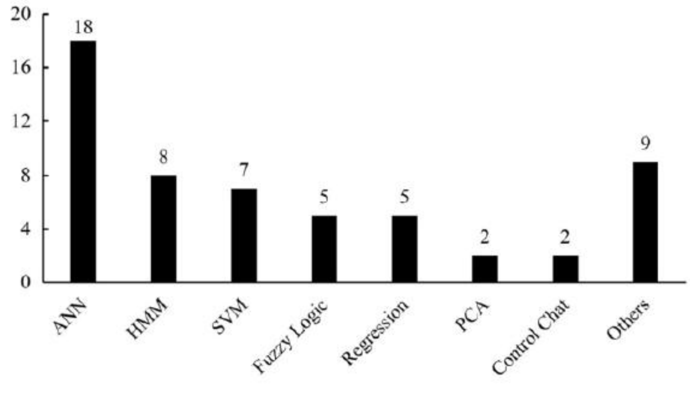
The usage of Acoustic Emission (AE) sensing, which is the measurement of different characteristics of acoustic (elastic) waves in solids, is another method reported by the author of Rehorn, Jiang, and Orban (2005). By the author, as the flank wear develops, the energy emitted by the AE signal increases, making it a handy tool for the detection of tool wear and even breakage. The same document also explained that the main frequency of the process, in the AE signal, gets higher at the end of the tool life. Mechanisms involving Root Mean Square (RMS), Skewness (Skew), and FFT are usually applied to evaluate the output presented by the AE signals.

Last but not least, the author of Rehorn, Jiang, and Orban (2005) indicates the usage of vibration sensing as a source of information for indirect measurement. This is usually done by the installation of accelerometers in the machine. Thus, the review of the behavior in some different bands of frequency of the signal can identify features of the measurement. This method of indirect measurement is mainly used in monitoring surface roughness (CHEN, J. et al., 1999; JANG et al., 1996). However, some papers showed some interesting results with the usage of this sensing method (DEY; STORI, 2005; DIMLA, D., 2002; ERTEKIN; KWON; TSENG, 2003). The main disadvantage of this method is that the sensor position and machine speed range have a great influence on the measurement (BAHR; MOTAVALLI; ARFI, 1997).

Different methodologies of indirect measurement were presented by the authors of Al-Sulaiman, Abdul Baseer, and Sheikh (2005), Sadílek et al. (2014) and Čuš-Uroš and Zuperl (2011). However, the major point to observe is that analyzing the outputs of the sensor for indirect measuring mostly requires the development of complex methods for the interpretation of the information. Thus, there is a high prevalence of statistical and smart system development for decision-making when coming to tool wear in milling processes.

The author of Cao, Xingwu Zhang, and Xuefeng Chen (2017) points out the strong trend in using ML to overcome this barrier. The paper also reported an increasing usage of ML techniques such as Artificial Neural Network (ANN), SVM, and Hidden Markov Model (HMM) to solve this issue. In the review presented by Zhou and Xue (2018), the author confirms this trend by listing the recent works regarding TCM for milling processes. This confirmation can be seen in the Figure 4.

Figure 4 – Recent TCM methods reported for milling process.



Source – Zhou and Xue (2018)

2.4 MACHINE LEARNING

Artificial Intelligence (AI) has received different definitions throughout history. Russell (2010) classified it into four different categories: Systems that act rationally, Systems that think rationally, Systems that think like humans and Systems that act like humans. Between all definitions, one of the most implied is "The branch of computer science that is concerned with the automation of intelligent behavior" or, in other words, AI is the field of study that purpose is to program computers to perform actions which demand human intelligence. ML, on the other hand, as described by Arthur Samuel (1988), is the "Field of study that gives the computer the ability to learn without being explicitly programmed".

AI techniques have been used by researchers, (DORNFELD, David A; DEVRIES, 1990) for the past few decades as a decision-making strategy of an TCM. In this context, several AI techniques have been used to diagnose tool wear in the past. The main AI techniques used for modeling and monitoring machining systems are ANN, fuzzy logic systems and neuro-fuzzy inference, which combines these two techniques. ML approaches, which is a subset of AI techniques, such as evolutionary algorithms or SVM, have been less widely used. However, they are gaining popularity in recent works.

The application in TCM of ML has been improved significantly in the past few years. In Kannatey-Asibu, Yum, and T.H. Kim (2017) a real-time TCM system was proposed to perform, through acoustic emission monitoring, wear classification of the cutting tool in a coroning process. It was proposed based on the concept of the classifier fusion method. This method shows significant results in classification rate with unity weighting.

Jinjiang Wang et al. (2017) published a study where a TCM was developed with an artificial intelligent model whose data were from a virtual tool wear sensing technique based on multi-sensory data fusion. This method involved the application of a ML technique called support vector regression.

Accordingly to Al-Zubaidi, Ghani, and Che Haron (2011), ML is divided into mainly two categories, supervised and unsupervised learning, having also other categories such as reinforcement learning (KAELBLING; LITTMAN; MOORE, 1996).

- Supervised learning: This approach is when the reference data is available to the algorithm. By other means, the computer contains a dataset with the inputs and target outputs. The learning process happens by using the dataset in the model multiple times. To determine the model's accuracy, the number of samples mapped correctly from input to output is used. The main examples of this approach are classification and regression.

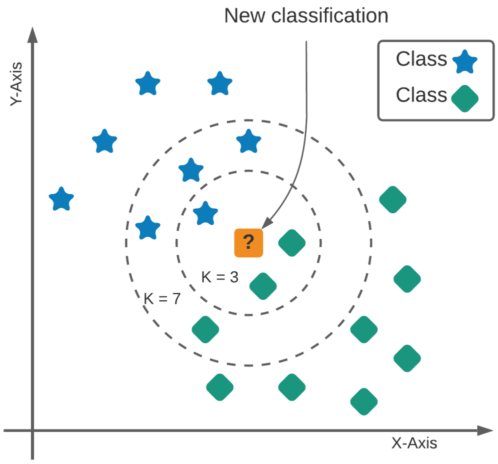
- Unsupervised learning: In this approach, the reference data isn't available to the algorithm. Therefore, the main task in this category is to observe and find patterns between different data inputs. It is possible to discover structures that are common to each group on the dataset when using this mechanism of ML. The main example of this approach is clustering.

2.4.1 K-Nearest Neighbour

As one of the most fundamental and simple classifications ML methods, K-Nearest Neighbour is the first choice for classification study when there is no or little prior knowledge about the data distribution. The KNN classification was developed from the need to perform discriminant analysis when reliable parametric estimates of probability are unknown or difficult to determine (PETERSON, 2009).

The KNN classifier is a type of classification commonly based on the Euclidean distance between the test sample and the specified training samples (PETERSON, 2009). The operation of the KNN is to use the distance function to evaluate the data point and select the class based on the "K" nearest distances of data. In other words, a class label is assigned based on a majority vote.

Figure 5 – KNN operation graphic.



Source – Author's Image

The Figure 5 explain how the KNN algorithm works and an example for $K = 3$ and $K = 7$. It is possible to see then how the data will be classified accordingly to the chosen number of neighbors meaning that the K parameter is vital for the tuning of the KNN ML model.

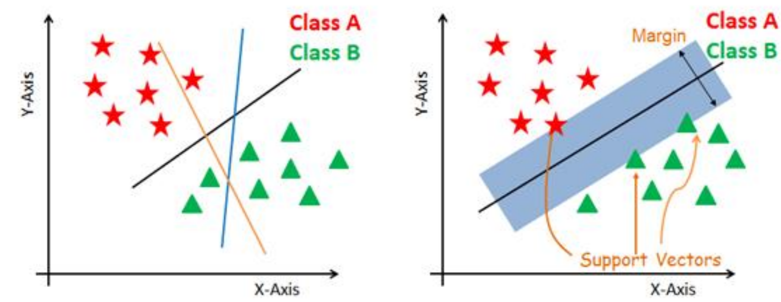
As for the distance function, the Euclidean distance was used. The Equation (1) shows the Euclidean distance function where $d(x, y)$ is the distance value and n is the number of samples.

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} \tag{1}$$

2.4.2 Support Vector Machine

A relatively new computational learning method, SVM is based on statistical learning theory and structural risk minimization principle which is in contrast to a classical learning approach such as ANN which is based on empirical risk minimization; an approach designed to minimize error on the training dataset. The SVM approach has been noticed to be efficient in large classification problems due to its capacity of handling very large feature spaces. The principal purpose of SVM technique is to explore an optimal hyperplane for maximizing the margin of separation of two classes (VAPNIK, V. N., 1999; BURGES, 1998).

Figure 6 – Support Vector Machine overview.



Source – Author's Image

The SVM algorithm constructs a hyperplane that will maximize the margin between the input classes. In the Figure 6 is possible to observe how the SVM algorithm works. There is an infinity of different hyperplanes that separate two classes of data and the SVM algorithm works to find the best hyperplane possible to separate it. The best possible hyperplane is the one with the maximum margin, which is the distance between the hyperplane itself and the support vector data points.

With the application of kernel functions, it becomes possible to use SVM algorithm in non-linear classification tasks. In this approach, the non-linear data is mapped into a higher dimensional space to make it linearly separable.

Conventional pattern recognition methods have shown that it is needed a higher number of sample data in order to develop these algorithms. On the other side, SVM algorithm can provide better generalization than ANN and other methods with a small number of data samples (VAPNIK, V., 1999). Due to the hardness of obtaining sufficient data in the industry, the SVM approach introduced an advantage in obtaining good results and therefore has become more used in similar problems.

2.4.3 Logistic Regression

Logistic Regression (LR) is a statistical method of analysis most commonly used to predict binary output problems, such as "yes" or "no", "1" or "0" or, in this case, a new tool or worn tool. Is also possible to use LR to predict multinomial classes, which means several possible outcomes as long as the number of the outcome is finite, however, it is not the normal use case.

The logistic model, in statistics, models the probability of an event taking place by having the logarithm odds of it be a linear combination of one or more predictors. The logistic model alongside with the probit model, are the most commonly used models for binary regression as said in Cramer (2002).

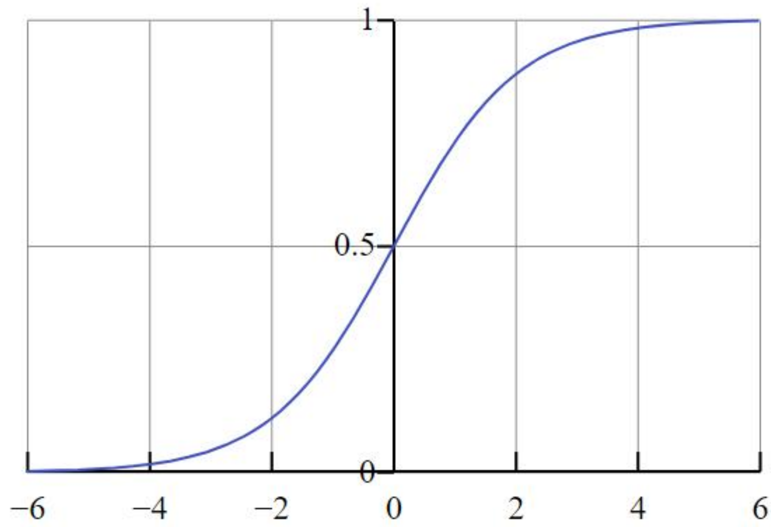
LR is an algorithm used in classification to predict the probability of an item belonging to a class, despite what its name suggests. As stated in Subasi (2020), LR is a simple and more efficient method for binary classification problems and it is a model which is easy to realize and achieves very good performance.

The LR uses a sigmoid function to map predictions and their probabilities,

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

the sigmoid function shown in the Equation (2) refers to an S-shaped curve that converts any value to a range between 0 and 1. The behavior of the described equation can be seen in Figure 7

Figure 7 – Logistic Regression curve.



Source – Chakraborty et al. (2019)

The Equation (3) represents the LR, where x is the input value, y is the predicted value, b_0 is the bias and b_1 is the coefficient for input x.

$$y = \frac{e^{(b_0 + b_1 X)}}{1 + e^{(b_0 + b_1 X)}} \tag{3}$$

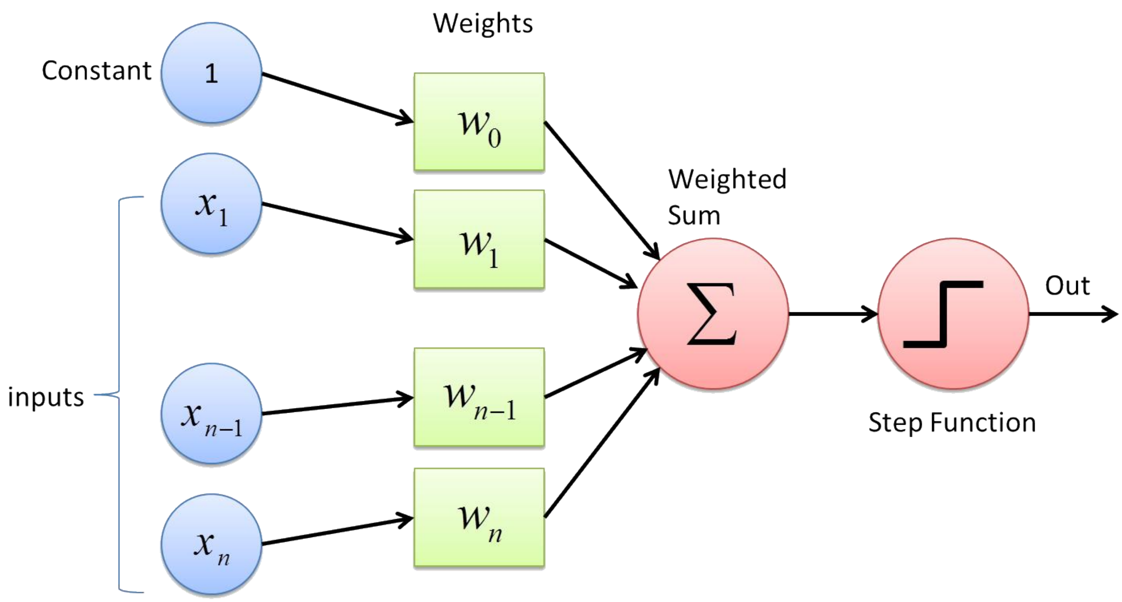
Due to the shortage of data, LR is most valuable in the medical and social sciences field, where it interprets results from experiments. LR is a simple and fast model, therefore it is also used in very large data sets.

2.4.4 Perceptron

The Perceptron (PER) technique was developed by Frank Rosenblatt in the period between 1950 and 1960. The technique is used in supervised learning and can be used to classify data from known inputs. The PER is used as a binary classifier, which is a function that can decide whether or not an input belongs to some specific class, as said in Freund and Schapire (1998). PER can be considered the simplest and the first type of ANN. A way to think about PER is that it is a device that takes decisions based on evidence.

PER is a mathematical model which receives several entries and produces only one binary output. It is formed by a single neuron that takes several data inputs and predicts a class label. The output is achieved by calculating the weighted sum of the inputs and bias.

Figure 8 – Perceptron scheme.



Source – DeepAI (2022)

The Figure 8 represents an overview of how the PER works. The step function, as seen in Figure 8, is one of the activation functions used in the PER model. The sigmoid function explained in the section before, can also be used as one of the activation functions for PER model.

The operation of the PER model starts with samples of the training dataset being shown to the model one at a time, the model then makes a prediction and the error is calculated. Before that, the weights of the model are updated in order to reduce the errors. This process is called the Perceptron updating rule.

The updating rule is repeated for all examples in the training dataset, called epoch. Then, the process of updating the model using examples is repeated for many epochs. The training stops whenever the error made by the model achieves a very low level or no longer improves the model, or a maximum number of epochs is performed.

2.5 FEATURE SELECTION

Feature selection, also known as variable selection, attribute selection, or variable subset selection, in ML and statistics is the process of selecting a subset of believed relevant features to be most useful to predict the desired target variable in model construction. In other words, as stated in Kuhn, Johnson, et al. (2013), feature selection is primarily focused on removing non-informative or redundant predictors from the model.

Among the numerous uses for feature selection, we can point out some of them:

- Pointed in James et al. (2013), simplification of models, making them easier to interpret by users and researchers.
- Shorter training time, as quoted by Liu (2010).
- Pointed by Kramer (1991), to avoid the curse of dimensionality, which is that when the dimensionality increases, the volume of the space increases so fast that the available data becomes sparse.
- To improve data compatibility with a learning model class, as seen in Kratsios and Hyndman (2021).

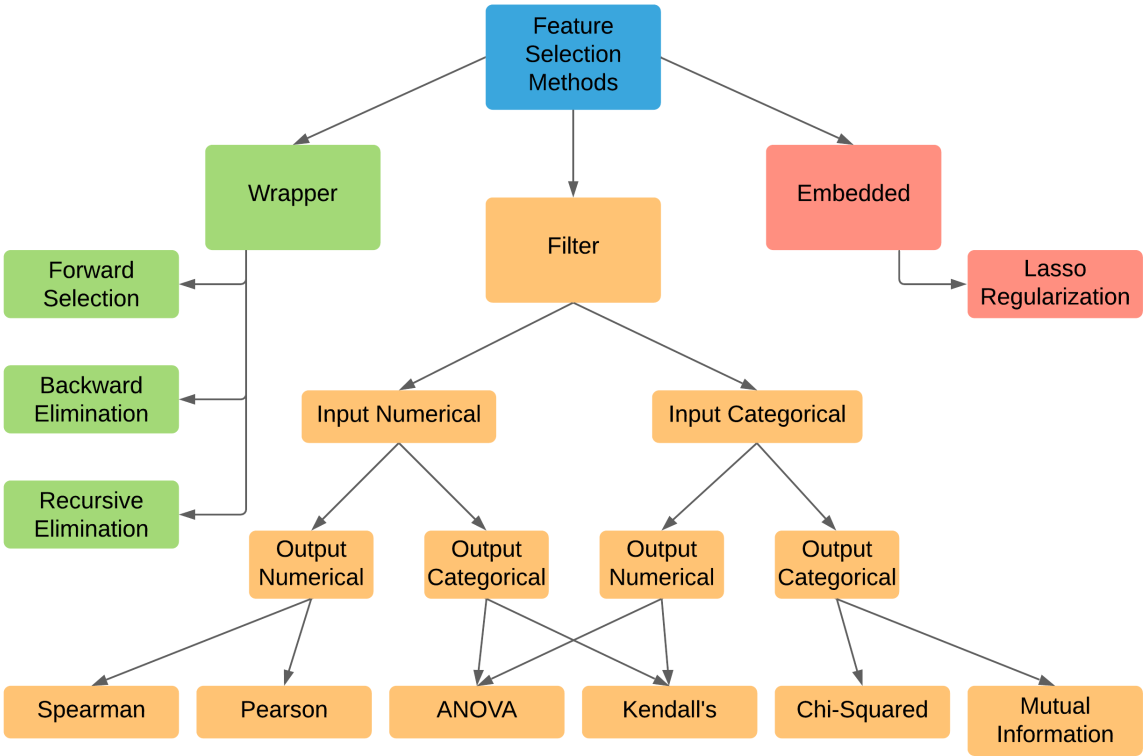
A high quantity of prediction modeling problems has a large number of variables, which can affect the training of the model by slowing the development and the training by itself and also increasing the amount of system memory required to process the ML model training. Additionally to it, including input features that are not relevant to achieve the target variable can affect the overall performance of the ML model. As cited in Kuhn, Johnson, et al. (2013) models based on regression estimate parameters for every term in the model. Therefore, having non-relevant variables for the model can add uncertainty to the prediction and then reduce the overall effectiveness of it.

The premise behind feature selection methods is that the input data contains features that are either redundant or irrelevant and, therefore, can be removed from the input without influencing much on the loss of relevant information for the ML model, as seen in Kratsios and Hyndman (2021). By Guyon and Elisseeff (2003), redundant and irrelevant are distinct notions since a relevant feature can be redundant in front of another relevant feature, being both of them strongly correlated with each other.

As seen in Kuhn, Johnson, et al. (2013), feature selection methods are divided into two distinct classes, supervised and unsupervised methods. The distinction between these classes has to do with whether features are selected based on the target variable or not. Supervised methods use the target variable to remove irrelevant variables from the input data. Unsupervised methods ignore the target variable to remove redundant variables from the input data.

Another way to divide feature selection methods is into wrapper, filter and embedded methods. These methods are almost always supervised and are evaluated based on the performance of a resulting model on a hold-out dataset. This division, with some examples of each class, can be seen in Figure 9. Wrapper and filter methods will be covered in the next sections.

Figure 9 – Feature selection division.



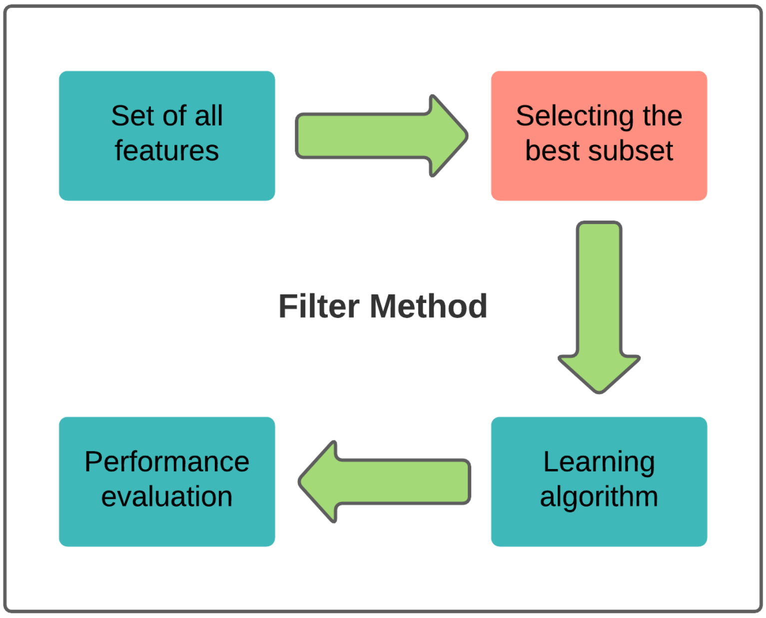
Source – Author's image

Finally, some ML algorithms perform feature selection automatically as part of the learning model. Those are recognized as embedded feature selection methods. In Kuhn, Johnson, et al. (2013), it was said that Tree-based and rule-based models, MARS, Lasso and random forest, for example, have built-in feature selection, meaning that they will only include predictors that help to maximize the accuracy.

2.5.1 Filter Methods

Generally used as a pre-processing step, the selection of a feature in filter methods is independent of any ML algorithms. Instead, features are selected based on their score in statistical techniques for correlation with the target variable. In other words, it evaluates the relationship between each input variable with the desired outcome variable, which will then be used as an input dataset to the model to be trained. The Figure 10 summarizes how filter feature selection methods work.

Figure 10 – Filter feature selection summary.



Source – Author's image

Filter feature selection methods suppress the least interesting variables. The advantages of using filter methods are, according to Yu and Liu (2003):

- This method is effective in computation time and power. In other words, it is faster than wrapper and embedded methods and uses less computational power to perform the feature selection.
- It is robust to over-fitting. Since it doesn't require a ML model to evaluate the features.
- Scales better to high-dimensional datasets than wrapper methods.

The standard procedure of filter feature selection methods is to evaluate the relationship between input variables and the target variables, meaning that it doesn't remove redundant features. In other words, filter feature selection methods don't deal with multicollinearity, it should be kept in mind when using filter feature selection methods.

According to studies on many filter feature selection methods and the type of data contained in the project, Pearson Correlation and Spearman Correlation methods were selected.

2.5.1.1 Pearson Correlation

Pearson Correlation, as its name indicates, is a measurement between two sets of data of its linear correlation. In other words, Pearson Correlation measures the strength and direction of linear correlation between two variables. It is calculated by the ratio between the covariance of two variables and the product of their standard deviations (WEISSTEIN, 2022).

The definition of covariance is the expected value, or mean, of the product of their deviations from their individual expected values (PARK, K. I.; PARK, M., 2018):

$$cov(X, Y) = E[(X - E[X])(Y - E[Y])] \tag{4}$$

where $E[X]$ is the expected value of X and $E[Y]$ is the expected value of Y, also known as the mean of X and Y.

The Equation (5) shows how Pearson Correlation is calculated where x_i is the X variable samples, y_i is the Y variable samples, $\overline{x}$ is the mean of values in X variable and finally, $\overline{y}$ is the mean of values in Y variable.

$$r_{xy} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum(x_i - \overline{x})^2 \sum(y_i - \overline{y})^2}} \tag{5}$$

As the result of the correlation between two variables can be values between -1 and 1, the Pearson Correlation is essentially a normalized measurement of the covariance. To interpret the result we must take into account two indicators, the signal and the value by itself. The positiveness of the result indicates the relation of the variables, being positive means that if one variable increases the other one tends to increase and vice versa. The value indicates the strength of the relation, being 0 as no correlation and 1 as perfect correlation.

2.5.1.2 Spearman Correlation

Spearman correlation, as just in Pearson correlation, measures the relationship between two sets of variables. Spearman correlation coefficient, which is normally denoted by the Greek letter ρ, is a non-parametric measure of the rank correlation between the variables using a monotonic function to describe the relationship. As Corder and Foreman (2011) stated, Spearman correlation is a statistical procedure to measure the relationship between two variables on an ordinal scale of measurement.

Rank correlation, such as Spearman and Kendall correlation, refers to methods that quantify the relationship between the rank of variables rather than the value of the variable. In other words, the data is ordered and then assigned an integer rank value for each data which will be used to calculate the correlation.

$$\rho = \frac{cov(R(x), R(y))}{\sigma_{R(x)}\sigma_{R(y)}} \tag{6}$$

The Equation (6) above mentioned denotes how Spearman correlation is calculated where $cov(R(x), R(y))$ is the covariance of the rank of X and Y variables, $\sigma_{R(x)}$ is the standard deviation of the rank of X variables and $\sigma_{R(y)}$ is the standard deviation of the rank of Y variables.

By Equation (6), the Spearman correlation is calculated just as the Pearson correlation between the rank of the variable instead of the actual value (MYERS; WELL; LORCH, 2013).

$$\rho = \frac{\sum((R(x) - \overline{R(x)})(R(y) - \overline{R(y)}))}{\sqrt{\sum(R(x) - \overline{R(x)})^2 \sum(R(y) - \overline{R(y)})^2}} \tag{7}$$

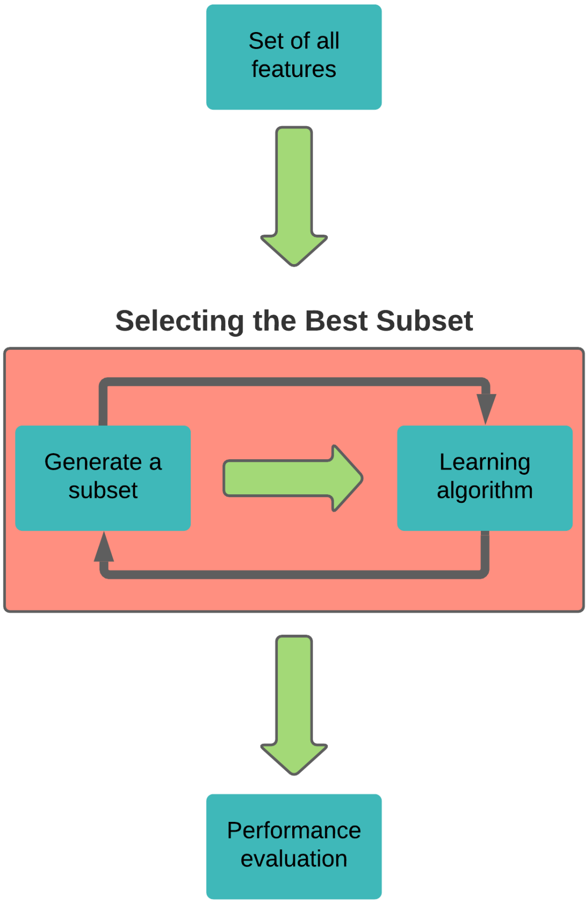
The Equation (7) indicates the complete Spearman correlation formula where $R(x)$ and $R(y)$ are the ranks of X and Y variables and $\overline{R(x)}$ and $\overline{R(y)}$ are the mean of the ranks of X and Y variables.

There are advantages when using Spearman correlation over Pearson correlation. One is that Spearman correlation is appropriate for both continuous and discrete variables, as described in Lehman (2005). The other one is that while the Pearson correlation assesses only linear relationships, the Spearman correlation assesses monotonic relationships, whether they are being linear or not.

2.5.2 Wrapper Methods

Unlike filter approaches, wrapper feature selection methods are dependent on a ML model to evaluate the features. By Kuhn, Johnson, et al. (2013) wrapper methods evaluate multiple models using procedures to add and/or remove features in order to find the optimal combination that maximizes model performance. In other words, wrapper methods train models using a subset of all features and based on the interference drawn from the previous model, it decides to add and/or remove features (PHUONG; LIN; ALTMAN, 2005). The Figure 11 shows how a wrapper feature selection method works.

Figure 11 – Wrapper feature selection summary.



Source – Author's image

The advantage of this method instead of the filter feature selection method is that wrapper feature selection methods deal with multicollinearity and find the optimal subset of features among the rules set for the wrapper method.

On the other hand, wrapper feature selection methods have disadvantages compared to filter methods. The main ones are:

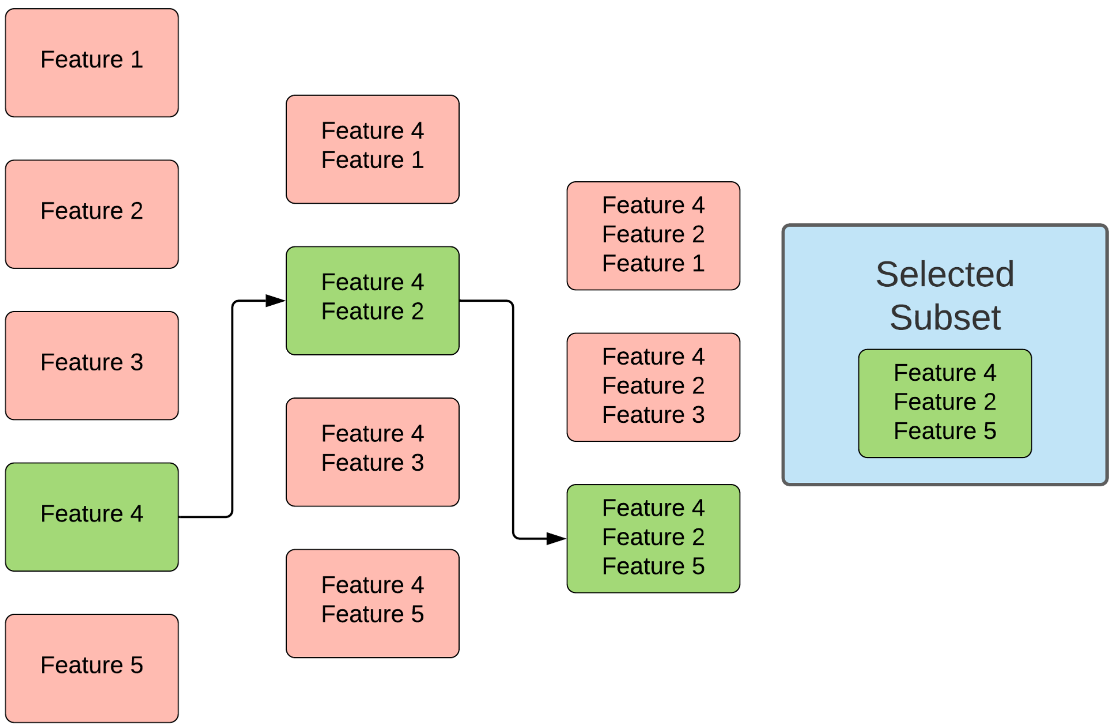
- Expensive computational time and cost when having a high amount of features.
- Increasing of overfitting risk when having an insufficient number of observations.

Forward in this work, it will be seen that Forward feature selection and Backward feature selection were selected as wrapper feature selection methods. Furthermore, these methods will be better explained.

2.5.2.1 Forward Feature Selection

Forward feature selection, unlike Pearson correlation and Spearman correlation, is a wrapper method of feature selection, which means that a ML model is used to select the features that are going to be used as input features for the main ML model. Forward feature selection is an iterative method to search for the best subset of features among all the features.

Figure 12 – Forward feature selection workflow.



Source – Author's image

In the Figure 12, the workflow of forward feature selection is presented. In Rejer and Lorenz (2013), the author explains how forward feature selection works. The process starts with an empty set of features selected. In each iteration of the process, a new feature is added to the collection of features selected until there is no improvement in the model used to select the features or the number of features chosen matches the predefined number of features to be determined by the process.

In other words, the method starts with training "N" times the model with only one feature, being "N" the number of all features. Then, all the "N" models are evaluated and the best feature is selected. The next step is to train "N-1" times models, but this time with the combination of the first selected feature and all the remaining features.

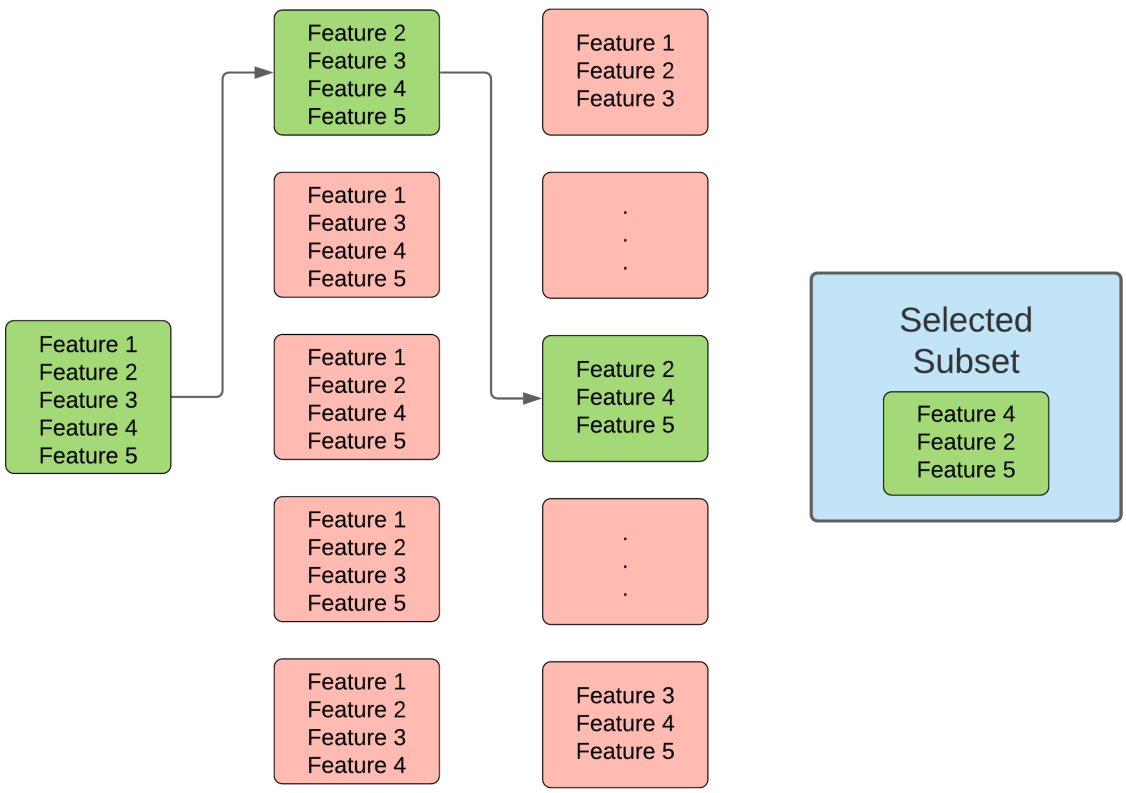
Therefore, the "N-1" models are evaluated and the best variety of the two features is selected. This process is repeated until one of the stop criteria mentioned before is matched.

Observe that although the process is reliable, it takes considerable time to run, depending on the number of features you have and which model is used as a classifier model. Also, the evaluation criteria is one of the parameters to be selected in order to produce an excellent forward feature selection. This work will use balanced accuracy as an evaluation metric for the forward feature selection method.

2.5.2.2 Backward Feature Elimination

Backward feature elimination, just as forward feature selection, is a wrapper feature selection method, which means that an internal ML model, called classifier model, is used to select the subset of features. Backward feature elimination also is an iterative method to select features. In a few words, backward feature elimination is the opposite of forward feature selection.

Figure 13 – Backward feature elimination workflow.



Source – Author's image

In the Figure 13, the workflow of backward feature elimination is presented. In Koller and Sahami (1996), the author explains how backward feature elimination works. The process starts with training the model classifier with all the "N" features and evaluating it. After that, a combination of "N-1" features is made, and the classifiers are trained for each set of combinations. All the combinations are evaluated and the feature that was dropped and produced the slightest change in the overall evaluation is then eliminated from the collection of features.

As forward feature selection, the process of backward feature elimination takes a considerable amount of time to run, depending on the number of features you have, the number of features that will be selected and the ML model used as classifier. In the same way as forward feature selection, backward feature elimination need an evaluation method to classify the models used in the selection of features and, in this work, balanced accuracy was used to evaluate the models from the classifier in backward feature elimination.

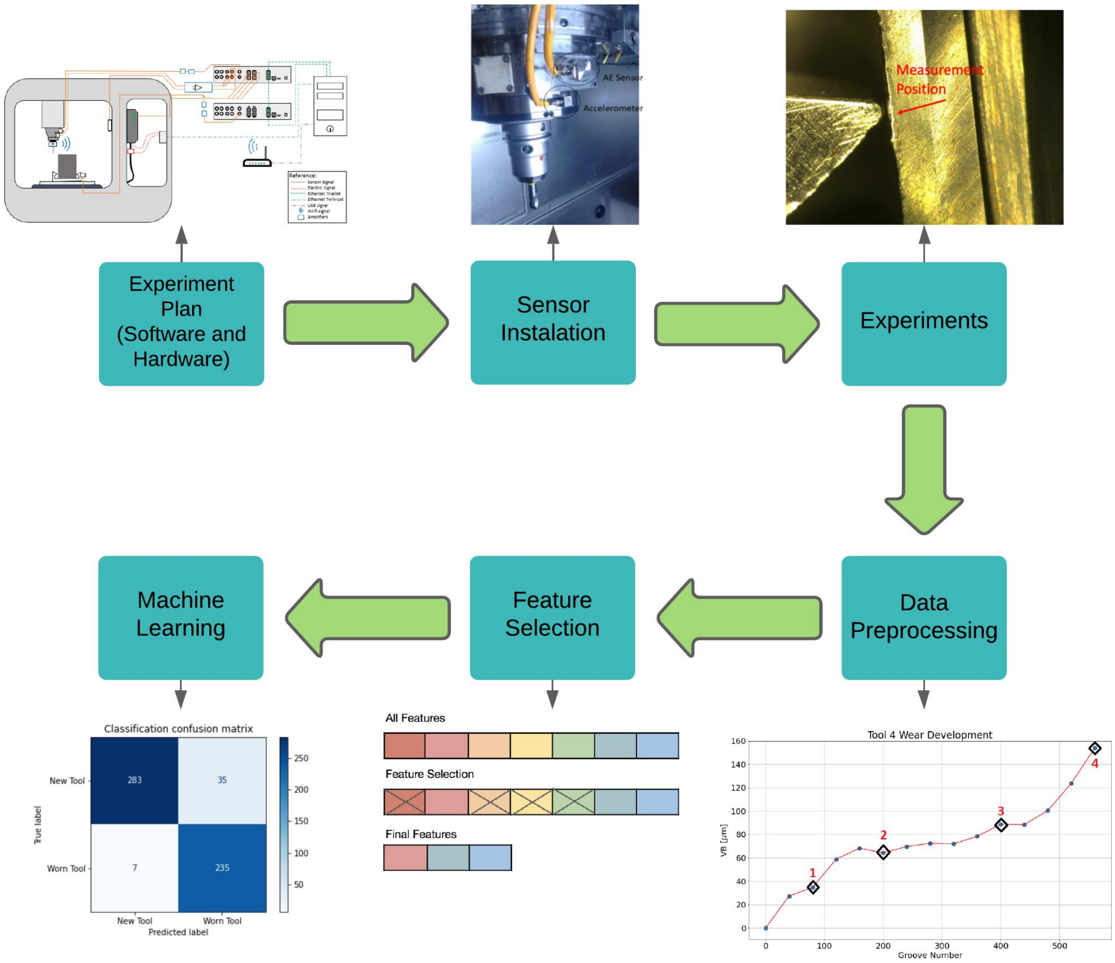
2.6 SOLUTION APPROACH

This project aims to study, analyze and summarize feature selection methods for a tool wear prediction system in high-speed precision milling processes. As described before, flank wear is the most prevalent type of wear and commands the other types of wear which, in normal conditions, occur in the final stages of flank wear. Therefore, this project will use flank wear as tool wear estimation.

This study will also use ML techniques to predict tool wear and evaluate the feature selection methods. In order to have a bigger view of the feature selection techniques employed in the project, four different types of ML techniques were used, being the ones explained in the section 2.4.

In addition, a set of experiments with different process parameters and direct measurement of flank wear of the tool was carried out to produce data to be used in the training of ML models. A multi-sensor system was installed in the CNC machine to conduct the best measurement strategy. The Figure 14 lists the methodology used in this project in order to get the conclusion about feature selection methods applied to tool wear prediction using ML techniques.

Figure 14 – Project process workflow.



Source – Author's image

This study explores every possible combinations of feature selection and ML techniques amongst the methods selected for this work, even the one used as internal classifiers in wrapper methods of feature selection. So, by analyzing the performance of the combinations, it is possible to identify the best combination of feature selection methods, ML technique and the number of features needed to produce a good result of tool wear prediction.

3 EXPERIMENT SETUP

In this chapter, a review of the components used in the experiment for the production of the data used in the work will be explained. Also, a description of the different setups of milling with different types of tools and machine variables will be covered. The chapter will then, explain the process of labeling the data in the microscope.

3.1 MACHINE

The CNC machine used in the scope of this work, see in Figure 15, is the HSC 55 linear from the company DMG MORI, designed for high-precision milling. It possesses a Heidenhain CNC system with three linear axes and a robust spindle that rotates up to 28000 RPM. The Table 1 provides more details about the CNC machine.

Figure 15 – DMG MORI HSC55 linear CNC Machine.

Source – Arena (2022)

Table 1 – CNC machine technical information.

DMG MORI HSC 55 linear		
Component	**Description**	**Unit**
Axes	3 (X, Y, Z)	-
Working area	(X, Y, Z) = (450, 580, 460)	mm
Max. Spindle Speed	28000	RPM
Max. Spindle Power	27	kW
Max. Spindle Torque	38	Nm
CNC System	HEIDENHAIN iTNC 530	-
Axes Encoders	HEIDENHAIN LC483	-
Axes Encoder Accuracy	± 5	µm
Axes Encoder Measuring Step	100	nm

Source – MORI (2022)

3.2 SENSORS

In this section, the sensors used in the experiments for gathering the data used in the project will be explained. Furthermore, the positioning of the sensors in the machine and the frequency of acquisition of each sensor will be quoted.

3.2.1 Encoder Position

The first source of information gathered for the project is the position of the axes while the process of milling is running. To get this information, the encoders from the CNC machine were used. Later this information will be used to help in the filtering of the data, explained in the next chapter.

The DMG MORI machine has an LC483 absolute positioning encoder. All three axes and the spindle of the machine will be recorded. In Figure 16 the encoder is presented and in the Table 2 technical information of the encoder is given.

Figure 16 – Heidenhain LC483 encoder.

Source – Heidenhain (2021a)

Table 2 – LC483 encoder technical information.

Heidenhain LC483 linear Encoder		
Component	**Description**	**Unit**
Calculation time	≤ 5	µs
Encoder Accuracy	± 5	µm
Encoder Measuring Step	100	nm

Source – Heidenhain (2021b)

3.2.2 Vibration

The vibration was another source of information gathered for the work. The removal of material from the workpiece provokes vibration in the spindle and in the table where the workpiece is placed.

For the acquisition of the vibration in the spindle body, the PCB 356B21 was chosen, and for the acquisition of the vibration in the workpiece table the PCB 356A15 was chosen, both sensors from the PCB Piezotronics company. The sensors record the vibration in the 3 axes separately.

In Figure 17 and Figure 18, the sensors are presented and in Table 3, technical information from the sensors was described.

Figure 17 – PCB 356B21 Accelerometer.

Source – Piezotronics (2022b)

Figure 18 – PCB 356A15 Accelerometer.

Source – Piezotronics (2022a)

Table 3 – Accelerometers technical information.

	PCB 356B21		PCB 356A15	
Component	**Description**	**Unit**	**Description**	**Unit**
Sensitivity (±10 %)	1.02	mV/(m/s^2)	10.2	mV/(m/s^2)
Measurement Range	±4905	m/s^2 pk	±490	m/s^2 pk
Frequency Range (Y or Z axis)(±5 %)	2 to 10000	Hz	2 to 5000	Hz
Frequency Range (X axis)(±5 %)	2 to 7000	Hz	1.4 to 6500	Hz
Resonant Frequency	$\geq$ 55	kHz	$\geq$ 25	kHz
Broadband Resolution	0.04	m/s^{2}rms	0.002	m/s^{2}rms

Source – Piezotronics (2022b,a)

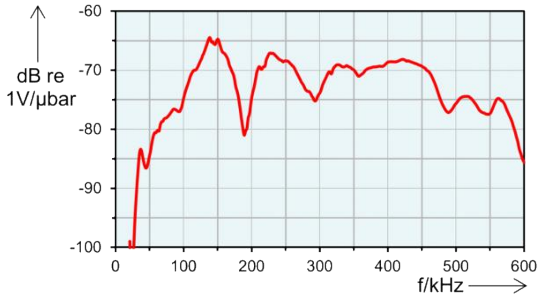
3.2.3 Acoustic Emission

Acoustic Emission (AE) was another source of information gathered during the milling process to generate data for the project. In this project, the AE will be measured in both the spindle body and workpiece table just as vibration.

The AE sensor chosen is the VS150-K3 from Vallen Systeme. The Figure 19 shows the sensor, Figure 20 shows the response graphic of the sensor and Table 4 shows technical information about it.

Figure 19 – VS150-K3 Acoustic Emission.

Figure 20 – VS150-K3 Response.



Source – Vallen (2022)

Table 4 – VS150-K3 Acoustic Emission technical information.

Vallen VS150-K3 AE Sensor		
Component	**Description**	**Unit**
Frequency Range (f peak)	100 to 450 (150)	kHz
Capacity	350	pF

Source – Vallen (2022)

3.2.4 Ultra-sonic Acoustic Emission

An Ultra-sonic microphone sensor was installed in the machine to record information of ultra-sonic acoustic emission of the milling process. The sensor selected was the PCB 130A24 from PCB Piezoeletronics. The Figure 21 shows the sensor and the Table 5 shows technical information about it.

Figure 21 – PCB Piezoeletronics 130A24 Microphone.

Source – Piezoeletronics (2022)

Table 5 – PCB 130A24 Microphone technical information.

PCB 130A24 Microphone Sensor		
Component	**Description**	**Unit**
Frequency Response (±3dB)	20 to 16000	Hz
Sensitivity (±dB)	1	V/Pa
Inherent Noise	20	μPa

Source – Piezoeletronics (2022)

3.2.5 Cutting Force

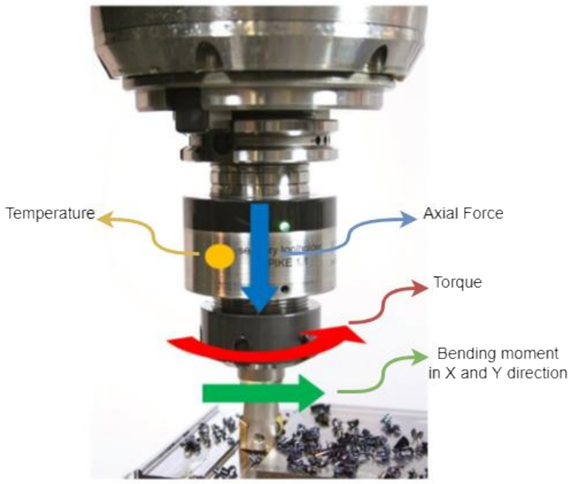
The Spike 1.2 tool holder from Pro-Micron company was selected to record information of cutting force due to the fact that a device installed in the spindle is more appropriate than one installed in the workpiece table based on the distance between the tool and the tool holder.

The Spike is a wireless smart tool holder which can record temperature, axial force, torque and bending moment in X and Y directions. Even though the Spike presents a large gamma of information about the tool, its application in the industrial environment is undesired due to its costs and the downtime of the machine to recharge the remote unit.

The Figure 22 shows the Spike tool holder, the Figure 23 shows the information that can be gathered with the Spike and the Table 6 summarizes technical information about it.

Figure 22 – Spike 1.2 tool holder.

Figure 23 – Spike 1.2 tool holder record signals.



Source – Pro-micron (2022a)

Table 6 – Pro-micron Spike 1.2 technical information.

PCB 130A24 Microphone Sensor		
Component	**Description**	**Unit**
Signals	Axial Force	N
	Torque	Nm
	Bending moment in X/Y	Nm
	Temperature	°C
Frequency Response	1600	Hz
Axial Force measuring range	60	kN
Torque measuring range	400	Nm
Bending Moment measuring range	400	Nm
Axial Force resolution	<5	N
Torque resolution	<0.03	Nm
Bending Moment resolution	<0.03	Nm

Source – Pro-micron (2022b)

3.3 V-BOX

The V-box is a Data Acquisition Device (DAQ) developed by Fraunhofer IPT for high-frequency data acquisition. Its goal is to provide support in the acquisition of multiple sensors at the same time, being possible to interconnect up to 30 V-boxes.

The V-Box, seen in Figure 24, has 8 analog inputs tuned in 80kHz and 2 channels for high-speed data acquisition tuned in 5MHz, called High Speed Analog Input (AIHS). It also has 4 analog encoder ports which can acquire signals from the encoder in the machine using a signal splitter.

Figure 24 – Fraunhofer V-Box.

Source – Fraunhofer (2022)

3.4 MICROSCOPE

To label the data during the experiments, a microscope was used to observe the tool and measure the wear of it. The microscope selected was the VHX-500F from the company Keyence. The Figure 25 shows the microscope and the Table 7 shows technical information about it.

Figure 25 – Keyence VHX-500F Microscope.

Source – Keyence (2022)

Table 7 – VHX-500F Microscope technical information.

VHX-500F Microscope	
Component	**Description**
Image Capture Device	1/1.8-inch, 2.11 million-pixel CCD image sensor Total pixels: 1688 (H) x 1248 (V) Effective pixels: 1628 (H) x 1236 (V) Virtual pixels: 1600 (H) x 1200 (V)
Electronic Shutter	AUTO, MANUAL, OFF, 1/15, 1/30, 1/60, 1/120, 1/250, 1/500, 1/1000, 1/2000, 1/5000
White Balance	Auto, Manual, One-push set, Preset (2700K, 3200K, 5600K, 9000K)

Source – Keyence (2022)

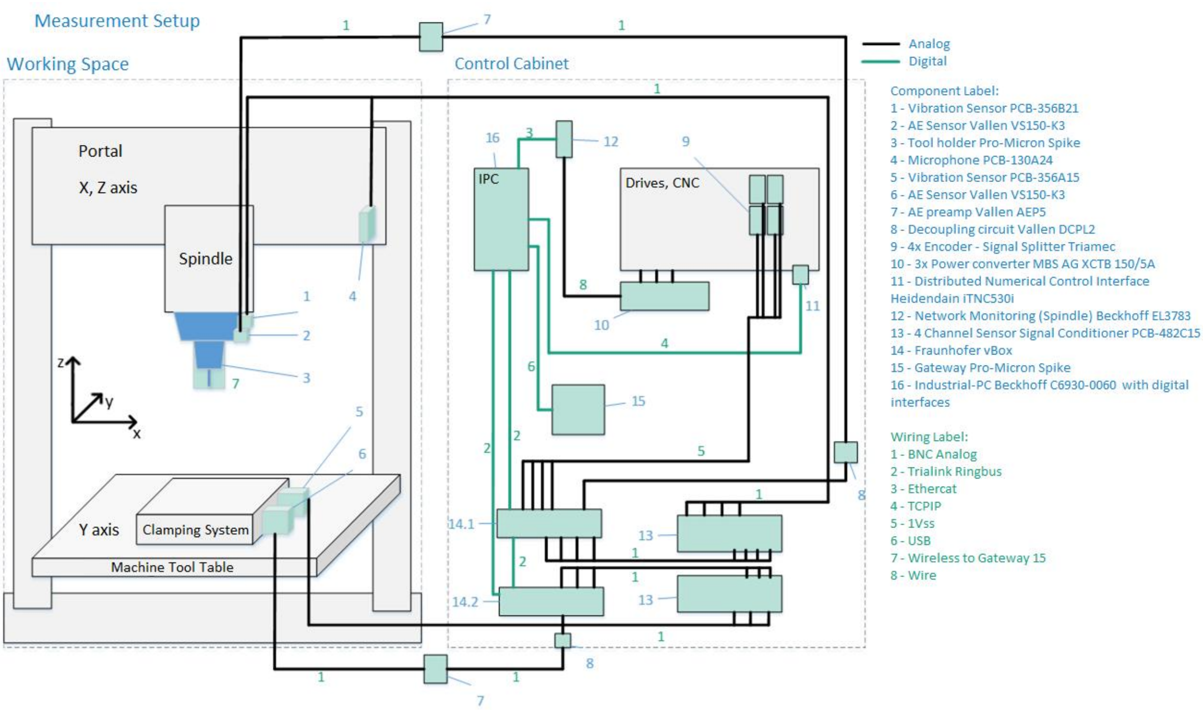
3.5 HARDWARE SETUP

Big experiments, like the one used in the scope of this project, consume time and resources mainly machine tools and materials. Therefore, the experiment done in this project were designed not only for the scope of this project but also to study other phenomena.

The data acquired in this experiment will also provide information for other projects. Thus, the experiment planning includes sensors and components that further will not be used in the scope of this work. Since only one V-Box is not enough to handle all the sensors alone, due to the V-Box characteristic, two V-Boxes are going to be used to acquire data from the sensors.

The Figure 26 displays the complete setup of the hardware with connections and amplifiers needed. The Table 8 complements the image mentioned before by denoting the V-Box connections and acquisition frequency.

Figure 26 – Machine hardware setup.



Source – Fraunhofer IPT

Table 8 – V-Boxes Connections.

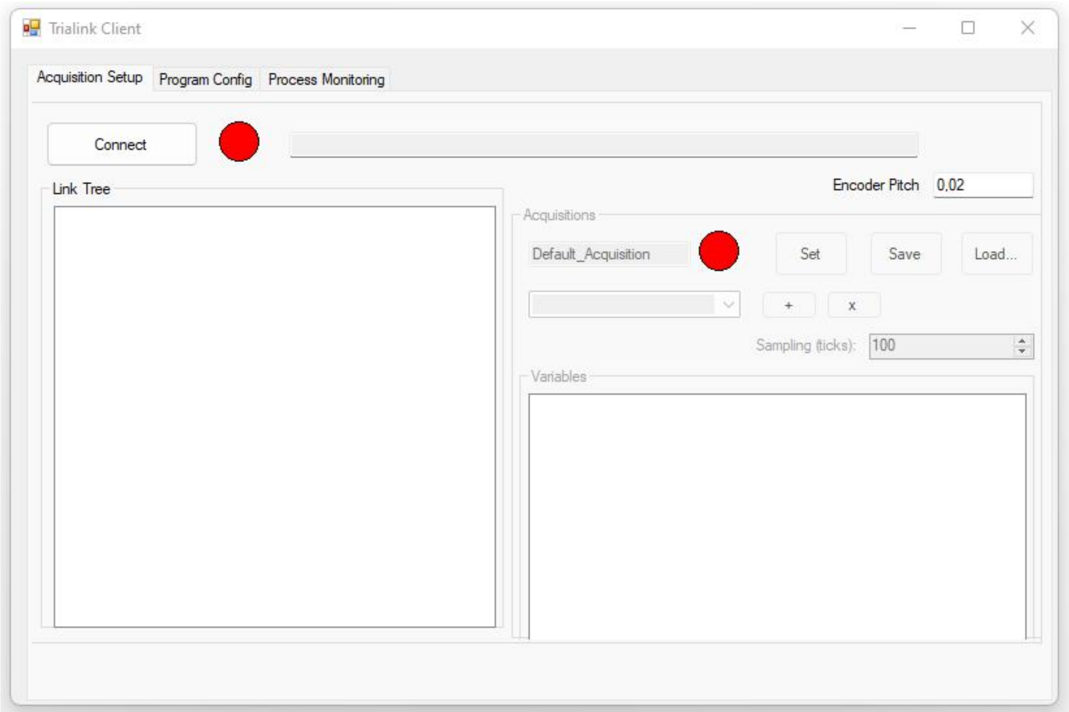
Source	Signal Handler	Destination	Frequency
AE - Spindle	AEP5 + DCPL2	AIHS1 - V-Box 1	100kHz
AE - Workpiece	AEP5 + DCPL2	AIHS1 - V-Box 2	100kHz
Accelerometer Spindle (X,Y,Z)	482C15	Analog Input (Ai)(1-3) - V-Box 1	50kHz
Accelerometer Table (X,Y,Z)	482C15	Ai(1-3) - V-Box 2	50kHz
Microphone	482C15	Ai4 - V-Box 1	50kHz
Encoders (X,Y,Z,Spindle)	Signal Splitter Triamec	Enc(1-4) - V-Box 1	50kHz

3.6 SOFTWARE SETUP

The information used in the scope of this project was from the V-Boxes. By saying that, the software used to gather and handle the information was the proprietary software from Fraunhofer IPT called Trialink Client. The system is responsible for communicating and handling the information from the V-Box.

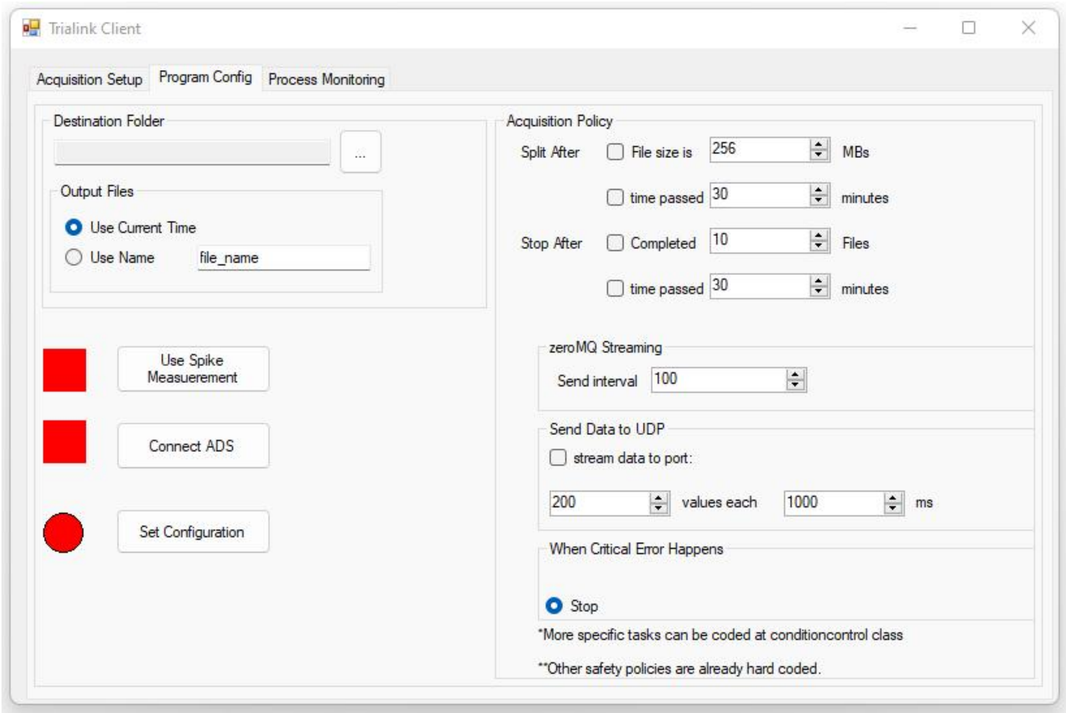
The images Figure 27, Figure 28 and Figure 29 show the pages of the software and its possible configurations. The Figure 30 gives an example of a file generated by the proprietary software with the signals acquired to the scope of this project.

Figure 27 – Acquisition Setup tab of Trialink Client.



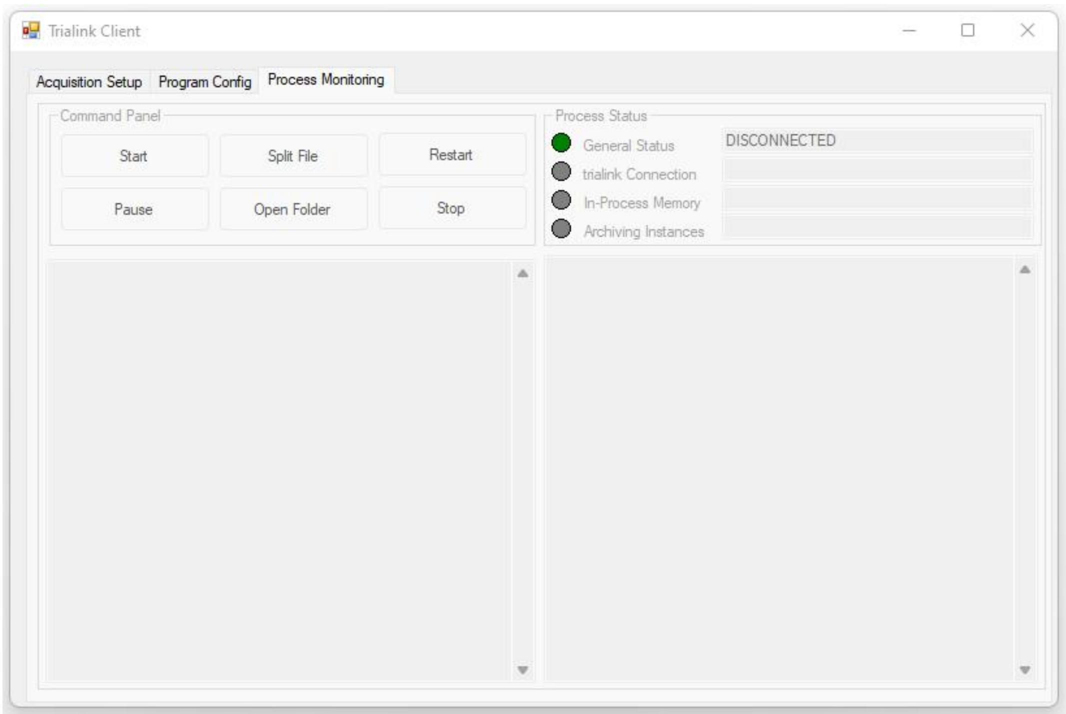
Source – Author's Image

Figure 28 – Program Configuration tab of Trialink Client.



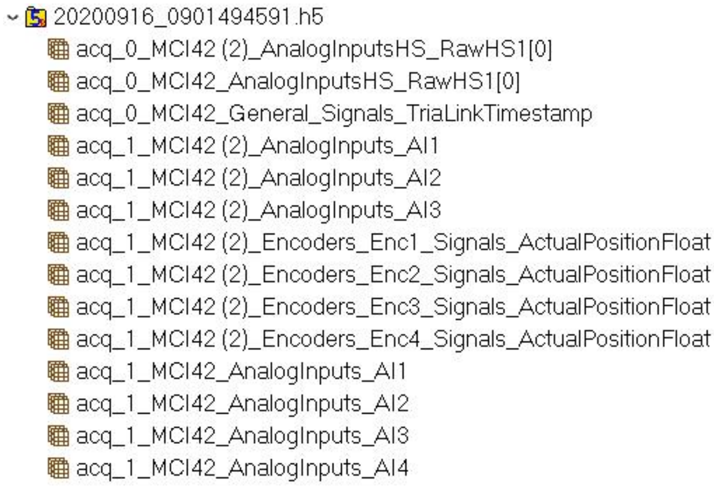
Source – Author's Image

Figure 29 – Process Monitoring tab of Trialink Client.



Source – Author's Image

Figure 30 – File generated.



Source – Author's Image

3.7 PROCESS SETUP

After finishing the hardware setup, the characteristics of the experiments and process were defined based on previous work done at Fraunhofer IPT. The selected tools for the experiments were the Pro Steel solid carbide roughing end mill HPC 8 mm and HOLEX Pro INOX solid carbide milling cutter HPC 8 mm, both from the manufacturer Holex. Further information about the tools can be found in the Table 9. The workpiece block is a steel plastic mold 40CrMnNiMo8-6-4.

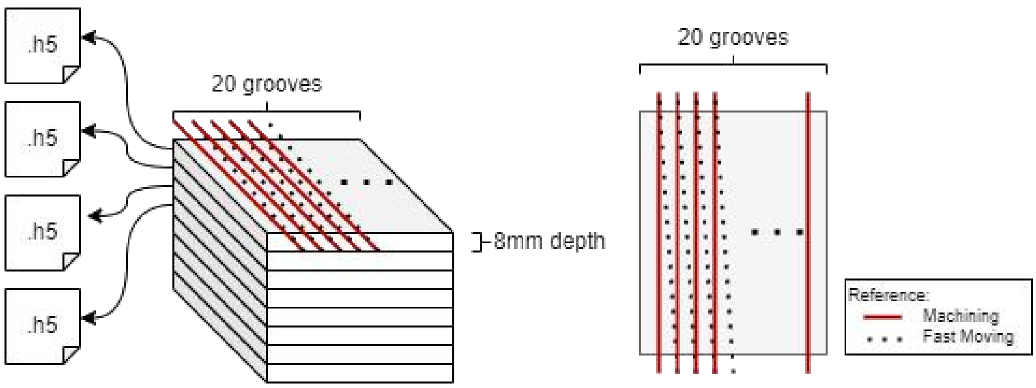
Table 9 – Tools technical information.

Tools Description		
Name	Pro Steel	Pro INOX
Article Number	2024148	20301510
Cutting Edge	8mm	10mm
Cutting Length	19mm	22mm
Overall Length	63mm	72mm
Number of Teeth	3	4
Coating	TiAIN	AlCrN
Tool Material	Solid Carbide	Solid Carbide

Source – Holex (2022b,a)

To experiment consists of making successive groves in the workpiece block to wear the tool as fast as possible. The cutting width used is 5mm for the tool with three teeth and 4mm for the tool with four teeth. Twenty grooves were produced in each depth layer in the workpiece due to the size of the block. The tool wear measurement was done after two complete layers for the tool with three teeth and after four entire layers for the tool with four teeth. The measurement of tool wear (VB) and the data labeling will be described in the next section.

Figure 31 – Process path planning.



Source – Author's Image

The experiment also combined different machine parameters seen in Table 10. A total of 8 tools were used in this experiment.

Table 10 – Machine parameters plan.

EX.ID	Tool teeth	Spindle speed [1/min]	Feed rate [mm/min]	Cutting width[mm]	Cutting depth[mm]	VB_{max} [μm]
1	3	6760	1160	5	8	110
2	3	6760	1160	5	8	110
3	3	7440	1160	5	8	110
4	3	7440	1160	5	8	110
7	4	6800	1290	4	8	110
8	4	7900	1290	4	8	110
11	4	6800	1290	4	8	110
12	4	7900	1290	4	8	110

4 DATA PRE-PROCESSING

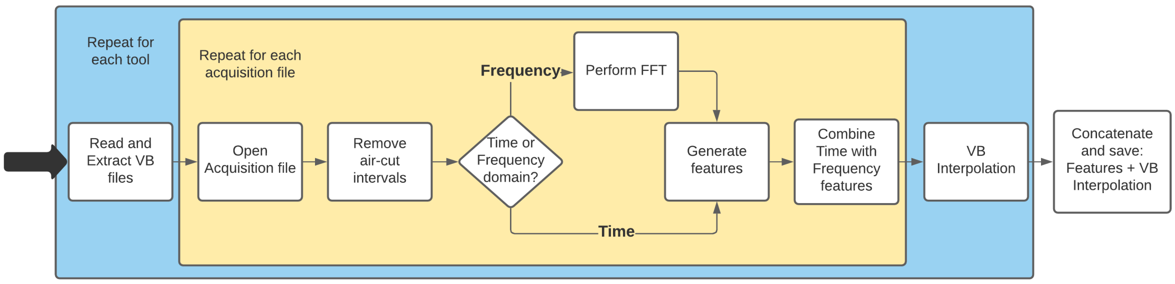
This chapter is responsible for the explanation of the path of the data acquired, since the gathering of it by the sensors installed in the machine until the dataset is ready to be applied in the feature selection methods to then further be used in a ML model to predict tool wear as well as describing the features extracted in the signals.

4.1 PRE-PROCESSING DATA FLOW

The performed experiments were done in a way that each acquisition data file has information of each level of milling, summing twenty grooves in each file. The Figure 31 shows each file, with extension .h5, generated in the milling process. For the tool, shown in Table 10, with three teeth, the flank wear of the tool was measured in the microscope after two layers of milling and for the tool with four teeth, after four layers of milling the tool was taken to the microscope to measure the flank wear, this choice was based on previous finished work.

The pre-processing stage is responsible for the extraction of information from the signals recorded during the milling process as well as the extraction of the flank wear data in the microscope in order to label the data further used in the ML models.

Figure 32 – Flowchart of the pre-processing stage.



Source – Author's Image

On Figure 32 a sequence of operations and actions performed in the pre-processing phase is presented. Those steps are done in order to extract the information from the acquisition signals and Flank Wear (VB) measurement files. The result of the data pre-processing stage is the complete dataset ready to be used in the feature selection and ML models to predict tool wear in the next steps of the project.

In the beginning, an analysis of the flank wear microscope measurement file for each tool is done, and for the points gathered is possible to estimate the flank wear curve for the referred tool and study the behavior of the tool during its lifetime. For that, polynomial interpolation is done.

The next step is to read the acquisition files generated during the milling process and, with the information on the encoder position of the spindle, the signals are filtered by removing the signal interval when the machine tool is far from the workpiece, meaning the tool is not milling at the time.

After that, time and frequency signals are separated. For the time signals, the features are extracted directly with no step needed before. On the other hand, for the frequency signals, the FFT is performed before the extraction of the features in the signal.

Finally, the time and frequency features are then combined with the interpolated tool wear curve, this way there will be one label, VB calculated from the interpolation, for each instant in the dataset.

This procedure is then repeated for all 8 tools used in this project, resulting in 8 files to be used further in the process. Furthermore, in the next sections, the steps presented above will be completely described.

4.2 FLANK WEAR MEASUREMENT AND EXTRACTION

4.2.1 Microscope data acquisition

As mentioned before and for Dolinšek and Kopač (2006), flank wear is the most typical and common wear of a tool under normal cutting conditions and other types of wear occur as a consequence of flank wear progress, therefore measuring the flank wear of a tool is used to indicate the total wear of a tool. In this work, measuring of flank wear will be adopted to indicate the actual tool wear.

The measuring procedure adopted consists of observing the wear surface length on the edge of the tool parallel to the cutting surface. This way it becomes possible to determine the measurement points and material loss of the tool. As presented in Table 9 the tools used in this project have a helical geometry, seen in Figure 33, meaning that additional care should be taken to measure precisely as possible the wear of the tool.

Figure 33 – Tool on Microscope.

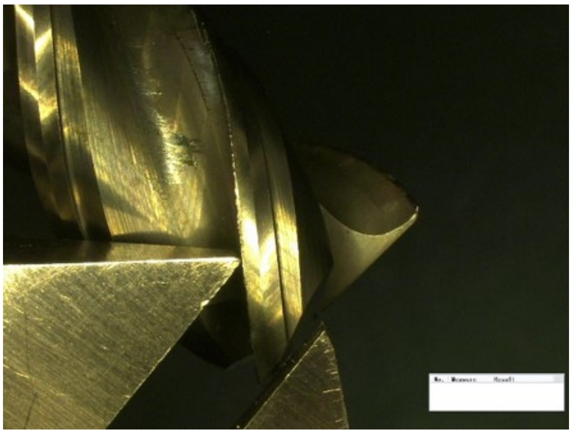
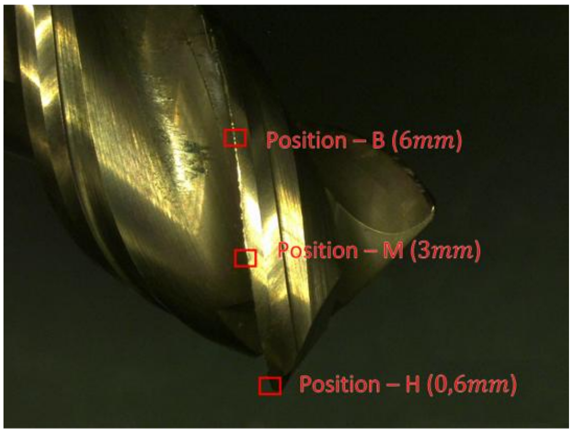


Figure 34 – Measurement positions.



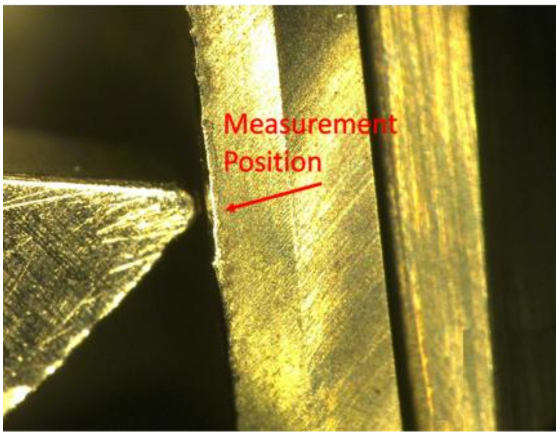
Source – Author's Images

On each tooth of the tool, a total of three measuring points are made as seen in Figure 34. The closest one to the tooltip (called H), approximately 0.6mm from the tip; The second one at a distance of 3mm from the tip (called M); and the third one at a distance of 6mm from the tip (called B). Therefore, for the tools used in this work, a total of nine and twelve points were measured for the tool of three and four teeth respectively. As recommended by the literature, the maximum value of these measured points will be adopted as tool wear.

To guarantee that the measuring points will be at the same exact point during the lifetime of the tool, a caliper is used to measure the distance of the tooltip to the point where the wear should be measured. In Figure 33 is possible to see the caliper with the tool.

Another consideration to be taken is that due to the tool having a helical geometry, different angles of viewing the tool could infer distortions in the measurement. Therefore, the light emitted by the microscope must be reflected in the tool at the exact measuring point. The Figure 35 shows the light position when doing a measurement.

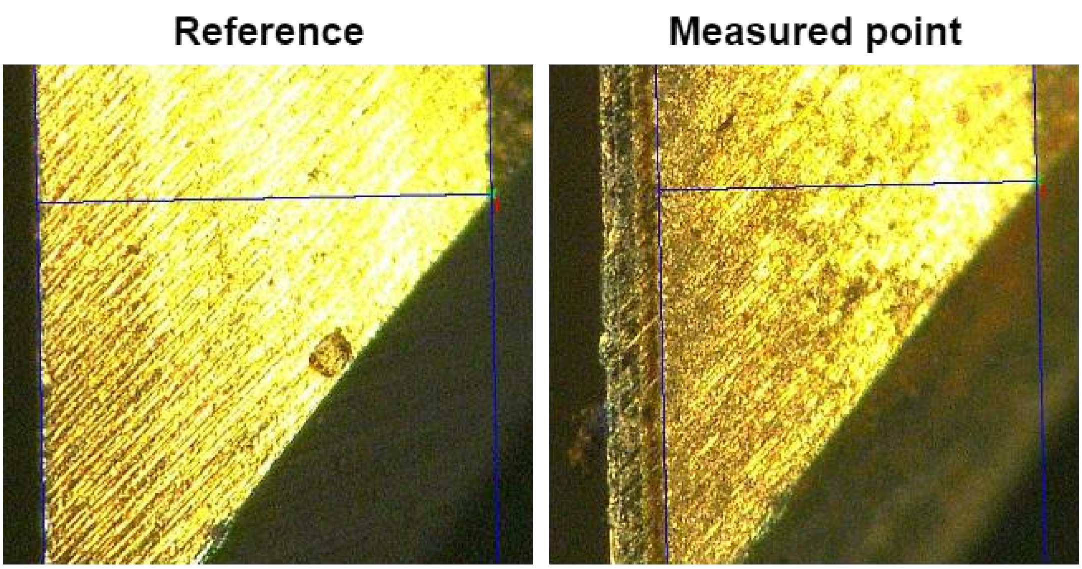
Figure 35 – Measurement point.



Source – Author's Image

At last, the flank wear is measured indirectly in the microscope. Mainly because of the deterioration of the tool it becomes more and more difficult to determine the wear of the tool. Thus, the strategy to measure the flank wear is: with the tool in perfect state, which means no usage at all, a reference is measured for each measurement point. After the usage of the tool, the flank wear is determined by the reference minus the value measured. The Figure 36 shows an example of the measurement for the same tool in different stages of wear. The blue lines in the Figure 36 are from the measurement system of the microscope.

Figure 36 – Reference and Measured point.



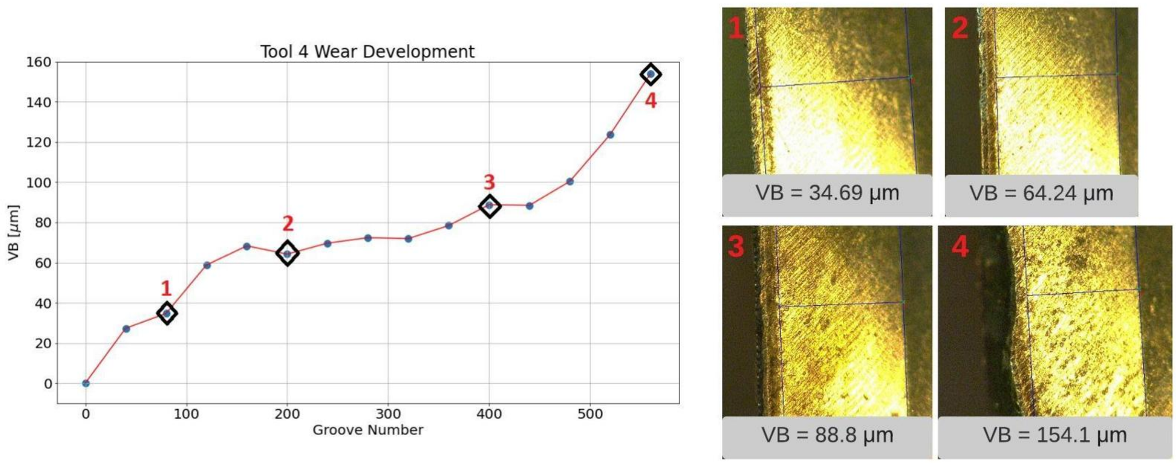
Source – Author's Image

4.2.2 Flank wear development

During the performance of the experiments, it was possible to confirm the pattern reported in Dolinšek and Kopač (2006) for the flank wear development curve. In the early stages of flank wear development - until around 30µm - there is a rapid development of the wear caused by material loss on the tip of the tool. Then the wear stabilizes and grows gradually until it reaches around 75µm when it starts to present other types of wear like chipping and crater wear.

After passing this point, the wear starts to grow rapidly again due to the intensification of chipping and crater wear. The tool also starts to lose material on the edge. At this point of degradation, it is possible to notice an increase in the noise produced by the CNC machine during the milling process. In the Figure 37 it is possible to observe such wear development alongside the microscope measurement pictures for the highlighted points in the curve.

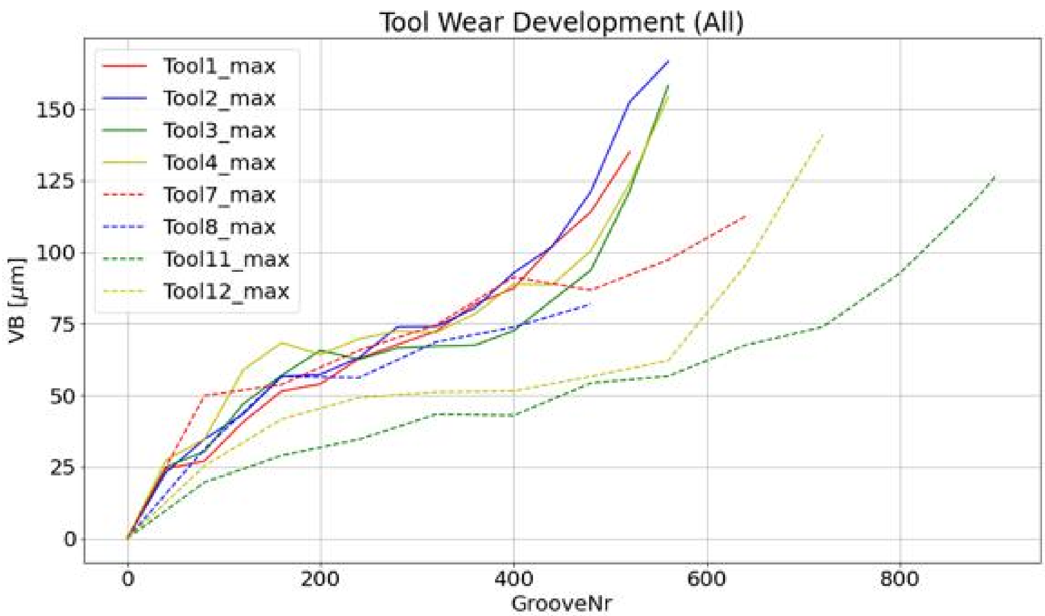
Figure 37 – Tool 4 wear development.



Source – Author's Image

As mentioned before and seen in Figure 37 with the increasing of the wear in the tool it becomes harder to observe and measure the flank wear. Due to this fact and experience gathered from previously finished work done at IPT the measurement system used in this work was done using a reference instead of measuring it directly.

Figure 38 – Tool wear development for all tools.



Source – Author's Image

The Figure 38 presents the wear development for all eight tools used in the scope of this project. Again it is possible to observe the wear development before described for all tools. It is also possible to observe the difference between the tools with three and four teeth, the one with more teeth lasts longer than the other one.

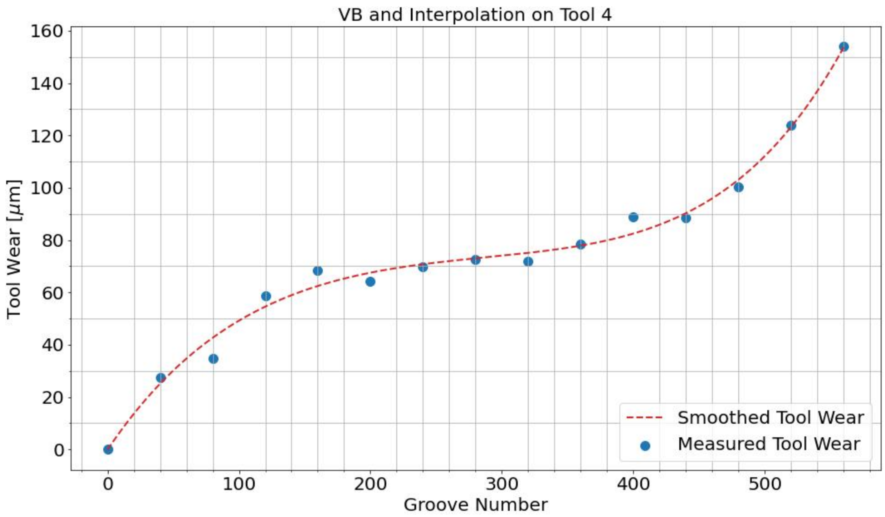
The pre-processing stage starts with obtaining the wear curve for each tool. As described in the section before, for each measurement there is a total of nine points for a tool with three teeth and twelve points for a tool with four teeth. Therefore the value chosen to represent the flank wear at that time of measurement is the maximum of the points measured, also seen in Figure 38.

Since it is unfeasible to have an all-time measuring of the flank wear, it is required an interpolation between the measured points in order to produce values of VB for the entire tool milling operation time. The most common approach is to maintain the previous value until the next one arrives, creating a curve of steps.

But in the scope of this project, a polynomial interpolation was chosen to generate the flank wear curve. Based on the behavior of the wear curve observed in Figure 37 and described by Dolinšek and Kopač (2006), an odd degree must be chosen for the polynomial function, this way, a 5^{th} degree polynomial function was selected.

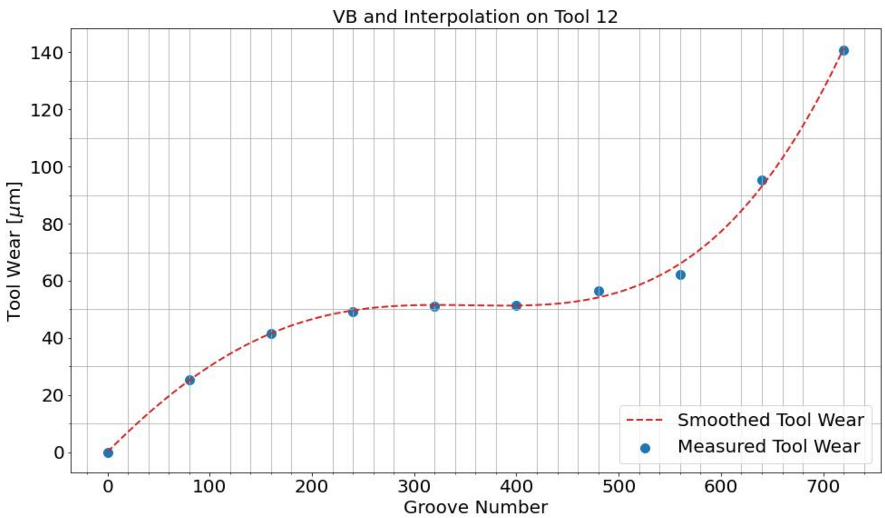
The 5^{th} degree polynomial interpolation was chosen due to the smoothness of the curve produced by it and helps filter the errors inserted by the microscope measurement procedure. The Figure 39 and Figure 40 represents the polynomial interpolation wear curve and the measured points for tool 4 and 12, with three and four teeth, respectively.

Figure 39 – Polynomial interpolation curve for tool 4.



Source – Author's Image

Figure 40 – Polynomial interpolation curve for tool 12.

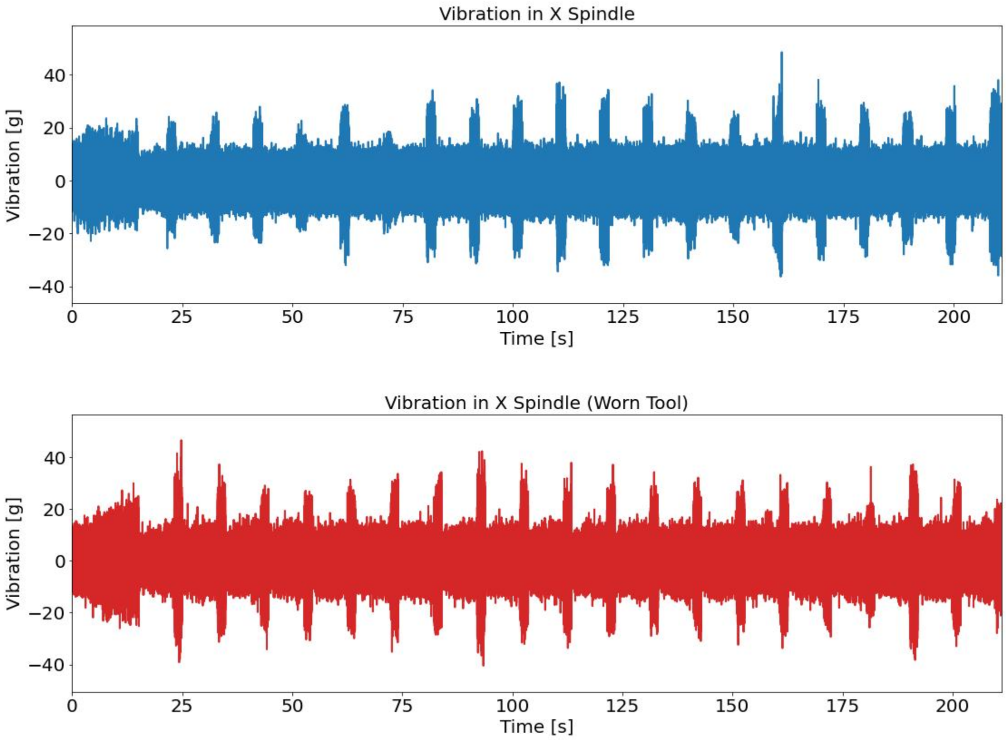


Source – Author's Image

4.3 AIR-CUT REMOVAL

The primary step done when working with data produced by an experiment is to observe and understand the behavior of such data. Having this in mind, a careful analysis was carried out in the data acquired from the milling process in the CNC machine.

Figure 41 – Vibration in X axis for New and Worn tool.



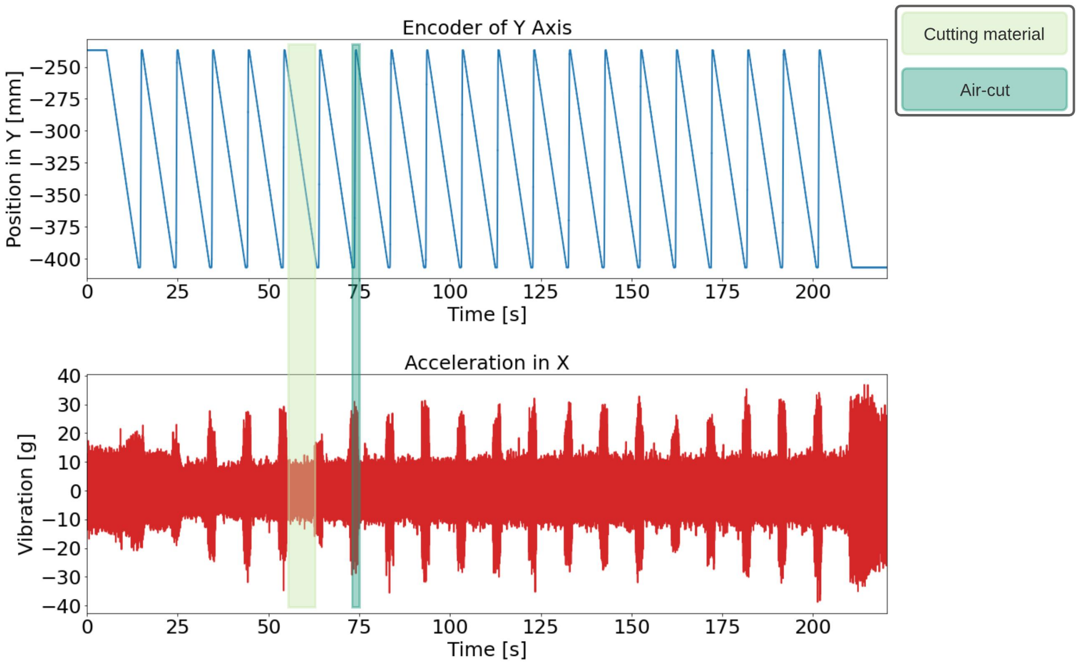
Source – Author's Image

The Figure 41 compares the same sensor signal, Vibration in X axis from the spindle, in the first twenty grooves with the last twenty grooves of the same tool during the milling process. During the analysis of the data, it was observed that some sensors acquired were not suitable for the project, this way some acquired sensors in the milling process were discarded from the project. The signals were mainly discarded due to the fact that the noise in the signal was prevalent above the process information, not being suitable for the project usage. Therefore, the selected signals were: Vibration in X, Y and Z from the spindle, Microphone and AE from the spindle.

Before the dataset is ready, a step of air-cut removal is necessary in order to delete data that is unnecessary for the project. Since the experiment, in other words, data production, is done by a sequence of grooves in the workpiece, at some point in the milling process, the tool will be elevated from the workpiece and return to the start position shifted accordingly as described in cutting width column in Table 10. This way the data will contain information of a movement not needed for the project.

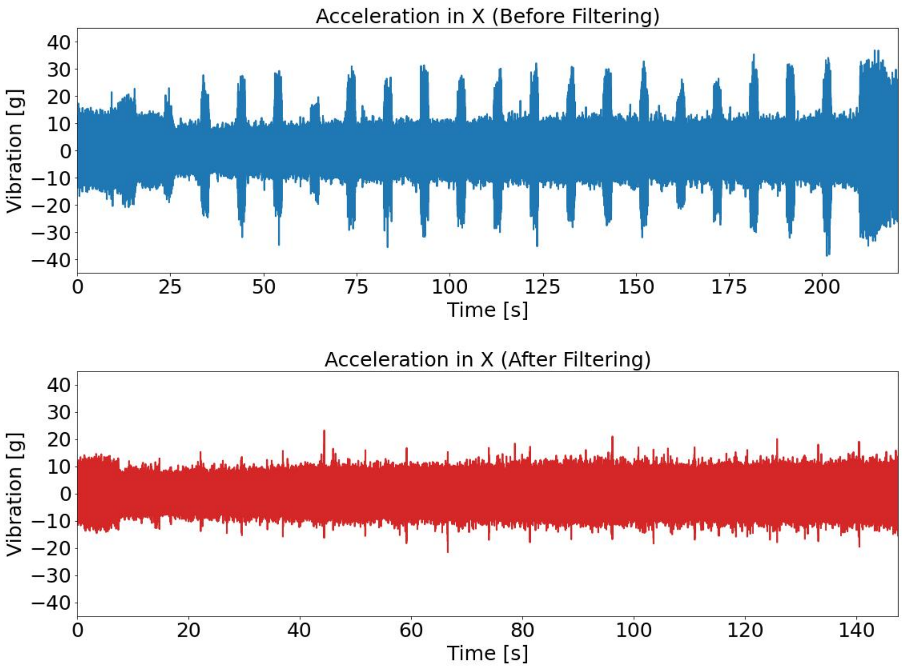
To perform the removal of such data, the information of Y axis encoder is used in order to evaluate the position of the tool during the process. It is only possible due to the fact that, since all information comes from the Fraunhofer V-Box, the information has the same timestamp, not being needed synchronization of different information.

Figure 42 – Air-cut removal from signal.



Source – Author's Image

Figure 43 – Before and After air-cut removal.



Source – Author's Image

The Figure 42 shows a comparison of the Y-axis encoder signal and X-axis accelerometer signal. The Figure 43 shows a comparison of the same signal, X-axis accelerometer, before and after the air-cut removal procedure.

4.4 FEATURE EXTRACTION

At this stage of the process, the time domain signals are ready to have features extracted to it while the frequency domain signal needs a previous step in order to be ready. The step needed is the execution of a FFT in order to reconstruct the spectrum of the signal in frequency without losing any information about the process.

The signals, which were acquired in the time domain, need to be transported to the frequency domain. Therefore, for the signals of Vibration in X, Y and Z axis and Microphone, the execution of FFT is needed to reconstruct the signal properties in the frequency domain. For the AE signal, the execution of the FFT is not needed due to the fact that the acquisition of it from the Fraunhofer V-Box is different from the other signals.

The result from the acquisition of the AE sensor from the Fraunhofer V-Box is an encrypted signal. Each encrypted value has 16 bits corresponding to the frequency and 16 other bits corresponding to the amplitude. Therefore, when decrypting the signal, the FFT is extracted and the signal is produced already in the frequency domain.

Figure 44 – Representation of the FFT.

Source – Author's Image

The Figure 44 represents the FFT performance where, due to the raw signal being recorded in 50kHz frequency, the chosen window size was 5000 samples or 0.1 seconds and the overlap selected was 2500 samples or 0.05 seconds.

As the last step to be done before the dataset is ready to be used in the ML models and since the work focuses on feature selection, an extraction of different features on the sensors signals is needed.

Besides the variation of the features and accordingly to the literature, a set of eight time domain features and one frequency domain feature were extracted. The Table 11 shows and describes the features extracted.

Table 11 – Features extracted.

	Features	**Description**
Time Domain	Arithmetic Mean (M)	$M = \frac{1}{n}\sum_{i=1}^{n} X_i$
	Root Mean Square (RMS)	$RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}$
	Variance (V)	$V = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n-1}$
	Skewness (Sk)	$Sk = \frac{1}{n}\frac{\sum_{i=1}^{n}(X_i - \mu)^3}{\sigma^3}$
	Kurtosis (Ku)	$Ku = \frac{1}{n}\frac{\sum_{i=1}^{n}(X_i - \mu)^4}{\sigma^4}$
	Signal Power (P)	$P = \frac{1}{n}\sum_{i=1}^{n} X_i^2$
	Peak-to-peak Amplitude (pp)	$pp = max(X_i) - min(X_i)$
	Crest Factor (CF)	$CF = \frac{max(X_i)}{RMS}$
Frequency Domain	Mean of Band Power (MBP)	$MBP = \frac{1}{n}\sum_{i=1}^{n} S(f)_i$

4.5 PRE-PROCESSING RESULTS

Finally, after all the steps before described, the data is ready to be used in feature selection methods and further in the ML models to predict tool wear. The final concatenated array with all features can be seen in Table 12.

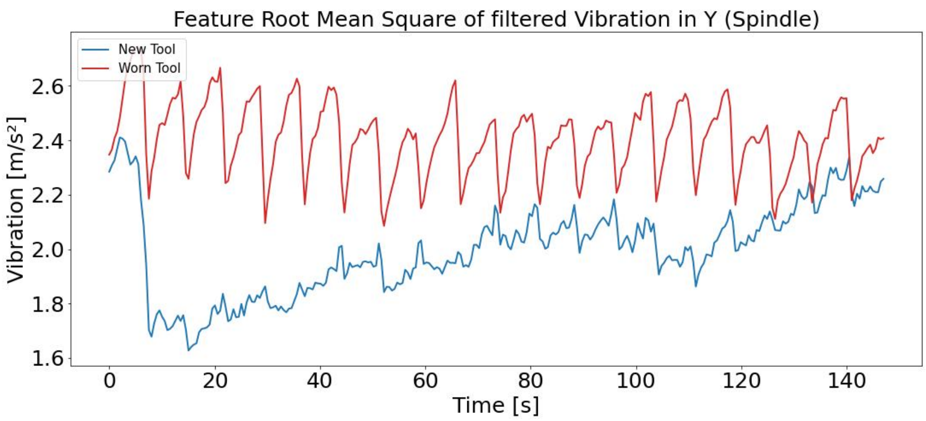
Table 12 – Final dataset array.

Time Domain Features																
Signal	Vibration in X (Spindle)								Vibration in Y (Spindle)							
Features	Mean	RMS	Variance	Skewness	Kurtosis	Signal Power	Peak to Peak	Crest Factor	Mean	RMS	Variance	Skewness	Kurtosis	Signal Power	Peak to Peak	Crest Factor
Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Signal	Vibration in Z (Spindle)								Microphone							
Features	Mean	RMS	Variance	Skewness	Kurtosis	Signal Power	Peak to Peak	Crest Factor	Mean	RMS	Variance	Skewness	Kurtosis	Signal Power	Peak to Peak	Crest Factor
Index	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31

Frequency Domain Features					
Signal	Vibration in X (Spindle)	Vibration in Y (Spindle)	Vibration in Z (Spindle)	Microphone	AE (Spindle)
Index	32 to 81	81 to 131	132 to 181	182 to 231	232 to 331

As can be seen from Table 12, and explained before, not all the sensors acquired in the experiment were used in the dataset. This fact is due to the information from the sensors installed in the workpiece clamp had a lot of noise which overcame the information of tool wear.

Figure 45 – Comparison between a new and worn tool signal.



Source – Author's Image

The Figure 45 shows us a difference between two conditions in the same signal, vibration in Y coordinate, of the same tool to observe the signal's behavior during its lifetime.

In the end, a total of 8 files were generated, one for each tool used in the experiment, which will be used later in feature selection and tool wear prediction with the ML models.

5 FEATURE SELECTION AND MACHINE LEARNING

This chapter is going to describe all the adjustments done in the ML and Feature Selection methods in order to obtain the widest range of possible adjustments. It will also point out the results of the adjustments done as well as some graphics and confusion matrix.

5.1 SCRIPT DESCRIPTION

Finally, after all the steps described, the dataset is ready to be used in the feature selection for further ML training in order to predict tool wear and evaluate the feature selection methods. The last remaining step in the project, before evaluation and conclusion of the subject, is to implement the feature selection methods and ML training script.

The Programming language used was Python due to the easiness to develop scripts and closeness to AI having a great number of methods already developed in this language. The resulting script is responsible for applying all the feature selection methods in the raw dataset and training all the ML methods for all the result datasets from the feature selection methods. In other words, it combines four feature selection methods with four ML training resulting in sixteen accuracy results every time the script is concluded.

The first step of the script is to load the data from the dataset resulting from the last chapter. After that, the four ML are trained with the dataset with all features to compare the accuracy result after the features are selected. Subsequently, the script runs the Pearson Correlation feature selection method and then, trains all four ML methods with the resulting dataset of features selected.

Secondly, the Spearman Correlation is done and the ML models are trained with the features selected by this method. In sequence, the forward feature selection method is applied and the ML models are trained. Lastly, the backward feature elimination method is then used and the models are trained with the resulting dataset of the feature selection method.

Finally, the script finishes testing and saving all accuracy from all sixteen results being: four ML models (KNN, SVM, LR e PER) for Pearson Correlation; four ML models for Spearman Correlation; four ML models for Forward feature selection for one ML model as internal classifier; and four ML models for Backward feature elimination for one ML model as an internal classifier.

The balanced accuracy method is used to evaluate the accuracy of each ML model. This method was selected due to the advantage of it being better than accuracy for unbalanced data and being the same as accuracy whenever the data is balanced.

The script also computes the time spent to run it and gets the features selected by each method of feature selection, saving them in a sheet file to be later analyzed.

5.2 FEATURE SELECTION AND MACHINE LEARNING TUNING

In order to evaluate all possible situations, many changes in the script were made. These changes, called tuning, were either in the feature selection methods or in the ML classifier models inside the wrapper types of feature selection methods. The ML models, which were used to train the data after the features were selected, remained the same for the complete training. This is due to the fact that the fine-tuning of the ML models to obtain a better accuracy is not needed. However, it is imperative that it remains the same to better analyze the impact of the feature selection methods on the overall accuracy. Once tuned, the ML classifier models didn't change either.

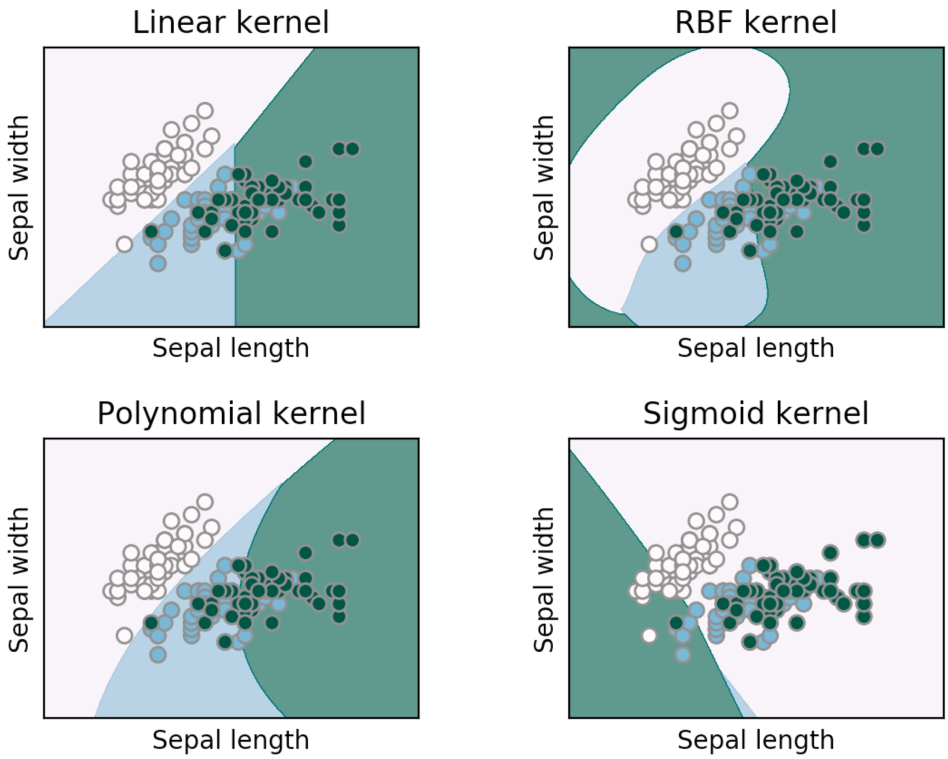
5.2.1 Machine learning tuning

The ML parameters, either for the models or for the internal classifiers are the same. This is done to make the tuning an easy process and because the project doesn't need to obtain the best result for the training.

For the KNN ML model, the parameters tuned were the number of neighbors, which was chosen as 25 neighbors, and the weight as distance, which assigns weights proportional to the inverse of the distance from the point. In other words, when the neighbor is closer to the analyzed point, it weighs more in deciding which class the analyzed point is.

The kernel parameter was the only one tuned for the SVM ML model. It was chosen as the linear kernel. This parameter selects the type of function that will be used to separate the hyperplane between the classes. The Figure 46 shows examples of different kinds of kernels used in classification with SVM models.

Figure 46 – Types of SVM Kernels.



Source – science (2022)

In the LR ML model, two parameters were tuned, the "random_state" and the "max_iter". The first one indicates the seed of randomness used in the LR model. The second one is due to the fact that, since the LR model is a regression one, it needs to have a stopping point to the convergence of the solvers; hence the max iteration, in this case, selected as 1000, the model will have in order to classify the observed point.

As for the PER ML model, the parameters tuned were "random_state" and "tol". The first one has the same purpose as the one with the same name in the LR model. The second one indicates the stop criteria used in the model; in this case, the stop criteria used is $1e-3$.

Those are the changes, or tunes, made in all the ML models used in this project, either as internal classifiers in the wrapper feature selection methods or in the ML models used to evaluate the accuracy of the features selected.

5.2.2 Feature selection tuning

As a final step of the project, before the evaluation of the ML models and feature selection methods, changes in the feature selection methods are conducted in order to present a wide range of situations to compare with each other.

For the filter feature selection methods, the filter's threshold is changed. With this change, either the Pearson Correlation or the Spearman correlation select more or fewer features. This threshold is the value of the correlation between the input and output of the feature selection method. The values change from zero, which means no correlation and therefore all features are selected in the filter methods, to the maximum correlation, which means only one feature is selected. For the purpose of this project, the values were changed from 0.15 to 0.6 with an increase of 0.05 each time the filter method was executed.

As for the wrapper methods of feature selection, the parameter changed was the number of features to be selected by the methods. This parameter is connected with the correlation of the input features with the output of the method, which means the more features selected, the less intense the correlation between features and output and the fewer features selected, the stronger the correlation and, in theory, higher the accuracy of the ML model. A total of eight different number of features to be selected were changed in this project, being: $10, 15, 20, 25, 50, 75, 100$ and 150.

Those small changes were the key to the project, making it possible to cover a wide range of different models and study the influence of feature selection methods in the accuracy of the ML models. The results of those changes are going to be explained in the next section.

5.3 RESULTS

After the combination of all parameters listed in the section above was implemented, a study was conducted in order to analyze the accuracy results of the models and select the feature selection method which had the most significant impact on the project.

As will be shown next, applying any feature selection method improved the overall accuracy of the ML models. With all methods, filter or wrapper methods, it was possible to increase the accuracy of the ML models compared with the accuracy using all features as input of the models.

Some of the combinations used to train the ML models significantly increased the computer usage, sometimes taking more than 14 hours to finish the combination. Even though the combination produced a higher accuracy result, it is not completely useful due to the time consumed to train the model. Others combinations had similar performance with smaller time spent to complete the combination.

Figure 47 – Filter methods results.

	K-Nearest Neighbour			Support Vector Machine			Logistic Regression			Perceptron			Threshold	Number of Features	
	PC	SC	All Features	PC	SC	All Features	PC	SC	All Features	PC	SC	All Features		PC	SC
Balanced Accuracy	0.9190	0.9211		0.8835	0.8452		0.9009	0.8886		0.8994	0.9595		0.15	264	267
	0.9096	0.9161		0.9233	0.8944		0.9385	0.8763		0.9479	0.9660		0.2	232	246
	0.9052	0.9096		0.9110	0.9291		0.9262	0.9074		0.9363	0.9493		0.25	215	225
	0.9096	0.9103		0.9088	0.9059		0.9515	0.9038		0.9291	0.9501		0.3	177	203
	0.9023	0.9182	0.8784	0.8944	0.8987	0.8386	0.9067	0.9103	0.8864	0.9052	0.8632	0.9132	0.35	144	178
	0.8994	0.9124		0.9117	0.9507		0.9211	0.9377		0.9247	0.9551		0.4	77	154
	0.8965	0.9168		0.9103	0.9559		0.9081	0.9269		0.8929	0.9168		0.45	39	123
	0.8661	0.9290		0.9052	0.8314		0.9102	0.8531		0.9059	0.8321		0.5	24	34
	0.8683	0.7475		0.9182	0.8271		0.9247	0.8242		0.8394	0.8220		0.55	13	8
	0.8242	0.8488		0.8726	0.9175		0.8994	0.8886		0.7475	0.6708		0.6	6	3

Source – Author's Image

Figure 48 – K-Nearest Neighbour model as internal classifier.

	K-Nearest Neighbour as internal classification model												Number of Features
	K-Nearest Neighbour			Support Vector Machine			Logistic Regression			Perceptron			
	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	
Balanced Accuracy	0.9471	0.8516		0.9283	0.7713		0.9030	0.7518		0.9175	0.7604		150
	0.9703	0.8582		0.8813	0.8509		0.8386	0.8307		0.8640	0.9153		100
	0.9718	0.8459		0.9059	0.7402		0.8965	0.7482		0.7996	0.6498		75
	0.9370	0.8046	0.8784	0.9038	0.9182	0.8386	0.9624	0.9074	0.8864	0.9399	0.8734	0.9132	50
	0.9682	0.8003		0.9667	0.7323		0.9718	0.7106		0.9067	0.7533		25
	0.9674	0.8473		0.9645	0.7641		0.9747	0.7735		0.9667	0.7562		20
	0.9226	0.8980		0.9703	0.9378		0.9718	0.9219		0.9515	0.9182		15
	0.9009	0.9450		0.9479	0.9067		0.9537	0.9088		0.8596	0.9407		10

Source – Author's Image

Figure 49 – Support Vector Machine model as internal classifier.

	Support Vector Machine as internal classification model												Number of Features
	K-Nearest Neighbour			Support Vector Machine			Logistic Regression			Perceptron			
	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	
Balanced Accuracy	0.9081	0.8089		0.9059	0.8169		0.9088	0.7995		0.9095	0.8950		150
	0.9349	0.8220		0.8980	0.8075		0.9001	0.7981		0.9045	0.8184		100
	0.9327	0.8119		0.9009	0.8379		0.9204	0.8314		0.8763	0.8104		75
	0.9378	0.8726	0.8784	0.9334	0.8213	0.8386	0.9544	0.8104	0.8864	0.9320	0.8082	0.9132	50
	0.9262	0.8097		0.9385	0.8155		0.9493	0.8198		0.9385	0.8654		25
	0.9219	0.8321		0.9472	0.8560		0.9537	0.8495		0.9255	0.8278		20
	0.9161	0.8864		0.9479	0.9009		0.9551	0.8828		0.9276	0.8618		15
	0.9240	0.9457		0.9436	0.9407		0.9674	0.9566		0.9602	0.9182		10

Source – Author's Image

Figure 50 – Logistic Regression model as internal classifier.

	Logistic Regression as internal classification model												Number of Features
	K-Nearest Neighbour			Support Vector Machine			Logistic Regression			Perceptron			
	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	
Balanced Accuracy	0.9153	0.8379		0.9096	0.8669		0.9030	0.8886		0.9370	0.8864		150
	0.9124	0.8980		0.8893	0.8596		0.8871	0.8886		0.9038	0.8799		100
	0.9117	0.9117		0.8792	0.8669		0.8792	0.8922		0.9139	0.8466		75
	0.9161	0.8741	0.8784	0.9356	0.9016	0.8386	0.9269	0.8929	0.8864	0.8777	0.9045	0.9132	50
	0.9443	0.8857		0.9609	0.8611		0.9624	0.8654		0.9414	0.8509		25
	0.9443	0.9067		0.9624	0.9211		0.9624	0.9146		0.9081	0.8806		20
	0.9450	0.8980		0.9624	0.9168		0.9602	0.9110		0.9530	0.9262		15
	0.9566	0.8915		0.9624	0.9161		0.9631	0.8835		0.9124	0.9016		10

Source – Author's Image

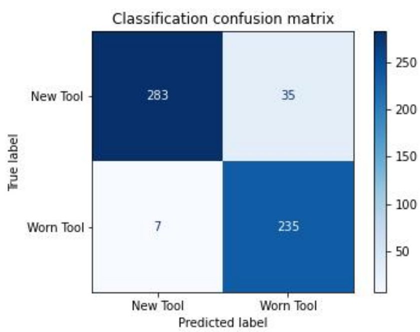
Figure 51 – Perceptron model as internal classifier.

	Perceptron as internal classification model												Number of Features
	K-Nearest Neighbour			Support Vector Machine			Logistic Regression			Perceptron			
	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	Forward	Backward	All Features	
Balanced Accuracy	0.8712	0.8734		0.9197	0.9313		0.9074	0.9110		0.9573	0.8357		150
	0.8792	0.9081		0.9457	0.9153		0.8944	0.8719		0.9501	0.8777		100
	0.9045	0.8994		0.9573	0.8234		0.8705	0.7967		0.8893	0.8488		75
	0.8640	0.8596	0.8784	0.9399	0.9284	0.8386	0.8821	0.8705	0.8864	0.8676	0.8821	0.9132	50
	0.9103	0.9030		0.9356	0.8444		0.9284	0.8104		0.9313	0.8111		25
	0.9197	0.9132		0.9356	0.8647		0.9262	0.8386		0.9197	0.8046		20
	0.9240	0.9349		0.9255	0.8300		0.9240	0.8292		0.9436	0.7308		15
	0.9182	0.9276		0.8958	0.7098		0.8958	0.7467		0.9096	0.9262		10

Source – Author's Image

Listed in the Figure 47, Figure 48, Figure 49, Figure 50 and Figure 51 are the values of the balanced accuracy of all combinations done. It also compares with the training of the same ML model with all features.

Figure 52 – Confusion Matrix K-Nearest Neighbour classifier.



Source – Author's Image

The confusion matrix showed in Figure 52 is just one of the many, each for one combination of the parameters listed in the before section, observed in the project and shows us the missed predictions the model made when training. A careful analysis of the confusion matrices was conducted to prove the results of each training ML model.

6 CONCLUSION

This document presented a complete study of the application of feature selection methods in ML models and the influence of it on the accuracy of the models for tool wear prediction in the high-precision milling process. The project covered a diverse range of topics throughout its realization.

The topics went from the planning and execution of the data production with the operation of a high precision CNC machine until the application of ML models to predict tool wear phenomena, going through the analysis of data and feature selection methods in the middle of the project, not to mention the massive preparation before the presented project steps.

The performed experiment acquired data from eight different signal sources: two AE sensors, positioned in the spindle body and workpiece clamp; two vibration sensors, placed in the workpiece clamp and spindle body; one microphone sensor, positioned inside the closed machine chamber; one encoder position sensor with X, Y, Z and Spindle signals; Spike tool holder, which recorded the force data applied in the tool; and ADS sensor, which was not part of the described project.

The number of signals processed when executing the experiments was 38, counting the signals not used to predict tool wear phenomena. Only 5 of the acquired signals were used in the ML prediction model.

Aside from the data from the ADS sensor, the remaining signals that weren't used to predict tool wear were discarded after an extensive analysis of all data acquired. Before the start of labeling the data, the process of doing it was changed due to the recognition of possible future problems.

The tool flank wear was measured in three different positions of each tooth, resulting in nine different measurements for the tool with three teeth and twelve measurements for the tool with four teeth, from which the maximum value among those was adopted as flank wear value to label the data.

From the five signals used in the prediction of tool wear, a total of 332 features were produced: 8 for each signal in the time domain, except AE sensor, and 50 for each sensor in the frequency domain, except the AE sensor which was 100 features of it.

As for feature selection, four different methods were explored in this project: two of them filter methods, more simple and easy to implement but with poor results compared to other methods; and two of them wrapper methods, more challenging to implement, produce better results but more costly for the computer.

For ML, four models were explored in this project: K-Nearest Neighbour, Support Vector Machine, Logistic Regression and Perceptron. All the parameters used in tuning the ML models were the same, preserving the models to compare the influence of the feature selection methods in the final accuracy of the prediction.

Overall, all feature selection methods could improve the models' accuracy. The most expressive accuracy improvement was in the SVM model with the Forward feature selection method using the KNN model as the internal classifier, which improved accuracy from 0.8386 to 0.9703. This is around a 15% of improvement in accuracy.

As a result, it is possible to observe that feature selection methods play an impressive role when speaking in Machine Learning and, for an implementation of a TCM system, it must be considered.

For future works, in order to better generalize the results, different parameters in the CNC process to acquire data can be explored as tool types, tool paths and others. Other feature selection methods could also be explored; LASSO Regularization is one of them, which is an embedded method. On the side of AI, a combination of feature selection methods with Deep-Learning techniques can be explored and a fine-tuning of the ML models in order to achieve the best accuracy for the model.

Furthermore, other sources of signals can also be explored, like a combination of signals which shows the force applied in the tool with the vibration of it. The results presented in this project can serve as guidelines for future studies in the field.

REFERENCES

ARENA, Industry. **DMG MORI HSC55 linear**. 2022. Available from: https://en.industryarena.com/dmgmori/products/high-speed-milling-machines-hsc--89/hsc-55-linear--123. Visited on: 23 June 2022.

BAHR, B.; MOTAVALLI, S.; ARFI, T. Sensor fusion for monitoring machine tool conditions. **International Journal of Computer Integrated Manufacturing**, Taylor Francis, v. 10, n. 5, p. 314–323, 1997. DOI: 10.1080/095119297131066. Available from: https://doi.org/10.1080/095119297131066.

BURGES, Christopher JC. A tutorial on support vector machines for pattern recognition. **Data mining and knowledge discovery**, Springer, v. 2, n. 2, p. 121–167, 1998.

BURKE, L.I. Competitive learning based approaches to tool-wear identification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 22, n. 3, p. 559–563, 1992. DOI: 10.1109/21.155957.

BYRNE, G.; DORNFELD, D.; INASAKI, I.; KETTELER, G.; KÖNIG, W.; TETI, R. Tool Condition Monitoring (TCM) — The Status of Research and Industrial Application. **CIRP Annals**, v. 44, n. 2, p. 541–567, 1995. ISSN 0007-8506. DOI: https://doi.org/10.1016/S0007-8506(07)60503-4. Available from: https://www.sciencedirect.com/science/article/pii/S0007850607605034.

CAO, Hongrui; ZHANG, Xingwu; CHEN, Xuefeng. The concept and progress of intelligent spindles: A review. **International Journal of Machine Tools and Manufacture**, v. 112, p. 21–52, 2017. ISSN 0890-6955. DOI: https://doi.org/10.1016/j.ijmachtools.2016.10.005.

CHAKRABORTY, Koyel; BHATTACHARYYA, Siddhartha; BAG, Rajib; HASSANIEN, Aboul Alla. 7 - Sentiment Analysis on a Set of Movie Reviews Using Deep Learning Techniques. In: DEY, Nilanjan; BORAH, Samarjeet; BABO, Rosalina; ASHOUR, Amira S. (Eds.). **Social Network Analytics**. [S.l.]: Academic Press, 2019. P. 127–147. ISBN 978-0-12-815458-8. DOI: https://doi.org/10.1016/B978-0-12-815458-8.00007-4. Available from: https://www.sciencedirect.com/science/article/pii/B9780128154588000074.

CHEN, J.C.; HUANG, L.H.; LAN, A.X.; LEE, S. Analysis of an effective sensing location for an in-process surface recognition system in turning operations. **Journal of Industrial Technology**, v. 15, n. 3, 1999. Available from: https://cdn.ymaws.com/atmae.site-ym.com/resource/resmgr/JIT/chen0499.pdf.

CORDER, Gregory W; FOREMAN, Dale I. **Nonparametric statistics for non-statisticians**. [S.l.]: John Wiley & Sons, Inc., 2011.

CRAMER, Jan Salomon. The origins of logistic regression. Tinbergen Institute Working Paper, 2002.

ČUŠ-UROŠ, Franci; ZUPERL, U. Real-Time Cutting Tool Condition Monitoring in Milling. **Strojniski Vestnik**, v. 57, Feb. 2011. DOI: 10.5545/sv-jme.2010.079.

DEEPAI. **Perceptron**. 2022. Available from: https://deepai.org/machine-learning-glossary-and-terms/perceptron. Visited on: 29 Sept. 2022.

DENKENA, Berend; BERGMANN, Benjamin; WITT, Matthias. Feeling Machine for Process Monitoring of Turning Hybrid Solid Components. **Metals**, v. 10, n. 7, 2020. ISSN 2075-4701. DOI: 10.3390/met10070930. Available from: https://www.mdpi.com/2075-4701/10/7/930.

DEY, S.; STORI, J.A. A Bayesian network approach to root cause diagnosis of process variations. **International Journal of Machine Tools and Manufacture**, v. 45, n. 1, p. 75–91, 2005. ISSN 0890-6955. DOI: https://doi.org/10.1016/j.ijmachtools.2004.06.018. Available from: https://www.sciencedirect.com/science/article/pii/S0890695504001609.

DIEI, E. N.; DORNFELD, D. A. Acoustic Emission from the Face Milling Process—the Effects of Process Variables. **Journal of Engineering for Industry**, v. 109, n. 2, p. 92–99, May 1987. ISSN 0022-0817. DOI: 10.1115/1.3187114. Available from: https://doi.org/10.1115/1.3187114.

DIMLA, D.E. The Correlation of Vibration Signal Features to Cutting Tool Wear in a Metal Turning Operation. **The International Journal of Advanced Manufacturing Technology**, v. 19, n. 10, p. 705–713, June 2002. ISSN 1433-3015. DOI: 10.1007/s001700200080. Available from: https://doi.org/10.1007/s001700200080.

DIMLA, Dimla E. Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. **International Journal of Machine Tools and Manufacture**, v. 40, n. 8, p. 1073–1098, 2000. ISSN 0890-6955. DOI: https://doi.org/10.1016/S0890-6955(99)00122-4. Available from: https://www.sciencedirect.com/science/article/pii/S0890695599001224.

DOLINŠEK, Slavko; KOPAČ, Janez. Mechanism and types of tool wear; particularities in advanced cutting materials. **Journal of Achievements in Materials and Manufacturing Engineering**, v. 19, n. 1, p. 11–18, 2006.

DORNFELD, David A; DEVRIES, MF. Neural network sensor fusion for tool condition monitoring. **CIRP annals**, Elsevier, v. 39, n. 1, p. 101–105, 1990.

ERTEKIN, Yalcin M.; KWON, Yongjin; TSENG, Tzu-Liang (Bill). Identification of common sensory features for the control of CNC milling operations under varying cutting conditions. **International Journal of Machine Tools and Manufacture**, v. 43, n. 9, p. 897–904, 2003. ISSN 0890-6955. DOI: https://doi.org/10.1016/S0890-6955(03)00087-7. Available from: https://www.sciencedirect.com/science/article/pii/S0890695503000877.

FRAUNHOFER. **Fraunhofer V-Box**. 2022. Available from: https://www.ipt.fraunhofer.de/de/kompetenzen/Produktionsmaschinen/praezisionstechnik-und-kunststoffreplikation/vbox.html. Visited on: 24 June 2022.

FREUND, Yoav; SCHAPIRE, Robert E. Large margin classification using the perceptron algorithm. In: PROCEEDINGS of the eleventh annual conference on Computational learning theory. [S.l.: s.n.], 1998. P. 209–217.

GUYON, Isabelle; ELISSEEFF, André. An introduction to variable and feature selection. **Journal of machine learning research**, v. 3, Mar, p. 1157–1182, 2003.

HEIDENHAIN. **Heidenhain LC483 encoder**. 2021a. Available from: https://siebert-automation.com/en/products/heidenhain-en/linear-encoders/lc-series/lc-483-en/heidenhain-lc483-ml720-id-557653-22/. Visited on: 24 June 2022.

HEIDENHAIN. **Heidenhain LC483 encoder**. 2021b. Available from: https://www.heidenhain.de/de_EN/products/linear-encoders/sealed-linear-encoders/for-numerically-controlled-machine-tools/lc-400-series/. Visited on: 23 June 2021.

HOLEX. **Holex Pro INOX Tool**. 2022a. Available from: https://www.hoffmann-group.com/GB/en/houk/Mono-machining/Solid-carbide-milling-cutters/HOLEX-Pro-INOX-solid-carbide-milling-cutter-HPC-AlCrN/p/203015-10?comingFromCategory=20&tId=561. Visited on: 28 June 2022.

HOLEX. **Holex Pro Steel Tool**. 2022b. Available from: https://www.hoffmann-group.com/GB/en/houk/Mono-machining/Solid-carbide-milling-cutters/HOLEX-Pro-Steel-solid-carbide-roughing-end-mill-HPC-TiAlN/p/202414-8?comingFromCategory=20&tId=561. Visited on: 28 June 2022.

IPT, Fraunhofer. **About Fraunhofer IPT**. 2021. Available from: https://www.ipt.fraunhofer.de/en/Profile.html. Visited on: 5 May 2022.

JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert. **An introduction to statistical learning**. [S.l.]: Springer, 2013. v. 112.

JANG, Dong Young; CHOI, Young-Gu; KIM, Hong-Gil; HSIAO, Alex. Study of the correlation between surface roughness and cutting vibrations to develop an on-line roughness measuring technique in hard turning. **International Journal of Machine Tools and Manufacture**, v. 36, n. 4, p. 453–464, 1996. ISSN 0890-6955. DOI: https://doi.org/10.1016/0890-6955(95)00074-7. Available from: https://www.sciencedirect.com/science/article/pii/0890695595000747.

KAELBLING, Leslie Pack; LITTMAN, Michael L; MOORE, Andrew W. Reinforcement learning: A survey. **Journal of artificial intelligence research**, v. 4, p. 237–285, 1996.

KANNATEY-ASIBU, Elijah; YUM, Juil; KIM, T.H. Monitoring tool wear using classifier fusion. **Mechanical Systems and Signal Processing**, v. 85, p. 651–661, 2017. ISSN 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2016.08.035. Available from: https://www.sciencedirect.com/science/article/pii/S0888327016303193.

KEYENCE. **Keyence VHX-500F**. 2022. Available from: https://www.keyence.com/products/microscope/digital-microscope/vhx-500f/models/vhx-500fe/. Visited on: 24 June 2022.

KOLLER, Daphne; SAHAMI, Mehran. **Toward optimal feature selection**. [S.l.], 1996.

KRAMER, Mark A. Nonlinear principal component analysis using autoassociative neural networks. **AIChE journal**, Wiley Online Library, v. 37, n. 2, p. 233–243, 1991.

KRATSIOS, Anastasis; HYNDMAN, Cody. Neu: A meta-algorithm for universal uap-invariant feature representation. **Journal of Machine Learning Research**, JMLR, v. 22, p. 92, 2021.

KUHN, Max; JOHNSON, Kjell, et al. **Applied predictive modeling**. [S.l.]: Springer, 2013. v. 26.

KURADA, S.; BRADLEY, C. A machine vision system for tool wear assessment. **Tribology International**, v. 30, n. 4, p. 295–304, 1997. ISSN 0301-679X. DOI: `https://doi.org/10.1016/S0301-679X(96)00058-8`. Available from: `https://www.sciencedirect.com/science/article/pii/S0301679X96000588`.

LEE, J.H; KIM, D.E; LEE, S.J. Statistical analysis of cutting force ratios for flank-wear monitoring. **Journal of Materials Processing Technology**, v. 74, n. 1, p. 104–114, 1998. ISSN 0924-0136. DOI: `https://doi.org/10.1016/S0924-0136(97)00256-2`.

LEHMAN, Ann. **Jmp For Basic Univariate And Multivariate Statistics: A Step-by-step Guide**. [S.l.]: Cary, NC : SAS Press, 2005.

LI, G.S.; LAU, W.S.; ZHANG, Y.Z. In-process drill wear and breakage monitoring for a machining centre based on cutting force parameters. **International Journal of Machine Tools and Manufacture**, v. 32, n. 6, p. 855–867, 1992. ISSN 0890-6955. DOI: `https://doi.org/10.1016/0890-6955(92)90035-F`.

LIU, Huan. Feature Selection. In: **Encyclopedia of Machine Learning**. Ed. by Claude Sammut and Geoffrey I. Webb. Boston, MA: Springer US, 2010. P. 402–406. ISBN 978-0-387-30164-8. DOI: `10.1007/978-0-387-30164-8_306`. Available from: `https://doi.org/10.1007/978-0-387-30164-8_306`.

MORI, DMG. **DMG MORI HSC55 linear manual**. 2022. Available from: `https://www.exfactory.com/Literature/DMG/HSC75-PDFa.pdf`. Visited on: 23 June 2022.

MYERS, Jerome L; WELL, Arnold D; LORCH, Robert F. **Research design and statistical analysis**. [S.l.]: Routledge, 2013.

O'DONNELL, Garret; YOUNG, Paul; KELLY, Kevin; BYRNE, Gerry. Towards the improvement of tool condition monitoring systems in the manufacturing environment. **Journal of Materials Processing Technology**, v. 119, n. 1, p. 133–139, 2001. PART 2: INTERNATIONAL CONFERENCE ON ADVANCES IN MATERIALS PROCESSING TECHNOLOGY. ISSN 0924-0136. DOI: `https://doi.org/10.1016/S0924-0136(01)00928-1`.

PALANISAMY, P.; RAJENDRAN, I.; SHANMUGASUNDARAM, S. Prediction of tool wear using regression and ANN models in end-milling operation. **The International Journal of Advanced Manufacturing Technology**, v. 37, n. 1, p. 29–41, 2008. ISSN 1433-3015. DOI: `https://doi.org/10.1007/s00170-007-0948-5`.

PARK, Kun Il; PARK, M. **Fundamentals of probability and stochastic processes with applications to communications**. [S.l.]: Springer, 2018.

PETERSON, Leif E. K-nearest neighbor. **Scholarpedia**, v. 4, n. 2, p. 1883, 2009.

PHUONG, Tu Minh; LIN, Zhen; ALTMAN, Russ B. Choosing SNPs using feature selection. In: IEEE. 2005 IEEE Computational Systems Bioinformatics Conference (CSB'05). [S.l.: s.n.], 2005. P. 301–309.

PIEZOELETRONICS, PCB. **PCB 130A24 Microphone**. 2022. Available from: `https://www.pcb.com/products?m=130a24`. Visited on: 24 June 2022.

PIEZOTRONICS, PCB. **PCB 356A15 Accelerometer**. 2022a. Available from: `https://www.pcb.com/products?m=356a15`. Visited on: 24 June 2022.

PIEZOTRONICS, PCB. **PCB 356B21 Accelerometer**. 2022b. Available from: `https://www.pcb.com/products?model=356b21`. Visited on: 24 June 2022.

PRO-MICRON. **Pro-Micron Spike Tool holder**. 2022a. Available from: `https://fccid.io/2AGNQ-STH24AA01A1/User-Manual/User-manual-2836341`. Visited on: 24 June 2022.

PRO-MICRON. **Pro-Micron Spike Tool holder**. 2022b. Available from: `https://www.pro-micron.de/spike/?lang=en`. Visited on: 24 June 2022.

RANGWALA, S.; DORNFELD, D. Sensor Integration Using Neural Networks for Intelligent Tool Condition Monitoring. **Journal of Engineering for Industry**, v. 112, n. 3, p. 219–228, Aug. 1990. ISSN 0022-0817. DOI: 10.1115/1.2899578. Available from: https://doi.org/10.1115/1.2899578.

REHORN, Adam G.; JIANG, Jin; ORBAN, Peter E. State-of-the-art methods and results in tool condition monitoring: a review. **The International Journal of Advanced Manufacturing Technology**, v. 26, n. 7, p. 693–710, 2005. ISSN 1433-3015. DOI: https://doi.org/10.1007/s00170-004-2038-2.

REJER, Izabela; LORENZ, Krzysztof. Genetic algorithm and forward method for feature selection in EEG feature space. **Journal of Theoretical and Applied Computer Science**, v. 7, n. 2, p. 72–82, 2013.

RUSSELL, Stuart J. **Artificial intelligence a modern approach**. [S.l.]: Pearson Education, Inc., 2010.

SADÍLEK, Marek; KRATOCHVÍL, Jiří; PETRU, Jana; ČEP, Robert; ZLAMAL, Tomas; STANČEKOVÁ, Dana. Cutting tool wear monitoring with the use of impedance layers. **Tehnicki Vjesnik**, v. 21, p. 639–644, June 2014.

SAMUEL, Arthur L. Some Studies in Machine Learning Using the Game of Checkers. I. In: **Computer Games I**. Ed. by David N. L. Levy. New York, NY: Springer New York, 1988. P. 335–365. ISBN 978-1-4613-8716-9. DOI: 10.1007/978-1-4613-8716-9_14. Available from: https://doi.org/10.1007/978-1-4613-8716-9_14.

SCHEFFER, C; HEYNS, PS. An industrial tool wear monitoring system for interrupted turning. **Mechanical Systems and Signal Processing**, Elsevier, v. 18, n. 5, p. 1219–1242, 2004.

SCIENCE, Towards data. **Types of SVM kernels**. 2022. Available from: https://towardsdatascience.com/multiclass-classification-with-support-vector-machines-svm-kernel-trick-kernel-functions-f9d5377d6f02. Visited on: 17 Aug. 2022.

SINGH, Nanua. **Systems Approach to Computer-Integrated Design and Manufacturing**. New York: John Wiley Sons., 1996.

SUBASI, Abdulhamit. Chapter 3 - Machine learning techniques. In: SUBASI, Abdulhamit (Ed.). **Practical Machine Learning for Data Analysis Using Python**. [S.l.]: Academic Press, 2020. P. 91–202. ISBN 978-0-12-821379-7. DOI: https://doi.org/10.1016/B978-0-12-821379-7.00003-5. Available from: https://www.sciencedirect.com/science/article/pii/B9780128213797000035.

AL-SULAIMAN, Faleh; ABDUL BASEER, Mohammed; SHEIKH, Anwar Khalil. Use of electrical power for online monitoring of tool condition. **Journal of Materials Processing Technology**, v. 166, p. 364–371, Aug. 2005. DOI: 10.1016/j.jmatprotec.2004.07.104.

VALLEN. **VS150-K3 Acoustic Emission**. 2022. Available from: https://www.vallen.de/sensors/watertight-sensors/vs150-k3/. Visited on: 24 June 2022.

VAPNIK, Vladimir. **The nature of statistical learning theory**. [S.l.]: Springer science & business media, 1999.

VAPNIK, Vladimir N. An overview of statistical learning theory. **IEEE transactions on neural networks**, IEEE, v. 10, n. 5, p. 988–999, 1999.

WANG, Jinjiang; XIE, Junyao; ZHAO, Rui; ZHANG, Laibin; DUAN, Lixiang. Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing. **Robotics and Computer-Integrated Manufacturing**, v. 45, p. 47–58, 2017. Special Issue on Ubiquitous Manufacturing (UbiM). ISSN 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2016.05.010. Available from: https://www.sciencedirect.com/science/article/pii/S0736584516301855.

WEISSTEIN, Eric. **Statistical Correlation.** 2022. Available from: https://mathworld.wolfram.com/StatisticalCorrelation.html. Visited on: 15 June 2022.

WILCOX, S.J.; REUBEN, R.L.; SOUQUET, P. The use of cutting force and acoustic emission signals for the monitoring of tool insert geometry during rough face milling. **International Journal of Machine Tools and Manufacture**, v. 37, n. 4, p. 481–494, 1997. ISSN 0890-6955. DOI: https://doi.org/10.1016/S0890-6955(96)00069-7. Available from: https://www.sciencedirect.com/science/article/pii/S0890695596000697.

YAN, Jihong; MENG, Yue; LU, Lei; LI, Lin. Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance. **IEEE Access**, v. 5, p. 23484–23491, 2017. DOI: 10.1109/ACCESS.2017.2765544.

YU, Lei; LIU, Huan. Feature selection for high-dimensional data: A fast correlation-based filter solution. In: PROCEEDINGS of the 20th international conference on machine learning (ICML-03). [S.l.: s.n.], 2003. P. 856–863.

ZHOU, Yuqing; XUE, Wei. Review of tool condition monitoring methods in milling processes. **The International Journal of Advanced Manufacturing Technology**, v. 96, n. 5, p. 2509–2523, 2018. ISSN 1433-3015. DOI: https://doi.org/10.1007/s00170-018-1768-5.

AL-ZUBAIDI, Salah; GHANI, Jaharah A.; CHE HARON, Che Hassan. Application of ANN in Milling Process: A Review. **Modelling and Simulation in Engineering**, Hindawi Publishing Corporation, v. 2011, p. 696275, May 2011. ISSN 1687-5591. DOI: 10.1155/2011/696275. Available from: https://doi.org/10.1155/2011/696275.