



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO E SISTEMAS
CURSO ENGENHARIA DE PRODUÇÃO MECÂNICA

João Vitor Vanzella dos Santos

**Inferência dos tempos-padrão de operações de inspeção em um laboratório de
metrologia**

Florianópolis

2025

João Vitor Vanzella dos Santos

Inferência dos tempos-padrão de operações de inspeção em um laboratório de metrologia

Trabalho de Conclusão de Curso submetido ao curso de Engenharia de Produção Mecânica do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel Engenharia de Produção Mecânica.

Orientador(a): Prof. Dr. Daniel Pacheco Lacerda

Florianópolis

2025

dos Santos, João Vitor Vanzella

Inferência dos tempos-padrão de operações de inspeção
em um laboratório de metrologia / João Vitor Vanzella dos
Santos ; orientador, Daniel Pacheco Lacerda, 2025.

108 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro
Tecnológico, Graduação em Engenharia de Produção Mecânica,
Florianópolis, 2025.

Inclui referências.

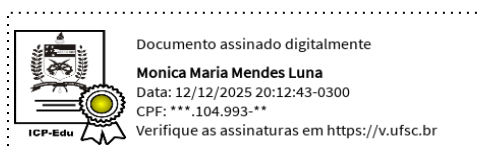
1. Engenharia de Produção Mecânica. 2. Tempo-padrão. 3.
Machine Learning. 4. Metrologia. 5. Variabilidade de
Processo. I. Lacerda, Daniel Pacheco. II. Universidade
Federal de Santa Catarina. Graduação em Engenharia de
Produção Mecânica. III. Título.

João Vitor Vanzella dos Santos

Inferência dos tempos-padrão de operações de inspeção em um laboratório de metrologia

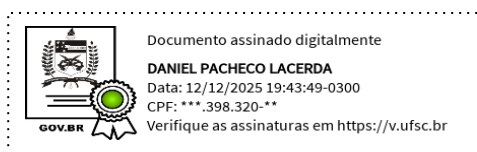
Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Bacharel em Engenharia de Produção Mecânica e aprovado em sua forma final pelo Curso Engenharia de Produção Mecânica.

Florianópolis, 27 de outubro de 2025.



Coordenação do Curso

Banca examinadora



Dr. Daniel Pacheco Lacerda

Orientador

Insira neste espaço
a assinatura

Prof.(a) [nome do(a) professor(a)], Dr.(a)

Instituição [nome da instituição]

Insira neste espaço
a assinatura

Prof.(a) [nome do(a) professor(a)], Dr.(a)

Instituição [nome da instituição]

Florianópolis, 2025.

Aos meus familiares, pelo apoio constante e por cada gesto de incentivo. Aos amigos feitos na faculdade, em especial os do GLean e da Ampera, pela parceria ao longo desta jornada. Aos amigos de longa data — Marco, Rodrigo, Marcelo, Roberto, Thiago, Matheus, Eduardo, Binckas e Rhenzo —, minha sincera gratidão.

AGRADECIMENTOS

Agradeço ao Prof. Daniel Lacerda, meu orientador, pela orientação rigorosa e pelas contribuições que elevaram a qualidade deste TCC. Ao Prof. Leandro Gauss, pelos comentários e indicações que ampliaram o escopo da pesquisa. E a todos do setor de Metrologia da Empresa Z, pela solicitude e pelo suporte técnico indispensável ao longo do trabalho.

RESUMO

Este trabalho desenvolve e valida um modelo de aprendizado de máquina para estimar, com precisão operacionalmente relevante, os tempos-padrão de inspeção de um laboratório de metrologia. Para isso, o estudo segue um fluxo de etapas reprodutível, envolvendo curadoria de dados, análise exploratória por correlação, seleção de preditores por abordagem *wrapper*, comparação sistemática de algoritmos e escolha do modelo para estimar o tempo normal, convertido em tempo-padrão com tolerâncias por máquina. Os tempos estimados são estruturados por recurso e por faixas de complexidade (combinação entre máquina e número de características medidas) em três patamares definidos como percentis de tempo (P50, P75, P95), que podem ser adotados como níveis de serviço conforme o risco de atraso aceitável. Uma análise complementar de regras de associação identifica combinações de máquina, tipo de inspeção e condições do processo mais associadas a alongamentos de tempo, apoiando ações de melhoria e o sequenciamento de inspeções mais críticas. Os resultados demonstram que é possível estimar tempos-padrão diretamente utilizáveis no planejamento de capacidade, no sequenciamento e na gestão de filas, inclusive em ambientes com medição automática limitada, oferecendo um método auditável e transferível para outras operações de apoio à produção.

Palavras-chave: Tempo-padrão; *Machine Learning*; Metrologia; Variabilidade de processo.

ABSTRACT

This study develops and validates a machine learning model to estimate inspection standard times for a metrology laboratory with operationally relevant accuracy. To this end, it follows a reproducible sequence of steps involving data curation, correlation-based exploratory analysis, predictor selection via a wrapper approach, systematic comparison of algorithms, and selection of the model used to estimate normal time, which is then converted into standard time with machine-specific allowances. The estimated times are organized by resource and by complexity bands (combining machine and number of characteristics measured) into three levels defined as time percentiles (P50, P75, P95). These can be adopted as service levels according to the acceptable risk of delay. A complementary association-rule analysis identifies combinations of machine, inspection type, and process conditions that are more strongly associated with time overruns, supporting improvement actions and the sequencing of more critical inspections. The results show that it is possible to estimate standard times that are directly usable for capacity planning, sequencing, and queue management, even in environments with limited automatic measurement, providing an auditable method that can be transferred to other production support operations.

Keywords: Standard time; Machine learning; Metrology; Process variability.

LISTA DE FIGURAS

Figura 1 - Evolução mensal da demanda da atividade de “ReExame”	17
Figura 2 - Evolução semanal da demanda da atividade de “ReExame”	18
Figura 3 – Formação do <i>corpus</i> de pesquisa preliminar	22
Figura 4 – Método de pesquisa.....	40
Figura 5 – Classificação dos calibradores antes do estudo	43
Figura 6 – Método de trabalho	44
Figura 7 – Influência de <i>setup</i> extra e necessidade de programação no tempo de <i>setup</i>	48
Figura 8 – Fluxo de inspeção	50
Figura 9 – Distribuição dos tempos para cada estação de inspeção	51
Figura 10 – Lógica de aprendizagem supervisionada	53
Figura 11 – Abordagem <i>wrapper</i> para seleção de subconjunto de preditores	54
Figura 12 – Percurso analítico.....	64
Figura 13 – <i>Heatmap</i> de correlações	66
Figura 14 – Previsto X Real para cada algoritmo de ML.....	69
Figura 15 – Proposta de Classificação dos Calibradores	79

LISTA DE QUADROS

Quadro 1 – Termos pesquisados em bases de dados	21
Quadro 2 – <i>Corpus</i> de pesquisa preliminar	22
Quadro 3 – Técnicas de obtenção de tempo padrão	30
Quadro 4 – Estratégias para ambientes com alta variabilidade	32
Quadro 5 – Regras de associação	34
Quadro 6 – Algoritmos de ML	38
Quadro 7 – Características da classificação anterior ao estudo	43
Quadro 8 - Processos do setor de metrologia	46
Quadro 9 – Descrição dos elementos de tempo.....	47
Quadro 10 – Síntese dos fatores para Cálculo das Tolerâncias	63
Quadro 11 – Questões para avaliação qualitativa.....	65
Quadro 12 – Preditores selecionados pelo método <i>wrapper</i> ($R^2 = 0,455$)	67

LISTA DE TABELAS

Tabela 1 - Dados Coletados.....	49
Tabela 2 – Fatores de esforço físico e mental	59
Tabela 3 – Fatores de recuperação	60
Tabela 4 – Fatores de monotonia.....	60
Tabela 5 – Fatores para condições ambientais	61
Tabela 6 – Resultados para cada algoritmo de ML	68
Tabela 7 – Variáveis Prevalentes dos Tempos-Padrão.....	70
Tabela 8 – Tempos-padrão médios por máquina.....	71
Tabela 9 – Tempos-padrão por máquina e características medidas	72
Tabela 10 – Fatores que elevam o tempo para o percentil 50	74
Tabela 11 – Fatores que elevam o tempo para o percentil 75	75
Tabela 12 – Fatores que elevam o tempo para o percentil 95	75

LISTA DE ABREVIATURAS E SIGLAS

ANN	<i>Artificial Neural Networks</i>
BSC	<i>Balanced Scorecard</i>
CART	<i>Classification and Regression Trees</i> (Árvores de Decisão)
CMM	<i>Coordinate Measuring Machine</i>
DHM	Modelagem Humana Digital
EDA	<i>Exploratory Data Analysis</i>
ERP	Planejamento de Recursos Empresariais
FIFO	<i>First In First Out</i>
FTE	<i>Full-Time Equivalents</i>
HMLV	<i>High-Mix, Low-Volume</i>
KDD	<i>Knowledge Discovery in Databases</i>
kNN	<i>k-Nearest Neighbors</i>
MES	<i>Manufacturing Execution System</i>
ML	<i>Machine Learning</i>
MLP	<i>Multilayer Perceptron</i>
MOST	<i>Maynard Operation Sequence Technique</i>
MP	Problema na Máquina
MTM	<i>Methods-Time Measurement</i>
NC	Não Conformidade
PCP	Planejamento e Controle da Produção
PN	<i>Part Number</i>
PP	Programa Pronto
PHH	Peça por Hora-Homem
PTS	Sistemas de Tempo Predeterminados
RMSE	<i>Root Mean Squared Error</i>
RSL	Revisão Sistemática da Literatura
SJR	<i>SCImago Journal Rank</i>
SLA	<i>Service Level Agreement</i>
SMED	<i>Single Minute Exchange of Die</i>
SVR	<i>Support Vector Regression</i>
SVM	<i>Support Vector Machines</i>
TAT	<i>Turnaround Time</i>

TPM	<i>Total Productive Maintenance</i>
WIP	Trabalho em Andamento

SUMÁRIO

1	INTRODUÇÃO	15
1.1	OBJETO E PROBLEMA DE PESQUISA.....	16
1.2	OBJETIVOS	20
1.3	JUSTIFICATIVA	20
1.4	ESTRUTURA DO TRABALHO	25
2	FUNDAMENTAÇÃO TEÓRICA.....	26
2.1	TEMPO-PADRÃO: CONCEITOS, TÉCNICAS E DESAFIOS	26
2.2	MODELAGEM ESTATÍSTICA E MACHINE LEARNING PARA A ESTIMATIVA DOS TEMPOS-PADRÃO	32
3	PROCEDIMENTOS METODOLÓGICOS.....	40
3.1	UNIDADE CONTEXTO E DE ANÁLISE.....	41
3.2	MÉTODO DE TRABALHO	43
3.3	COLETA DE DADOS	46
3.4	DEFINIÇÃO DO ESCOPO DA ANÁLISE.....	49
3.5	ANÁLISE DE DADOS	52
4	ANÁLISE DOS RESULTADOS	66
4.1	AVALIAÇÃO DOS ALGORITMOS DE MACHINE LEARNING.....	66
4.2	INFERÊNCIA DOS TEMPOS-PADRÃO	71
4.3	FATORES EXPLICATIVOS DO TEMPO DE INSPEÇÃO.....	74
5	DISCUSSÃO DOS RESULTADOS	78
6	CONCLUSÕES.....	84
	REFERÊNCIAS.....	87

1 INTRODUÇÃO

No âmbito da engenharia de métodos, os estudos de tempos e movimentos permitem dissecar as tarefas de um processo, identificando oportunidades para obter eficiência por meio da eliminação de desperdícios e de movimentos desnecessários (Gomes Ribeiro *et al.*, 2019). No início do século XX, Taylor enfatizou três princípios básicos para racionalizar o desempenho: definir o que fazer (tarefa padronizada), quanto tempo leva para fazer (tempo determinado via cronometragem) e como fazer (método ideal) (Breznik *et al.*, 2023; Çaklt; Dağdeviren, 2023). A integração entre o estudo de movimentos, focado na eficiência do método de trabalho, e a medição de tempos permite quantificar o desempenho. Essa quantificação é fundamental para avaliar e elevar a produtividade em ambientes industriais e de serviços (Çaklt; Dağdeviren, 2023).

Um conceito central derivado desses estudos é o tempo-padrão. Em termos formais, parte-se do tempo normal, entendido como o tempo de execução de uma tarefa por um operador qualificado, seguindo o método especificado em ritmo padrão, e acrescentam-se tolerâncias para necessidades pessoais, fadiga e atrasos inevitáveis (Teshome; Meles; Yang, 2024). Do ponto de vista gerencial, tempos-padrão precisos são importantes para planejamento e controle, pois permitem: i) elaborar cronogramas realistas; ii) balancear linhas de produção; iii) calcular capacidades; iv) estimar custos; e v) avaliar a produtividade objetivamente (Çaklt; Dağdeviren, 2023).

Apesar da importância da precisão dos tempos-padrão, a sua determinação não é trivial. Tradicionalmente, a obtenção desses tempos requer a realização de estudos por meio de técnicas de medição direta (como a cronometragem com cronômetro e a filmagem) ou medição indireta (como sistemas de tempos pré-determinados ou análises comparativas) conduzidas por analistas especializados. Essas abordagens apresentam desafios práticos, por exemplo, a cronometragem e observação diretas podem ser custosas e demoradas, exigindo um ambiente controlado que nem sempre está disponível, além de dependerem intensivamente da experiência e do rigor de quem realiza o estudo (Çaklt; Dağdeviren, 2023). Em muitos casos, especialmente em operações complexas ou em empresas de pequeno porte, os métodos tradicionais de medição do trabalho mostram-se inadequados ou impraticáveis para garantir um tempo-padrão preciso, por muitas vezes desconsiderarem fatores que influenciam os tempos de ciclo, como tipos de máquina, habilidades dos funcionários, material e tipo de produto (Çaklt; Dağdeviren, 2023; Onaran; Yanik, 2020). Consequentemente, é comum que organizações recorram a estimativas

ou julgamentos empíricos e subjetivos para estipular a duração de suas atividades (Urban; Koblasa; Mendřický, 2024).

Diante desse panorama, não se trata de eliminar a necessidade de medir tempos, mas de reduzir de forma significativa o volume de medições manuais exigido. Em vez de conduzir novos estudos de tempos para cada variação de processo, é possível realizar uma campanha inicial de medições em uma fração das operações e, a partir desses dados históricos, treinar modelos de aprendizado de máquina para estimar tempos-padrão de tarefas similares. Nesse contexto, técnicas de *Machine Learning* (ML) supervisionado passam a complementar a cronoanálise tradicional: o modelo aprende a relação entre atributos das operações e o tempo gasto, permitindo generalizar essas estimativas para combinações ainda não medidas. Assim, parte do esforço de estudo de tempos é substituída por inferência preditiva, reduzindo o trabalho de campo sem abrir mão de uma base empírica. Estudos recentes corroboram o potencial dessa abordagem, ao mostrar que algoritmos de aprendizado de máquina podem estimar tempos-padrão com boa precisão e apoiar ganhos de produtividade por meio de um planejamento mais acurado dos processos (Çaklı; Dağdeviren, 2023).

Sendo assim, esta pesquisa se insere na temática do estudo dos tempos e movimentos em áreas de apoio para a melhoria da eficiência operacional. Na seção a seguir, detalham-se o objeto de estudo e a problemática de pesquisa associada.

1.1 OBJETO E PROBLEMA DE PESQUISA

O gerenciamento eficaz das organizações, seja na esfera industrial ou de serviços, depende intrinsecamente da capacidade de mensurar o desempenho de seus processos (Kaplan; Norton, 1992). A ausência de métricas confiáveis e precisas compromete a tomada de decisões, tornando o direcionamento estratégico vulnerável a vieses e suposições (Kaplan; Norton, 1992). Logo, um sistema de indicadores de desempenho, ou *Key Performance Indicators* (KPIs), é o elo entre o estado atual dos processos e as escolhas sobre como alocar recursos de maneira eficiente (Neely; Gregory; Platts, 1995; Zhu *et al.*, 2017). A seleção e o monitoramento dos KPIs corretos permitem que as empresas identifiquem fraquezas operacionais e capitalize em oportunidades de melhoria que, de outra forma, permaneceriam invisíveis (Zhu *et al.*, 2017).

Contudo, a aplicação consistente de métricas de desempenho enfrenta um desafio notório em funções indiretas, estudos recentes indicam que essas funções tendem a ser marginalizadas nos programas de melhoria contínua, com a atenção gerencial frequentemente voltada para os processos produtivos diretos (Altmeier *et al.*, 2023). Essa marginalização ocorre

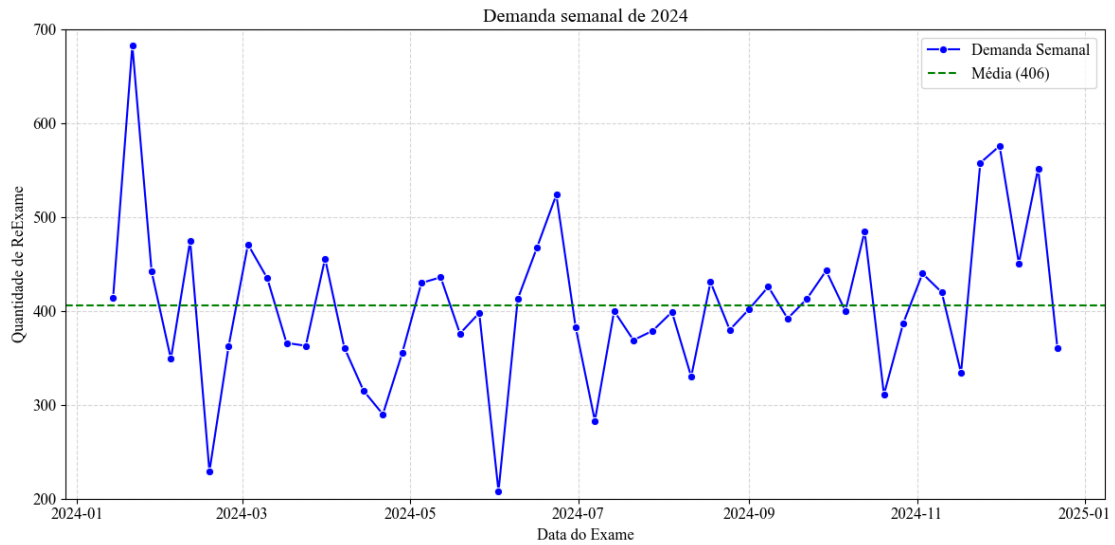
não por sua falta de importância, mas pela maior complexidade de mensuração e pela reduzida padronização, que dificulta a realização de comparações e *benchmarks* (Altmeier *et al.*, 2023). Paradoxalmente, essas funções de apoio consomem uma parcela expressiva dos recursos operacionais de uma empresa, representando cerca de 8% a 12% dos custos totais de operações e ocupando entre 30% a 35% do total de equivalentes de tempo integral (*Full-Time Equivalents - FTEs*) (Altmeier *et al.*, 2023).

Esse problema se manifesta no laboratório de metrologia da Empresa Z, *locus* deste trabalho. Operando em um ambiente *high-mix/low-volume* (HMLV), o setor de metrologia lida com uma ampla diversidade de calibradores, cada qual com requisitos e procedimentos de inspeção distintos. Agrava-se esse quadro pelo crescimento (Figura 1) e pela variabilidade da demanda de inspeções observada no período (Figura 2), pressionando a capacidade, a programação e os prazos. Esse cenário dificulta, torna imprecisa a determinação de tempos-padrão e, utilizando os métodos tradicionais, economicamente inviável.



Fonte: Elaborado pelo autor

Figura 2 - Evolução semanal da demanda da atividade de “ReExame”



Fonte: Elaborado pelo autor

Cabe registrar que as Figura 1 e Figura 2 consideram exclusivamente a atividade de ReExame, única com recorrência diária e registros sistemáticos no período analisado. Na prática, o setor de metrologia também absorve outras demandas de apoio à linha, que são não programadas, heterogêneas e competem pelos mesmos recursos. Essas entradas provocam trocas de prioridade, *setups* adicionais e interrupções, ampliando significativamente a variabilidade de tempos e a dispersão de carga. Assim, as variações observadas tendem a subestimar a variabilidade total do sistema.

A abordagem convencional de medição do trabalho, como a cronometragem manual, embora fundamental para os estudos de tempos e movimentos, apresenta desafios práticos expressivos neste contexto. A cronometragem direta é custosa, demorada e exige um ambiente controlado, e depende fortemente da experiência e do rigor do analista. Com mais de 10.000 *Part Numbers* (PNs) e diferentes procedimentos de medição, o engenheiro de tempos e métodos da Empresa Z estima que um estudo de tempos pelos métodos tradicionais exigiria cerca de um ano de trabalho. Diante deste esforço, a organização precisa recorrer a "estimativas ou julgamentos empíricos" para estipular a duração de suas atividades (Urban; Koblasa; Mendřický, 2024), resultando em métricas de produtividade imprecisas e subjetivas.

A avaliação de desempenho atual da Empresa Z baseia-se na contagem de calibrações, desconsiderando as diferenças de complexidade entre os serviços e, consequentemente, distorce o balanceamento da carga de trabalho. Na prática, a falta de tempos-padrão impede que o laboratório de metrologia conheça com clareza a própria capacidade de atendimento. Sem referências temporais por máquina e por tipo de inspeção, o planejamento e controle da

produção (PCP) não consegue traduzir a carteira de calibrações em carga de trabalho, nem estimar com precisão quantos inspetores são necessários por turno ou qual o impacto de oscilações de demanda. Os indicadores por recurso tornam-se enviesados, pois comparam inspetores e máquinas apenas pelo número de calibrações concluídas, ignorando diferenças de complexidade entre serviços e rotas de inspeção. Isso dificulta identificar gargalos reais, priorizar melhorias e avaliar o efeito de investimentos ou mudanças de método, já que não há uma base temporal para comparar “antes e depois”. Além disso, prazos prometidos à produção e às áreas clientes passam a depender de julgamentos empíricos, aumentando a probabilidade de atrasos, retrabalhos de programação e conflitos entre áreas quando a demanda supera a capacidade efetiva do laboratório.

Além de estabelecer tempos-padrão para as inspeções, é importante compreender quais fatores contribuem para a sua variação. Estudos em metrologia coordenada mostram que escolhas de estratégia de medição — por exemplo, quantidade de pontos amostrados, percurso da sonda e velocidade de varredura — influenciam simultaneamente o tempo de medição e a produtividade do processo, reforçando que o tempo não depende apenas do volume inspecionado, mas também de como a tarefa é executada (Magdziak, 2022).

Em um contexto amplo, laboratórios de metrologia da indústria automotiva realizam dezenas de milhares de calibrações por ano, segundo o *National Institute of Standards and Technology* (State Laboratory Program Workload Survey, 2022). Por um lado, estudos científicos em células de inspeção por CMMs (*Coordinate Measuring Machines*) evidenciam que a falta de padronização nos métodos de medição aumenta a variabilidade dos resultados, consome parte significativa do tempo útil e gera gargalos de capacidade (Aggogeri; Torino; Levi, 2008). Por outro lado, pesquisas industriais mostram que a racionalização e padronização da sequência de pontos de medição pode reduzir o ciclo de inspeção em até 12 % (Salman *et al.*, 2016).

A implementação de modelos de ML não está isenta de desafios. A qualidade e quantidade dos dados para treino são fatores críticos para o sucesso de qualquer modelo preditivo. Em ambientes de fabricação com poucos dados, é comum basear-se em estimativas empíricas ou aproximações grosseiras, em função da dificuldade de coletar o volume de dados necessário para o ML (Urban; Koblasa; Mendřický, 2024). No caso da Empresa Z, além de não haver sensoramento nas máquinas do setor de metrologia para registrar automaticamente tempos de ciclo, não há registros de algumas atividades realizadas no setor nem carimbos de tempo padronizados (início/fim, *setup*, medição e espera), além de ampla heterogeneidade entre

famílias e programas de medição. Esses fatores dificultam a obtenção de dados, e aumentam ruído e viés de amostragem.

Diante desse contexto, o problema de pesquisa central é a ausência de tempos-padrão confiáveis nas rotinas de inspeção do departamento de metrologia da Empresa Z. Essa lacuna torna incerto o balanceamento da carga de trabalho, eleva custos, dificulta a visualização dos resultados de melhoria e aumenta os riscos de atrasos.

Em decorrência, a questão de pesquisa que orienta este trabalho é: Como desenvolver e um modelo de *Machine Learning* que estime, com precisão operacionalmente relevante, os tempos-padrão de inspeção no processo de ReExame do laboratório de metrologia da Empresa Z?

1.2 OBJETIVOS

O objetivo geral deste trabalho é estimar os tempos-padrão de inspeção para avaliação da produtividade a partir de *machine learning*. Além disso, estas são as contribuições colaterais do trabalho:

- a) Analisar comparativamente as aplicações de *machine learning* para a estimação dos tempos-padrão em setores de apoio;
- b) Classificar os calibradores em grupos por semelhança;
- c) Averiguar os principais fatores que levam a um aumento no tempo de medição dos calibradores;
- d) Propor um modelo para averiguação dos tempos-padrão em setores de apoio à produção;

1.3 JUSTIFICATIVA

Nesta seção apresentam-se as justificativas para a realização deste estudo, organizadas em duas frentes: acadêmica e empresarial. Para fundamentar este estudo, realizou-se uma revisão sistemática da literatura (RSL). Esse procedimento permite mapear, sintetizar e avaliar criticamente as evidências disponíveis sobre o tema, consolidando o conhecimento acumulado e revelando limitações e lacunas que orientam novas investigações (Ermel, 2021).

O protocolo adotado contemplou a definição de bases de dados relevantes em engenharia e manufatura, a formulação de descritores com operadores booleanos e a aplicação de filtros de escopo. Priorizaram-se publicações acadêmicas com foco em aplicações de

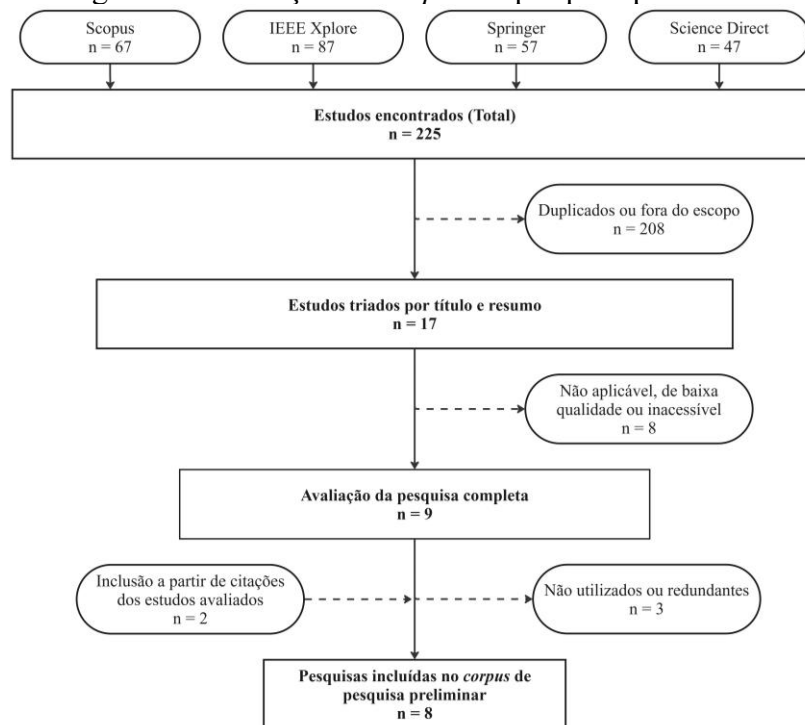
aprendizado de máquina relacionadas à estimativa de tempos-padrão, preferencialmente relacionados a inspeções de qualidade/metrologia.

O Quadro 1 resume as estratégias de busca e os resultados por base, com adaptações necessárias à sintaxe de cada plataforma. No conjunto, foram identificados 225 registros. A seguir, procedeu-se à triagem em três etapas: i) remoção de duplicatas e leitura de títulos e resumos para verificar aderência ao escopo; ii) leitura integral para confirmação de qualidade e elegibilidade; iii) Por fim, eliminou-se artigos redundantes ou aqueles que não foram utilizados arbitrariamente. A Figura 3 apresenta o fluxo de formação do corpus preliminar, conforme orientações metodológicas de (Ermel, 2021).

Quadro 1 – Termos pesquisados em bases de dados

Base de dados	Termos de busca	Documentos encontrados
SCOPUS	TITLE-ABS-KEY ("machine learning" AND ("standard time" OR "time standard" OR "time study")) AND (LIMIT-TO (SUBJAREA , "ENGI"))	67
IEE XPLORE	"machine learning" AND ("standard time" OR "time standard" OR "time estimation") AND ("productivity" OR "manufacturing" OR "quality")	87
Springer	"machine learning" AND ("standard time" OR "time standard" OR "time estimation") AND ("productivity" OR "manufacturing") AND ("metrology" OR "quality inspection")	57
Science Direct	"machine learning" AND ("standard time" OR "time standard" OR "time estimation") AND ("productivity" OR "manufacturing") AND ("metrology" OR "quality inspection")	47
Total de documentos		225

Fonte: elaborado pelo autor (2025)

Figura 3 – Formação do *corpus* de pesquisa preliminar

Fonte: adaptado de Ermel (2021)

Quadro 2 – *Corpus* de pesquisa preliminar

Título	Autores	Journal
Comparative analysis of machine learning algorithms for predicting standard time in a manufacturing environment	Çaklıt e Dağdeviren (2023)	Artificial Intelligence for Engineering Design, Analysis and Manufacturing
Enhancing process lead time forecasting with machine learning and upstream process data: A case study in wind tower manufacturing	Flores-Huamán <i>et al.</i> (2025)	Computers & Industrial Engineering
Identification and prediction of standard times in machining for precision steel tubes through the usage of data analytics	Weller <i>et al.</i> (2023)	Procedia CIRP
Introduction to Work Study	Kanawaty (1992)	International Labour Office
Machine learning applied to forecasting the manufacturing time of new products prototypes and ETO products: An exploratory study	Rosa; Gonçalves; Barbalho (2025)	International Journal of Production Economics
Machine Learning in Small and Medium-Sized Enterprises, Methodology for the Estimation of the Production Time	Urban; Koblasa; Mendřický (2024)	Applied Sciences
Predicting Cycle Times in Textile Manufacturing Using Artificial Neural Network	Onaran e Yanık (2020)	Advances in Intelligent Systems and Computing
XGBoost: A Scalable Tree Boosting System	Chen e Guestrin (2016)	Association for Computing Machinery

Fonte: elaborado pelo autor (2025)

Além da atualidade da maior parte do corpus, os artigos foram veiculados em periódicos com métricas sólidas na área de engenharia. Três deles estão em títulos de quartil Q1 no SCImago: *Computers & Industrial Engineering* (SCImago Journal Rank, SJR, 2024 = 1,628, Q1; H-index 176), *International Journal of Production Economics* (SJR 2024 = 2,833, Q1; H-index 248) e *Association for Computing Machinery* (SJR 2024 = 2,246, Q1, H-index 146). Os demais foram publicados em periódicos bem ranqueados: *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* (SJR 2024 = 0,574, Q2; H-index 62) e *Applied Sciences* (SJR 2024 = 0,521, Q2; H-index 162). No SCImago, o SJR mede a influência científica de um periódico com base no prestígio das citações, o H-index (índice de Hirsch) é o maior número h tal que o periódico possui h artigos com pelo menos h citações cada, e os quartis classificam os títulos do Q1 (top 25%) ao Q4 por categoria temática. Esses indicadores reforçam que se trata de um esforço recente e em consolidação, acolhido por veículos de referência na área.

Do ponto de vista acadêmico, este trabalho contribui ao deslocar o foco dos estudos de predição de tempos, frequentes em operações de manufatura, para funções de apoio menos exploradas. A literatura sobre a aplicação de *machine learning* (ML) na manufatura tem se concentrado, predominantemente, em aplicações na linha de produção. Uma revisão de literatura recente, que analisou 114 artigos, evidenciou que a maioria dos estudos foca em funções como controle de qualidade, monitoramento de processos e predição de propriedades, com o controle de qualidade sendo o domínio mais pesquisado (Ördek; Borgianni; Coatanea, 2024). Em contraste, este trabalho foca na dimensão temporal da metrologia de calibração, isto é, na estimativa de tempos-padrão de inspeção em um laboratório HMLV, tema menos explorado em comparação ao uso de ML para decisão “conforme/não conforme” ou para previsão de propriedades.

Quando o foco é a estimativa de tempos, o viés em direção à produção direta persiste. Trabalhos como o de (Çaklı; Dağdeviren, 2023), que compararam algoritmos para prever tempos de soldagem, ou os de Chehade; Sylla; Diallo, Abdoul Rahime Doremus (2024) e Sun *et al.* (2022), que aplicaram ML para estimar tempos de fabricação e usinagem, respectivamente, permanecem ancorados em operações de linha de produção, ignorando setores de apoio. Há lacunas até mesmo na literatura de metrologia, que privilegia a aplicação de *virtual metrology* (VM) para prever parâmetros de qualidade do produto, mas a duração das inspeções remanesce negligenciada (Dreyfus *et al.*, 2022; Liu *et al.*, 2024).

O trabalho não se limita a um estudo de caso pontual, mas se propõe a oferecer uma referência para a aplicação de aprendizado de máquina em funções de apoio à produção. A

partir da análise de um laboratório de metrologia, busca-se avaliar a viabilidade de estimar tempos-padrão com ML em um ambiente complexo e variável e, com isso, delinear uma metodologia replicável que possa ser aproveitada em estudos futuros e em outras áreas operacionais indiretas.

Perante a gestão, a utilidade deste estudo consiste em fornecer tempos previstos confiáveis por família de calibradores, insumo direto para metas internas de *turnaround time* (TAT), entendido como o tempo decorrido entre o recebimento do item no laboratório e a conclusão da inspeção, bem como para o desenho de agendas e janelas de atendimento. Laboratórios de metrologia públicos divulgam metas explícitas de TAT, por exemplo de 10 dias úteis (*Washington State Department of Agriculture*, 2025), condicionando o agendamento e a priorização do trabalho, prática que depende de estimativas consistentes, mas que muitas vezes ainda são estimadas pelo *know-how* dos inspetores e gestores, usual na Empresa Z. Além disso, guias e manuais associados à ISO/IEC 17.025 destacam que a direção deve alocar recursos compatíveis com o TAT calculado e deve monitorar sua aderência, reforçando a necessidade de parâmetros de tempo para o planejamento de capacidade (ISO, 2017).

Para a empresa Z, o estudo é importante para estender ao setor de metrologia o indicador peças por hora-homem (PHH), padrão da planta para mensurar a produtividade. Sem esse indicador, torna-se difícil quantificar o impacto das iniciativas de melhoria, comparar períodos, e justificar a alocação de recursos a novos projetos no setor. Além disso, o sequenciamento de ReExames é consequência do sistema de *flow racks* no setor, a ordem dos exames é *First in First Out* (FIFO), gerando gargalos e assimetrias entre inspetores e máquinas, que ora ficam ociosos, ora sobrecarregados. A estimativa de tempos de inspeção viabiliza o cálculo do PHH e oferece base objetiva para sequenciamento e balanceamento.

A adoção de *machine learning* para tempos-padrão traz diversos benefícios potenciais. Primeiro, permite reduzir custos e tempo ao evitar estudos de cronometragem extensivos em campo para cada novo produto. Em vez de medir todas as combinações possíveis, mede-se apenas uma parcela das operações e o modelo estima os tempos das demais, o que é especialmente relevante em contextos de alto volume de variedade (Çaklıt; Dağdeviren, 2023). Segundo, contribui para o planejamento futuro, ao permitir estimar de forma rápida e antecipada o tempo de novas operações ainda na fase de desenvolvimento ou preparação da produção, usando essas estimativas em estudos de balanceamento de linha, dimensionamento de capacidade, definição de prazos e análise de custos. Terceiro, esses modelos podem aprender com os dados históricos da empresa, ajustando-se às especificidades do ambiente produtivo em questão, capturando fatores como eficiência média dos operadores da empresa, características

dos equipamentos, etc., mais abrangentemente do que uma fórmula manual conseguiria. Por fim, a reprodutibilidade é alta: as estimativas podem ser rapidamente revisadas e recalibradas à medida que novos dados são incorporados, seja por coleta automática, ou por novos estudos de tempo.

Este estudo pode servir como referência para estimativas de tempo em outros setores de apoio, para além da metrologia. Os princípios centrais, como estratificar o portfólio em famílias; padronizar definições; mapear *drivers* de carga; e estimar tempos de serviço para compará-los; são os mesmos para outras funções indiretas como manutenção, qualidade, engenharia, gestão da produção e logística interna. Na prática, a transposição ocorre substituindo “calibradores” pelos objetos de trabalho de cada função. Em manutenção, a unidade pode ser a ordem de serviço.

1.4 ESTRUTURA DO TRABALHO

Este estudo está organizado em 6 seções, sendo esta primeira a Introdução. Nesta seção foram contextualizados o ambiente de laboratório de metrologia e o problema tratado, além de serem apresentados o problema de pesquisa, os objetivos do trabalho, a justificativa, o escopo e as contribuições esperadas.

Na seção 2 apresenta-se a fundamentação teórica necessária ao desenvolvimento do trabalho, abrangendo conceitos de estudo de tempos e determinação de tempos-padrão, variabilidade e vieses em cronometragem/avaliação de ritmo, padronização e qualidade de dados operacionais, bem como noções de modelagem preditiva aplicadas à gestão de operações.

A seção 3 descreve os procedimentos metodológicos: caracterização do contexto empírico (setor de metrologia), critérios de seleção e coleta dos dados, etapas de preparação e curadoria, desenho do experimento, métricas de avaliação e estratégias de validação adotadas. Na seção 4 detalham-se os resultados da abordagem proposta, e são apresentados os tempos-padrão estimados pelo estudo.

A seção 5 discute os resultados obtidos, estimativas de ganho para planejamento e programação, recomendações operacionais e evidências do potencial de aplicação na empresa Z, incluindo limitações práticas e diretrizes de implantação.

Por fim, a seção 6 apresenta as conclusões, respondendo aos objetivos, destacando contribuições acadêmicas e gerenciais, além de expor limitações do estudo e sugestões para pesquisas futuras e para a aplicação dos modelos em ambientes produtivos reais.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo será abordado o referencial teórico utilizado para a elaboração desta pesquisa. Inicialmente serão abordados conceitos gerais sobre tempo-padrão, em seguida desafios em processos de alta variabilidade, por fim será apresentado como é feita a abordagem de *machine learning* em literaturas recentes.

2.1 TEMPO-PADRÃO: CONCEITOS, TÉCNICAS E DESAFIOS

O conceito de tempo-padrão surgiu no início do século XX, nos estudos de tempos e movimentos de Frederick W. Taylor e de Frank e Lillian Gilbreth. Taylor, por meio da administração científica, introduziu a cronometragem sistemática para determinar o “tempo definido” de cada tarefa, buscando sua “melhor maneira” de execução (Kalne; Mehendale, 2022); (Çaklt; Dağdeviren, 2023). Paralelamente, os Gilbreth desenvolveram o *motion study*, decompondo tarefas em micromovimentos para eliminar etapas desnecessárias e aperfeiçoar métodos de trabalho (Stephens, 2024). Esses esforços fundamentaram o conceito de tempo-padrão, definido como o tempo necessário para que um operador médio, em condições normais e ritmo padrão, conclua uma tarefa, incluindo folgas para necessidades pessoais, fadiga e contingências inevitáveis (International Labour Office; Kanawaty, 1992). Ao longo das décadas, o tempo-padrão consolidou-se como elemento essencial para o planejamento e controle nas organizações: sem referências temporais confiáveis, é difícil elaborar cronogramas de produção, balancear linhas, estimar capacidades ou calcular custos e prazos de entrega (Çaklt; Dağdeviren, 2023).

No desdobramento do tempo-padrão, as tolerâncias têm a função de ajustar o tempo normal às condições reais de trabalho, contemplando necessidades pessoais, fadiga (física e mental) e demoras inevitáveis (Silva, A V; Coimbra, 1980). A literatura enfatiza que não se trata de “folga” arbitrária, mas de um componente técnico destinado a preservar desempenho sustentável e a neutralizar variações não controláveis do ambiente (Barnes, 1977; Kanawaty, 1992). Para operacionalizar esses acréscimos, manuais e guias propõem classificações qualitativas do esforço e do contexto, Silva e Coimbra (1980) distinguem graus de esforço físico e mental, relacionando cada grau a percentuais orientativos e a fatores de recuperação quando há exigências fisiológicas ou cognitivas que requeiram pausas. Assim, tarefas com posturas estáticas, atenção visual contínua, tomada de decisão sob incerteza ou monotonia prolongada tendem a demandar tolerâncias mais altas que operações mecânicas simples, pois elevam o

desgaste e o risco de erro — algo especialmente pertinente a ambientes de metrologia, em que a acurácia, a inspeção minuciosa e a interação com sistemas podem gerar carga mental e interrupções frequentes (Kanawaty, 1992; Silva, A V; Coimbra, 1980).

A literatura também diferencia duas vias de obtenção dessas tolerâncias. A primeira é normativa/analítica, baseada em tabelas e listas de verificação que traduzem condições ergonômicas e organizacionais em classes de referência — abordagem útil para padronizar critérios e garantir consistência entre áreas (Barnes, 1977; Kanawaty, 1992; Silva, A V; Coimbra, 1980). A segunda é empírica, via amostragem do trabalho e observação sistemática dos tempos de interrupções e pausas em populações de tarefas similares; quando a organização possui dados históricos confiáveis, essa via permite calibrar os percentuais às condições locais e às mudanças de processo (Das, 1990). Em ambos os casos, recomenda-se documentar claramente a origem das tolerâncias, os critérios de enquadramento (grau de esforço, fatores ambientais, demandas cognitivas) e as hipóteses de aplicação, de modo que o tempo-padrão não seja subestimado, pressionando o operador; nem superestimado, prejudicando o planejamento, e permaneça auditável ao longo do tempo (Barnes, 1977; Das, 1990; Kanawaty, 1992).

Com os critérios de tolerâncias estabelecidos, a literatura mostra diferentes formas de obter o tempo-padrão a partir de abordagens complementares, que variam conforme a natureza do processo, a possibilidade de observação direta e a disponibilidade de dados: i) medição direta do trabalho real; ii) análise por decomposição com tempos padronizados; iii) inferência a partir de dados, seja por reuso de históricos ou por modelos de aprendizado de máquina; e iv) observação amostral para estimar parcelas produtivas/improdutivas em ciclos irregulares, a seguir detalham-se essas abordagens.

A cronometragem clássica consiste em observar diretamente o ciclo de trabalho com um cronômetro, segmentar a tarefa em elementos com início e fim bem definidos, e registrar múltiplas repetições para obter o tempo básico. Essa técnica é, particularmente, indicada em processos repetitivos e de ciclo estável, pois exige repetições suficientes para reduzir a variabilidade amostral e produzir estimativas confiáveis do desempenho normal (Kanawaty, 1992). Na prática, a qualidade do resultado depende do treinamento do analista, do tamanho amostral, e do controle de vieses como efeito Hawthorne (mudança de comportamento sob observação) (Goodwin *et al.*, 2017; Kalne; Mehendale, 2022). Em processos com múltiplos fatores influentes, a aplicação direta de cronometragem pode reduzir a acurácia dos tempos previstos, pois o estudo foca o movimento e desconsidera variáveis de contexto relevantes (Onaran; Yanık, 2020).

Com o avanço das abordagens de engenharia de métodos, surgiram os Sistemas de Tempos Predeterminados (PTS), que eliminam a necessidade de cronometragem direta em cada tarefa. O principal é o MTM (*Methods-Time Measurement*) que atribui um tempo-base a cada movimento elementar (Morlock *et al.*, 2017). A adoção de PTS permite estimativas rápidas de tempos-padrão, com alta confiabilidade, desde que o catálogo de elementos e seus tempos padronizados sejam aplicados corretamente (Snakly; Furler; Aylan, 2019). Essas técnicas são amplamente empregadas em balanceamento de linhas de produção, planejamento de capacidade, orçamentação de custos de mão-de-obra e avaliações de produtividade, pois fornecem métricas consistentes para identificar gargalos, dimensionar recursos e basear decisões em dados concretos (Teshome; Meles; Yang, 2024).

A vídeo-cronoanálise filma a operação e, no pós-processamento, permite analisar quadro a quadro os elementos do ciclo, capturando micro-movimentos e pausas que podem escapar ao cronômetro; estudos de visão computacional mostram que é possível classificar elementos e estimar automaticamente o tempo de cada elemento a partir do vídeo (Ji; Pannakkong; Buddhakulsomsiri, 2022; Luciano *et al.*, 2022). Resultados em manufatura indicam concordância estatística entre tempos estimados por modelos baseados em vídeo e valores de referência obtidos por analistas (Ji; Pannakkong; Buddhakulsomsiri, 2022).

Esse método é indicado quando há ciclos curtos ou movimentos sutis, em casos que a observação em tempo real leva à perda de detalhes, e quando se deseja auditar e reutilizar registros para treinamento e melhoria de método (Ji; Pannakkong; Buddhakulsomsiri, 2022). A literatura expõe que os algoritmos conseguem segmentar operações e atribuir tempos a elementos com boa precisão em cenários reais (Ji; Pannakkong; Buddhakulsomsiri, 2022).

Quando há uma base de tempos históricos coletados, por exemplo, em *Manufacturing Execution Systems* (MES), é possível estimar tempos-padrão para novas tarefas similares sem refazer todo o estudo, construindo uma linha de base que refletem tendências e melhorias ao longo do tempo (Ki *et al.*, 2023). Um exemplo de aplicação é o trabalho realizado por Ki *et al.* (2023) que propõe uma abordagem para calcular a taxa de mudança do tempo de produção a partir de dados históricos de tempo-padrão, com a finalidade de prever capacidade futura e planejar investimentos conforme o processo evolui.

Em ML, treinam-se modelos preditivos que aprendem a relação entre características da tarefa e o tempo realizado em históricos, produzindo estimativas de tempo-padrão para novas ordens sem cronometragem peça a peça (Çaklt; Dağdeviren, 2023). A abordagem é indicada quando há diversidade de itens e volume de dados suficiente para treinar e validar modelos, além de necessidade de atualização contínua (Çaklt; Dağdeviren, 2023). Em cenários HMLV,

o ML também pode suportar decisões de sequenciamento e dimensionamento ao integrar-se com simulação e técnicas de agrupamento de tarefas, reduzindo incertezas associadas à alta variabilidade (Gödri, 2022). No caso analisado por Weller *et al.* (2023), vários modelos de ML superam, em R^2 , a prática de PTS; porém, os autores reforçam que o desempenho é dependente do contexto, ou seja, não há um algoritmo que traz uma performance superior sempre.

O *work sampling* estima a proporção do tempo gasto em diferentes atividades por meio de observações aleatórias ao longo do turno; a fração de observações em cada estado aproxima a percentagem de tempo naquele estado, possibilitando calcular o tempo-padrão (Buchmeister; Herzog, 2024; Kanawaty, 1992). É indicado para processos longos ou de ciclo irregular, nos quais cronometrar cada unidade seria impraticável; tempo e dinheiro gasto podem ser de 35 a 80% inferior quando comparado com o método tradicional de tomada de tempo por cronômetro (Buchmeister; Herzog, 2024). Revisões recentes reforçam o uso do método para compreender perdas e utilização em sistemas produtivos, guiando decisões de dimensionamento de recursos (Buchmeister; Herzog, 2024).

Esse método também apresenta desvantagens. O registro das observações pode envolver certo grau de subjetividade, gerando inconsistências na coleta e na interpretação dos dados; há dificuldades para registrar postos individuais, sobretudo quando estão fisicamente afastados; e não se capturam diferenças individuais entre trabalhadores quando a observação é feita sobre um grupo (Buchmeister; Herzog, 2024). Além disso, o que se obtém é uma visão panorâmica das atividades, sem detalhes sobre tarefas ou processos específicos (Buchmeister; Herzog, 2024).

A Modelagem Humana Digital (DHM) simula operador em 3D para projetar e analisar tarefas. Revisões sobre DHM na indústria mostram o uso da tecnologia para avaliar métodos e alternativas de desenho de trabalho em fase de desenvolvimento, reduzindo retrabalho físico (da Silva, Adailton Gonçalves; Gomes; Winkler, 2022; Schall Mark C.; Fethke; Roemig, 2018). Quanto à integração com sistemas de tempos pré-determinados, há evidência acadêmica recente: Mazareinezhad; Sekkay e Imbeau (2025) propõem e evidenciam um framework de integração do Maynard *Operation Sequence Technique* (MOST) dentro de sistemas de DHM, com validação em campo comparando as estimativas geradas pelo método com tempos observados e com o próprio MOST manual; o estudo mostra a aplicação no DELMIA *Ergonomic Workplace Design*, automatizando a extração de parâmetros de tempo a partir do cenário 3D. Além disso, o artigo registra que o software Jack da Siemens “permite análise integrada de ergonomia e de tempo usando padrões MTM-1”, reforçando que existem soluções de mercado acoplando DHM e PTS (Mazareinezhad; Sekkay; Imbeau, 2025).

Quadro 3 – Técnicas de obtenção de tempo padrão

Método	Síntese	Indicação	Aplicação
Cronometragem	Observação direta do ciclo de trabalho com cronômetro; divisão da tarefa em elementos e aplicação de fator de desempenho e folgas para chegar ao tempo-padrão.	Processos repetitivos de ciclo estável; quando é possível cronometrar várias repetições.	(Rehman <i>et al.</i> , 2019)
Sistemas Predeterminados (MTM, MOST, etc.)	Decomposição da tarefa em movimentos elementares com tempos padronizados. Soma-se esses tempos para obter o tempo-padrão, sem necessidade de cronometragem direta em cada ocorrência.	Operações repetitivas e de volume elevado; planejamento de linhas de produção.	(Breznik; Buchmeister; Vujica Herzog, 2023)
Vídeo-cronoanálise	Filmagem da operação e análise quadro a quadro para extrair tempos de cada elemento, identificando micro-movimentos e pausas imperceptíveis em tempo real, melhorando precisão e detalhamento do estudo de tempos.	Tarefas com movimentos sutis ou ciclos curtos, onde é difícil captar detalhes ao cronometrar.	(Luciano <i>et al.</i> , 2022)
Dados históricos	Uso de banco interno de tempos medidos em operações passadas para estimar tempos-padrão de novas tarefas similares, acelerando o processo de obtenção de tempos; requer ajuste de fatores de correção quando há diferenças no contexto	Processos com histórico consistente de medições; quando não é viável medir diretamente, e a precisão não precisa ser muito grande.	(Ki <i>et al.</i> , 2023)
<i>Machine Learning</i>	Treinamento de modelos preditivos com dados históricos de tempos e características das tarefas para estimar automaticamente o tempo das atividades.	Ambientes com volume de dados históricos e grande diversidade; quando métodos tradicionais são inviáveis.	(Çaklı; Dağdeviren, 2023)
<i>Work Sampling</i> (amostragem do trabalho)	Observações aleatórias do posto e, pelo percentual de tempo produtivo vs. improdutivo, calcula o tempo normal e adiciona folgas para chegar ao tempo-padrão.	Processos longos ou com ciclos irregulares, onde cronometrar cada peça seria inviável.	(Nuroctavia <i>et al.</i> , 2024)
Modelagem Humana Digital (DHM) + PTS	Simula o operador em 3D por meio de softwares como <i>Ergonomic Workplace Design</i> (DELMIA) ou Jack (Siemens), e gera automaticamente os tempos de cada movimento usando códigos MOST/MTM embutidos no software.	Fases de projeto ou quando não há linha física montada; permite testar layouts antes de investir.	(Mazareinezhad; Sekkay; Imbeau, 2025)

Fonte: Elaborado pelo autor (2025)

Processos produtivos caracterizados por alta variabilidade – frequentemente descritos como ambientes *high-mix/low-volume* (HMLV), de ampla variedade de produtos com volume reduzido por tipo – apresentam desafios distintos em comparação a operações de mix restrito e alto volume. Nesses contextos, há flutuação da demanda e ampla heterogeneidade dos itens, tornando a gestão diária desafiadora diante de indicadores de desempenho conflitantes, por exemplo, manter alta pontualidade de entregas, reduzir *lead-time* e maximizar a utilização de recursos simultaneamente (Gödri, 2022). Em ambientes HMLV típicos, os tamanhos de lote

tendem a ser pequenos, o volume de ordens é elevado e a variabilidade dos produtos está em crescimento, exigindo dos sistemas produtivos flexibilidade, capacidade de resposta rápida e eficiência de custos acima do usual (Gödri, 2022). Tais características são encontradas tanto em setores de manufatura sob encomenda quanto em operações de inspeção e metrologia, onde uma instalação pode lidar com uma grande diversidade de instrumentos ou peças a serem calibrados/verificados, cada qual com requisitos e tempos de processamento distintos.

Um dos principais obstáculos nesses processos de alta variabilidade é viabilizar, em termos de esforço e abrangência, a determinação de tempos-padrão confiáveis para as atividades. Metodologias tradicionais de cronometragem e sistemas de tempos predeterminados (como o MTM) são conceitualmente adequados para definir tempos-padrão, mas demandam observação intensiva e preparação detalhada, o que se torna pouco factível quando há milhares de combinações pouco repetitivas.

Além disso, qualquer abordagem baseada em dados depende da qualidade das medições. Erros e inconsistências no processo de medição se propagam para as estimativas. Frequentemente, não há dados históricos suficientes de cada variante de produto para estimar tempos com precisão, e quando existem estimativas, elas podem não representar fielmente a realidade de um processo altamente variável (Machado *et al.*, 2019). Por exemplo, Machado *et al.* (2019) analisaram o impacto da variabilidade nos tempos-padrão em linhas de montagem e observaram que usar um único valor fixo para simular operações pode introduzir erros significativos; nas simulações, erros superiores a 6% surgiram quando a variabilidade real dos tempos de ciclo não foi adequadamente considerada, prejudicando a confiabilidade do modelo.

Além de complicar a medição do tempo-padrão, a variabilidade elevada afeta diretamente a medição de desempenho operacional. Estudos apontam que a variabilidade deteriora o desempenho de sistemas produtivos ao aumentar os tempos de ciclo médios e o trabalho em andamento (WIP), bem como ao elevar a imprevisibilidade dos resultados. Isso reduz o nível de serviço e força a alocação de buffers de tempo ou de estoque maiores para absorver flutuações (Alfieri; Castiglione; Pastore, 2024).

Frente a esses desafios, pesquisadores propõem diversas estratégias para lidar com a alta variabilidade em processos produtivos, visando melhorar a estimativa de tempos e o desempenho do sistema, conforme expõe o Quadro 4.

Quadro 4 – Estratégias para ambientes com alta variabilidade

Estratégia	Descrição	Referências
Classificação e Agrupamento de Tarefas (<i>Group Technology</i>)	Identificação de similaridades entre produtos ou operações para agrupá-los em famílias homogêneas, criando tempos-padrão por família e reduzindo a variância interna.	(Deuse; Konrad; Bohnen, 2013); (Gödri, 2022)
Buffers e Controle de WIP	Uso de buffers (estoque intermediário ou tempo de folga) para absorver flutuações. Dimensionamento de WIP via análise estatística da variabilidade e adoção de sistemas <i>pull</i> (Kanban) com <i>takt time</i> alinhado à demanda, estabilizando o <i>throughput</i> em cargas irregulares.	(Herps <i>et al.</i> , 2022)
Medições Estatísticas Avançadas e Análise Multivariada	Aplicação de ferramentas estatísticas além de média e variância para capturar a distribuição completa dos tempos de processamento. Uso de <i>design of experiments</i> e regressão múltipla para identificar fontes críticas de variabilidade.	(Alfieri; Castiglione; Pastore, 2024)
<i>Machine Learning</i> e Sistemas Inteligentes	Combinação de simulação de produção com algoritmos de ML para otimizar sequenciamento e dimensionamento de lotes em tempo real. Modelos preditivos treinados com históricos variados para estimar tempos-padrão automaticamente e <i>clustering</i> de tarefas, adaptando previsões conforme novos dados.	(Gödri, 2022)

Fonte: Elaborado pelo autor (2025)

Com os conceitos e limitações dos métodos tradicionais estabelecidos, a seção seguinte apresenta a modelagem estatística e o uso de *machine learning* empregados para estimar os tempos-padrão no contexto analisado.

2.2 MODELAGEM ESTATÍSTICA E MACHINE LEARNING PARA A ESTIMATIVA DOS TEMPOS-PADRÃO

Revisões recentes indicam expansão acelerada do uso de ML em funções de manufatura e, especificamente para tempos-padrão, há evidência de que a medição direta pode ser impraticável em certos contextos, sendo então proposto o uso de ML para estimar o tempo (Çaklt; Dağdeviren, 2023; Ördek; Borgianni; Coatanea, 2024). A ideia central é utilizar dados disponíveis do processo para treinar modelos que predigam o tempo de execução de uma tarefa, dispensando (total ou parcialmente) a necessidade de estudos de tempo tradicionais para cada nova instância (Çaklt; Dağdeviren, 2023). A seguir, abordam-se técnicas de mineração de dados e, posteriormente, os modelos supervisionados aplicados à estimação de tempos.

Mineração de dados (*data mining*) é o conjunto de métodos que extraem padrões úteis e não triviais a partir de dados, inserindo-se no processo mais amplo de *Knowledge Discovery in Databases* (KDD) — que inclui etapas de seleção, pré-processamento, transformação, mineração e avaliação/implantação do conhecimento descoberto (Fayyad; Piatetsky-Shapiro; Smyth, 1996). Em manufatura, revisões sistemáticas mostram um uso crescente de mineração

de dados em processos, manutenção, qualidade e planejamento, com ganhos como redução de custos, aumento de produtividade e melhor responsividade — especialmente quando integrada a dados de chão de fábrica (MES) e de gestão (ERP) (Harding *et al.*, 2005). Esses trabalhos classificam aplicações por tipo de conhecimento buscado, destacando que a escolha do método deve refletir o objetivo analítico (explicar padrões x prever resultados) e as restrições dos dados (volume, granularidade, rotulagem) (Choudhary; Harding; Tiwari, 2009).

Para estimativa e gestão de tempos operacionais, técnicas de análise de dados podem cumprir papéis complementares. Em um primeiro nível, métodos de mineração de dados são empregados com finalidade exploratória e descritiva, revelando regularidades estruturais e combinações de fatores de processo associadas a tempos elevados, sem objetivo direto de predição. Em um segundo nível, modelos de aprendizado de máquina supervisionado utilizam esses dados históricos para estimar tempos em novas instâncias, apoiando decisões de planejamento e capacidade, abordagem cuja relevância vem sendo destacada em levantamentos recentes sobre aplicações de data mining e ML em manufatura (Choudhary; Harding; Tiwari, 2009). Revisões destacam que combinar tarefas descritivas e preditivas é especialmente valioso: padrões descritivos explicam regularidades do processo e orientam melhoria; modelos preditivos estimam resultados para novos casos, apoiando planejamento e tomada de decisão (Choudhary; Harding; Tiwari, 2009; Han; Kamber; Pei, 2011; Harding *et al.*, 2005).

Entre as tarefas descritivas, as regras de associação são o resultado de processos de mineração que buscam padrões frequentes nos dados, produzindo proposições do tipo “se—então”. A qualidade dessas regras é usualmente avaliada por suporte, confiança e *lift*. O suporte indica com que frequência, em termos relativos, a combinação de elementos da regra aparece no conjunto de dados; a confiança representa a probabilidade de o consequente ocorrer quando o antecedente já ocorreu; e o *lift* mede a relação entre essa confiança e a probabilidade de o consequente ocorrer independentemente do antecedente, podendo indicar alavancagem, indiferença ou repulsão entre antecedente e consequente (Han; Kamber; Pei, 2011).

As regras podem ser obtidas pelo *A priori*, algoritmo que explora o fato de que, se um conjunto de itens ocorre com frequência, então qualquer parte dele também deve ocorrer com frequência; com isso, é possível podar rapidamente combinações que não valem a pena testar e focar nos padrões realmente frequentes, dos quais se extraem as regras fortes (Agrawal; Srikant, 1994; Han; Kamber; Pei, 2011). Com esse eixo descritivo estabelecido, na sequência apresentam-se os modelos supervisionados para previsão de tempos, mantendo as regras como apoio à interpretação e à tomada de decisão.

Quadro 5 – Regras de associação

Medida	Definição	Interpretação
Suporte	Percentual de casos em que antecedente e consequente aparecem juntos no conjunto de dados.	Suportes maiores sugerem padrão amplamente presente e potencialmente mais acionável; suportes menores indicam padrão raro. A interpretação depende do contexto do tamanho e granularidade da base
Confiança	Entre os casos em que o antecedente ocorre, percentual em que o consequente também ocorre.	$\geq 0,90$ muito alta; 0,75–0,89 alta; 0,60–0,74 aceitável; $< 0,60$ fraca.
Lift	O quanto a ocorrência conjunta de antecedente e consequente é mais (ou menos) provável do que seria por acaso se não houvesse relação entre eles.	$> 2,0$ associação forte; 1,5–2,0 moderada; 1,1–1,49 fraca; $\approx 1,0$ independência; $< 1,0$ associação negativa.

Fonte: Adaptado de Agrawal e Srikant (1994); Han, Kamber e Pei (2011)

A regressão linear múltipla ajusta uma equação linear que relaciona a variável-alvo a vários preditores, estimando coeficientes por mínimos quadrados (Roustaei, 2024; Urban; Koblasa; Mendřický, 2024). Ela serve para descrever relações e prever valores esperados do alvo para combinações de preditores, sendo um ponto de partida transparente e interpretável (Çaklt; Dağdeviren, 2023; Roustaei, 2024). Na prática, tende a funcionar bem quando os efeitos são aproximadamente proporcionais e aditivos (Ho; Cole, 2023). Em cenários com não linearidades e interações significativas entre variáveis de processo, estudos recentes mostram que árvores e *ensembles* costumam superar a regressão linear em acurácia, o que reforça seu papel como baseline e a necessidade de comparar modelos em dados reais (Çaklt; Dağdeviren, 2023; Urban; Koblasa; Mendřický, 2024).

A regressão Ridge usa os mesmos princípios da regressão linear, mas adiciona um termo de penalização L2 para contrabalançar o impacto de expressivas variâncias, o que encolhe os coeficientes, sem chegar a zero, e estabiliza a regressão quando há multicolinearidade entre preditores (Çaklt; Dağdeviren, 2023; Magklaras *et al.*, 2024). É útil quando o objetivo é obter previsões estáveis em bases tabulares com diversos descritores correlacionados (Magklaras *et al.*, 2024).

A regressão Lasso, assim como a Ridge, usa penalização para conter *overfitting*; porém, por empregar L1, pode reduzir coeficientes a zero e realizar seleção de variáveis embutida (Çaklt; Dağdeviren, 2023; Tibshirani, 1996). O modelo desempenha bem quando o sinal é esparsa (poucos preditores com efeitos moderados), cenário no qual estudos clássicos mostram vantagem da Lasso sobre alternativas como Ridge ou seleção de subconjuntos; quando os efeitos são numerosos e pequenos, a Ridge tende a ir melhor (Emmert-Streib; Dehmer, 2019; Tibshirani, 1996).

O Elastic Net combina penalizações L1 e L2 no mesmo critério de ajuste: a L1 permite seleção de variáveis (coeficientes podem ir a zero) e a L2 promove encolhimento estável dos coeficientes. Além disso, o Elastic Net apresenta o chamado *grouping effect*: tende a manter juntas variáveis altamente correlacionadas, mitigando a instabilidade da Lasso nesse cenário (Zhou, 2013; Zou; Hastie, 2005). Na prática, o Elastic Net é indicado quando há muitos preditores com correlação entre si ou mesmo situações $p \gg n$, pois busca um equilíbrio entre parcimônia (via L1) e estabilidade (via L2). Revisões recentes destacam seu uso como alternativa robusta à Lasso em presença de dependência entre covariáveis; aplicações empíricas recentes também mostram cenários em que o Elastic Net supera Lasso e Ridge em desempenho preditivo sob multicolinearidade (Freijeiro-González; Febrero-Bande; González-Manteiga, 2022; Nur; Jaya; Siswanto, 2023).

Árvores de decisão (CART) constroem previsões por particionamento recursivo binário: a cada passo escolhe-se a variável e o ponto de corte que melhor separam os dados (Çaklt; Dağdeviren, 2023; Retzlaff *et al.*, 2024). O treinamento começa no conjunto inteiro e, um passo de cada vez, escolhe a pergunta que melhor separa os registros em dois grupos mais “parecidos” quanto ao alvo; depois repete o mesmo procedimento dentro de cada grupo, criando novos ramos; o processo para quando não há mais benefício relevante ou quando se atinge limites definidos (Retzlaff *et al.*, 2024). São considerados modelos interpretáveis porque representam uma estrutura hierárquica de regras do tipo se–então–senão, que pode ser facilmente visualizada, compreendida e interpretada por humanos (Retzlaff *et al.*, 2024).

Modelos de árvore de decisão funcionam bem como *baseline* interpretável quando se deseja expor, em regras “se–então”, como combinações de atributos levam à previsão; a estrutura hierárquica facilita revisar quais variáveis e faixas de valor mais influenciam o resultado (Retzlaff *et al.*, 2024). Elas capturam não linearidades e interações automaticamente, exigindo pouco preparo dos dados, mas são instáveis (pequenas mudanças no conjunto de treino podem gerar árvores bem diferentes) e tendem a sobreajustar se crescerem demais; por isso, é recomendável limitar a complexidade e usar validação cruzada para escolher esses parâmetros (Last; Maimon; Minkov, 2002).

Random Forest treina muitas árvores de decisão diferentes e, no fim, combina as respostas dessas árvores (Breiman, 2001). Cada árvore é treinada em uma amostra aleatória dos registros e, a cada divisão dentro da árvore, considera-se apenas um subconjunto aleatório das variáveis para escolher a melhor “pergunta”, esse procedimento faz com que as árvores “errem de jeitos diferentes” e, ao juntar todas, o modelo final fique menos sujeito a oscilações do que uma árvore única, generalizando melhor (Breiman, 2001).

Na prática, em dados tabulares com não linearidades e interações entre atributos de processo, Random Forest costuma superar uma árvore única e aparecer entre os modelos mais consistentes sem exigir muito pré-processamento; estudos aplicados à estimativa de tempos/produção em ambiente industrial relatam desempenho competitivo frente a regressões e a outros algoritmos, o que justifica seu uso como referência prática em comparações (Çaklı; Dağdeviren, 2023). Em contraste com expressões polinomiais de grau elevado, que aumentam a variância do ajuste e tendem a capturar ruído (*overfitting*), o Random Forest reduz variância ao combinar muitas árvores mantidas pouco correlacionadas; na formulação original, o erro de generalização da floresta depende justamente da força média das árvores e da correlação entre elas (Breiman, 2001; James *et al.*, 2014).

Para prever um novo caso, o k-nearest neighbors (kNN) procura, nos dados de treino, os k exemplos mais parecidos por meio da distância entre as variáveis, e usa a média das respostas desses vizinhos como previsão (Çaklı; Dağdeviren, 2023). É um método não paramétrico: não supõe forma funcional da relação entre preditores e alvo; parte apenas da ideia de que exemplos “próximos” tendem a ter respostas semelhantes (Arora; Khanduja; Bansal, 2022). Como a noção de “proximidade” depende de uma métrica de distância, as escalas das variáveis influenciam fortemente o resultado; por isso, é comum padronizar/normalizar os preditores antes do kNN e selecionar k por validação cruzada (Arora; Khanduja; Bansal, 2022). Também, em alta dimensionalidade as distâncias tendem a se concentrar (pontos “mais próximos” e “mais distantes” ficam parecidos), o que pode tornar o kNN pouco informativo; estudos mostram esse efeito e discutem como a escolha de métrica e a presença de dimensões irrelevantes agravam o problema (Aggarwal; Hinneburg; Keim, 2001).

Artificial Neural Networks (ANNs) são computacionais formados por muitos elementos simples (“neurônios”) conectados por ligações com pesos; cada neurônio calcula uma combinação ponderada das entradas e aplica uma função de ativação, em geral não linear (Montesinos López; Montesinos López; Crossa, 2022). As redes são organizadas em camadas e a informação pode fluir apenas para frente (redes *feedforward*) ou incluir realimentação no tempo (redes recorrentes) (Montesinos López; Montesinos López; Crossa, 2022). O treinamento ajusta os pesos por métodos de gradiente baseados em retropropagação do erro, que indicam como alterar os parâmetros em cada camada para reduzir a perda (Lin; Li, 2023).

A Support Vector Regression (SVR) aceita pequenos desvios entre o valor previsto e o observado dentro de uma faixa de tolerância ϵ , que pode ser visualizada como um “corredor” ao redor da curva de predição (Smola; Schölkopf, 2004). O objetivo é encontrar uma curva simples que mantenha a maior parte dos pontos dentro desse corredor, penalizando apenas quem

fica fora; são justamente esses pontos “fora do corredor” que determinam a solução, por isso recebem o nome de “vetores de suporte” (Smola; Schölkopf, 2004). O parâmetro C ajusta o equilíbrio entre manter a curva simples e pagar por violações do corredor, enquanto ϵ define quão larga é essa faixa de tolerância (Çaklt; Dağdeviren, 2023; Smola; Schölkopf, 2004).

Gradient Boosting constrói o modelo por etapas: parte de uma previsão inicial e, a cada iteração, calcula o negativo do gradiente da perda nas previsões atuais, ajusta um novo componente para aproximar esse sinal e o adiciona ao modelo; com isso, o método consegue otimizar diretamente perdas diferenciáveis (Natekin; Knoll, 2013). Quando os blocos básicos são árvores de decisão, tem-se o Gradient Tree Boosting, algoritmo de ML usado com frequência por cientistas de dados para alcançar resultados expressivos em diversos desafios de ML (Chen; Guestrin, 2016).

Dentre os algoritmos de Gradient Boosting, XGBoost é um conjunto (*ensemble*) de árvores de decisão, projetado para ser altamente escalável (Bentéjac; Csörgő; Martínez-Muñoz, 2021). O LightGBM é voltado à criação de um algoritmo computacionalmente eficiente, baseado na pré-computação de histogramas dos atributos (*features*), como no XGBoost (Bentéjac; Csörgő; Martínez-Muñoz, 2021). Por fim, o CatBoost modifica o cálculo dos gradientes para evitar o deslocamento da predição (*prediction shift*), com o objetivo de melhorar a acurácia do modelo (Bentéjac; Csörgő; Martínez-Muñoz, 2021).

No estudo de Çaklt e Dağdeviren (2023), foram comparados 13 algoritmos de ML distintos para estimar tempos-padrão em um ambiente fabril, incluindo modelos lineares (regressão linear simples e regularizada) e não lineares (k-NN, Random Forest, redes neurais, etc.). Os resultados expõem que vários desses modelos conseguem obter adequada acurácia preditiva; especificamente, o algoritmo k-NN destacou-se pelos menores erros de previsão em uma aplicação envolvendo uma fábrica de ônibus, superando inclusive técnicas como redes neurais e regressão avançada (Çaklt; Dağdeviren, 2023). Nesse estudo, as variáveis de entrada mais influentes para estimar o tempo-padrão de montagem foram a quantidade de operações de solda e o número de peças/componentes do produto, informações que refletem a complexidade do produto e do processo (Çaklt; Dağdeviren, 2023).

Quadro 6 – Algoritmos de ML

Algoritmo	Indicação	Limitações	Referências
Regressão linear múltipla	Relações aproximadamente lineares; <i>baseline</i> interpretável.	Sensível a outliers e multicolinearidade; perde acurácia com não linearidades/interações fortes.	Çaklt; Dağdeviren (2023); Ho; Cole (2023); Roustaei (2024); Urban; Koblasa; Mendřický (2024)
Regressão Ridge	Previsões estáveis com muitos preditores correlacionados.	Não zera coeficientes (não seleciona variáveis); requer padronização e ajuste de λ .	Çaklt; Dağdeviren (2023); Magklaras <i>et al.</i> (2024)
Regressão Lasso	Seleção de variáveis e parcimônia; bom com sinal esparsos.	Instável com alta correlação; piora quando há muitos efeitos pequenos.	Çaklt; Dağdeviren (2023); Tibshirani (1996)
Elastic-net	Muitos preditores correlacionados ou $p \gg n$; equilíbrio entre seleção e estabilidade; tende a agrupar variáveis correlatas.	Dois hiperparâmetros (λ e α) para calibrar; menos esparsos que Lasso puro.	Freijeiro-González; Febrero-Bande; González-Manteiga (2022); Nur; Jaya; Siswanto (2023); Zhou (2013)
Árvore de decisão (CART)	Regras se-então interpretáveis; captura não linearidades e interações simples com pouco pré-processamento.	Alta variância/instabilidade; sobreajusta se muito profunda; requer poda e validação.	Çaklt; Dağdeviren, (2023); Retzlaff <i>et al.</i> (2024)
Random forests	Tabular com não linearidades/interações; mais consistente que uma árvore única; pouco pré-processamento.	Menos interpretável; suaviza extremos.	Breiman (2001); Çaklt; Dağdeviren (2023); James <i>et al.</i> (2014)
K-nearest neighbors	Padrões locais; bom baseline em bases pequenas/médias.	Requer normalização; sofre em alta dimensionalidade; escolha de k.	Aggarwal; Hinneburg; Keim (2001); Arora; Khanduja; Bansal (2022); Çaklt; Dağdeviren (2023)
Redes neural artificial (ANN)	Não-linearidades complexas; melhora com mais dados.	Exige <i>tuning</i> / regularização; menos interpretável; sensível à escala; risco de <i>overfitting</i> .	Lin; Li (2023); Montesinos López; Montesinos López; Crossa (2022)
Support Vector Regression (SVR)	Não-linearidades moderadas; boa com poucos outliers.	Sensível a $C/\epsilon/\gamma$; escalar features; custo cresce com N.	Çaklt; Dağdeviren (2023); Smola; Schölkopf (2004)
Gradient Boosting (XGBoost, LightGBM, CatBoost)	Alta acurácia em dados tabulares; captura interações complexas; variantes rápidas/eficientes.	Tuning fino (learning rate, profundidade); risco de <i>overfit</i> .	Bentéjac; Csörgő; Martínez-Muñoz (2021); Chen; Guestrin (2016); Natekin; Knoll (2013)

Fonte: Elaborado pelo autor (2025)

Entretanto, a implementação de modelos de ML não está isenta de desafios. A qualidade e quantidade dos dados de treino são fatores críticos para o sucesso de qualquer modelo preditivo. Em ambientes de fabricação com poucos dados, ainda é comum basear-se em estimativas empíricas ou aproximações grosseiras, dada a dificuldade de coletar volume de dados necessário para o ML (Urban; Koblasa; Mendřický, 2024). Pesquisas recentes têm proposto metodologias específicas para contornar essa limitação, como a categorização de tipos

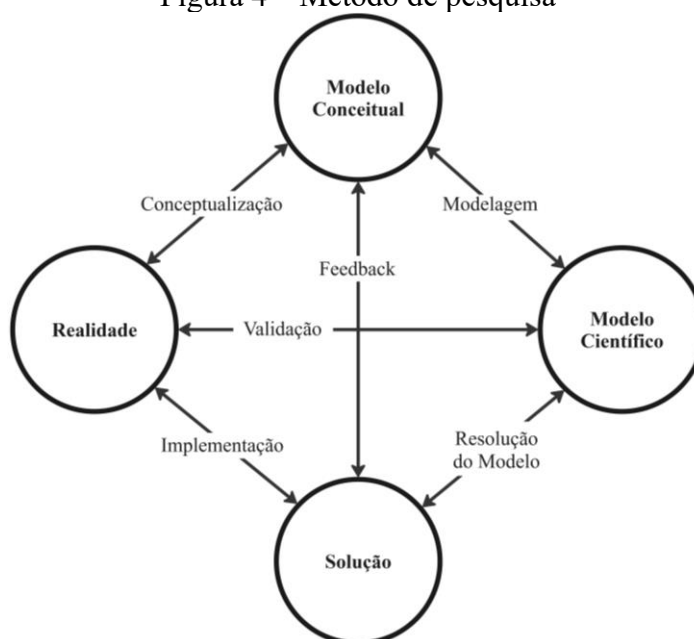
de *jobs* e técnicas de ampliação de base de dados, tornando possível aplicar ML mesmo com dados limitados (Urban; Koblasa; Mendřický, 2024). Por exemplo, uma abordagem recente adaptada para PMEs empregou classificação de tipos de produto e partição de dados segundo um teorema de convergência, combinando estatística e ML para estimar tempos de produção com poucos dados disponíveis (Urban; Koblasa; Mendřický, 2024). Os resultados indicaram que, ao segmentar corretamente os dados e escolher modelos adequados, é viável obter estimativas razoáveis; contudo, certas previsões falharam quando não havia correlação forte entre as variáveis de entrada e o tempo. No caso citado, a estimativa de *lead-time* não obteve sucesso devido à fraca dependência entre os atributos escolhidos e o resultado (Urban; Koblasa; Mendřický, 2024).

Em suma, o uso de modelos estatísticos e de *machine learning* desponta como uma solução promissora para estimar tempos-padrão em contextos de alta variedade e dinâmica. Conforme discutido na literatura, quando a mensuração tradicional é difícil ou custosa, técnicas preditivas baseadas em dados podem assumir esse papel com precisão adequada (Çaklıt; Dağdeviren, 2023).

3 PROCEDIMENTOS METODOLÓGICOS

Este estudo caracteriza-se como pesquisa empírica, mais especificamente do tipo empírico-prescritiva, porque parte de dados reais do processo são usadas para construir e estimar modelos. A partir deles, propõe-se modelos decisórios para apoiar o planejamento e a operação (Bertrand; Fransoo; Udenio, 2023). Na tipologia de pesquisas quantitativas baseadas em modelos, a classe empírico-prescritiva é aquela que percorre o ciclo completo apresentado na Figura 4, indo além da mera descrição para orientar a ação gerencial (Bertrand; Fransoo; Udenio, 2023).

Figura 4 – Método de pesquisa



Fonte: Adaptado de Bertrand, Fransoo e Udenio (2023)

Nesse enquadramento empírico-prescritivo, adota-se a modelagem baseada em casos: parte-se de um caso delimitado do mundo real para compreender o fenômeno em seu contexto e extrapolar os resultados por generalização analítica — isto é, conectando as evidências do caso a proposições teóricas, e não a uma população por inferência estatística (Flyvbjerg, 2006; Yin, 2018). Essa forma de generalização é reconhecida como cientificamente útil quando o caso é teoricamente motivado e analisado com rigor (Flyvbjerg, 2006).

Do ponto de vista do processo de pesquisa, o estudo segue o ciclo de modelagem proposto por Bertrand; Fransoo; Udenio (2023), orientado pelas transformações mostradas na Figura 4. O ponto de partida é a conceptualização, em que a situação observada no laboratório é abstraída em termos de recursos, fluxos e variáveis relevantes para o problema de tempos-

padrão. Em seguida ocorre a modelagem, na qual esse esquema conceitual é traduzido em um modelo formal: definição dos dados, seleção das variáveis de entrada e saída e escolha da classe de modelos de aprendizado de máquina a ser estimada. A etapa seguinte é a resolução do modelo, quando os algoritmos são treinados e validados sobre os dados históricos para gerar estimativas de tempo normal e tempos-padrão por percentis. Depois, realiza-se a validação, comparando as previsões com tempos observados e discutindo os resultados com especialistas do laboratório, o que produz feedback para eventuais ajustes no modelo conceitual e científico. A implementação é tratada aqui de forma prospectiva, por meio da proposição de como os tempos estimados podem ser incorporados ao planejamento e aos indicadores do laboratório.

A rigorosidade decorre de práticas como definição clara do objeto do caso, encadeamento lógico entre proposições e evidências, análises “dentro do caso” e iterações controladas entre teoria e dados, que são padrão em pesquisas que constroem modelos a partir de casos (Eisenhardt; Graebner, 2007; Voss; Tsikriktsis; Frohlich, 2002).

A escolha de modelagem baseada em casos é coerente com o objetivo do estudo, produzir uma prescrição a partir de dados reais de um sistema operacional específico, seguindo o ciclo proposto por Bertrand, Fransoo e Udenio (2023) típico da pesquisa baseada em casos em operações (Voss; Tsikriktsis; Frohlich, 2002). Esse encadeamento visa apoiar decisão gerencial no contexto do processo analisado. Além disso, revisões metodológicas em *operations management* recomendam explicitamente os estudos de caso para desenvolver e testar modelos próximos da prática, conectando evidências a proposições úteis à gestão (Voss; Tsikriktsis; Frohlich, 2002).

Do ponto de vista de rigor, a literatura define rotinas para construir explicações sólidas “a partir do caso”: análise dentro do caso, lógica de replicação (quando houver múltiplos casos) e ligação transparente entre dados e proposições, o que sustenta contribuições teóricas e uso prático mesmo em caso único bem fundamentado (Eisenhardt; Graebner, 2007). Em suma, para um problema operacional com acesso profundo a dados do processo e necessidade de decisões contextualizadas, a modelagem baseada em casos é adequada e rigorosa, alinhando relevância prática e solidez metodológica (Bertrand; Fransoo; Udenio, 2023; Yin, 2018).

3.1 UNIDADE CONTEXTO E DE ANÁLISE

A Empresa Z é uma multinacional automotiva de grande porte, com operações em mais de 60 países e cerca de 418 mil colaboradores no mundo. Globalmente, opera em múltiplas áreas de mobilidade e manufatura avançada, com centenas de milhares de colaboradores e

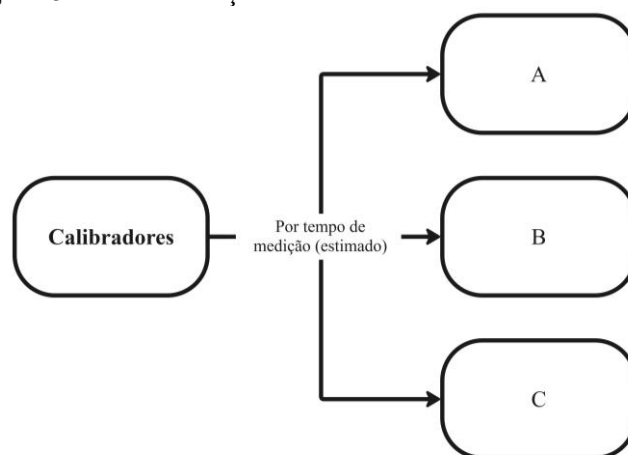
faturamento anual na ordem de dezenas de bilhões. Em 2024, registrou faturamento global de € 90,3 bilhões e investimento em P&D de € 7,8 bilhões, com aproximadamente 86,8 mil profissionais dedicados a pesquisa e desenvolvimento, o que sustenta alto grau de digitalização fabril (ERP/MES, coleta e tratamento de dados de processo). No Brasil, a empresa possui um contingente de mais de 8 mil colaboradores em diversas unidades, onde linhas produtivas processam milhares de peças por dia e o laboratório de metrologia atende um portfólio de alto mix e rigor metrológico. Essa escala, somada à maturidade de dados operacionais, reforça a relevância e a viabilidade empírica do estudo.

A Empresa Z é adequada como campo empírico desse trabalho por três razões principais e complementares. Primeiro, pela relevância prática, pois o laboratório de metrologia opera em ambiente de alta variedade e a falta de tempos-padrão confiáveis afeta diretamente o PCP, o nivelamento de carga e o atendimento a prazos. Segundo, pela viabilidade empírica, tendo em vista a autorização e acesso ao histórico de atividades registradas na base de dados da empresa, fornecendo insumos para modelagem e validação. Também foi permitido acompanhar e registrar as atividades realizadas pelos inspetores, *in loco*. Terceiro, pelo potencial de implementação, a equipe gestora engajada e infraestrutura digital permitem incorporar os tempos previstos aos sistemas, viabilizando medir ganhos reais. Além de ser um caso representativo de funções de apoio em manufatura discreta (o que amplia a transferibilidade dos resultados), esse contexto permite percorrer o ciclo conceitualização–modelagem–solução–validação de completamente.

A Empresa Z possui diversas salas de inspeção responsáveis por assegurar a qualidade dos lotes fabricados, garantindo a conformidade com os padrões estabelecidos na planta. Nesse contexto, o setor de metrologia desempenha um papel central, sendo o principal responsável pela calibração e manutenção da qualidade dos dispositivos de medição utilizados na linha de produção. Atendendo a uma demanda semanal superior a 400 calibradores, o setor conta com uma estrutura composta por máquinas de medição por coordenadas (CMM), equipamentos de medição óptica e ampla variedade de instrumentos de medição manual. A precisão exigida nos processos é elevada, e as tolerâncias dimensionais das peças inspecionadas chegam à ordem de décimos de micrômetro.

A classificação ocorre baseada em um estudo interno realizado em 2014, que classificou os calibradores em 3 grupos principais, baseado no *know-how* dos operadores quanto à dificuldade de se realizar o processo de inspeção. A dificuldade de inspeção é menor na categoria A, e mais alta na categoria C, mas não há um documento sobre os detalhes dessa classificação.

Figura 5 – Classificação dos calibradores antes do estudo



Fonte: Elaborado pelo autor, adaptado da Empresa Z (2025)

Quadro 7 – Características da classificação anterior ao estudo

Classificação	Características	Tempo de Inspeção
A	Calibradores simples, normalmente com até 2 características a serem medidas, as quais são simples. Para máquinas automáticas, os programas para esses calibradores costumam estar prontos.	≤ 5 minutos • Exceto quando há algum tipo de não conformidade.
B	Calibradores de média complexidade, normalmente com até 5 características a serem medidas, ou com poucas características, mas de difícil acesso, ou que exigem inspeção em mais de uma máquina.	≤ 12 minutos • Pode extrapolar com facilidade caso haja não conformidade ou necessidade de fazer algum ajuste prévio à inspeção.
C	Alta complexidade de inspeção, comumente exigem inspeção em mais de uma máquina e possuem mais de 5 características a serem medidas. Podem requerer fixação especial nas máquinas, ou outros tipos de cuidado prévio à inspeção que aumentam o tempo de <i>setup</i> significativamente.	≥ 15 minutos • Pode durar até 1 hora quando há não conformidade.

Fonte: Elaborado pelo autor, adaptado da Empresa Z (2025)

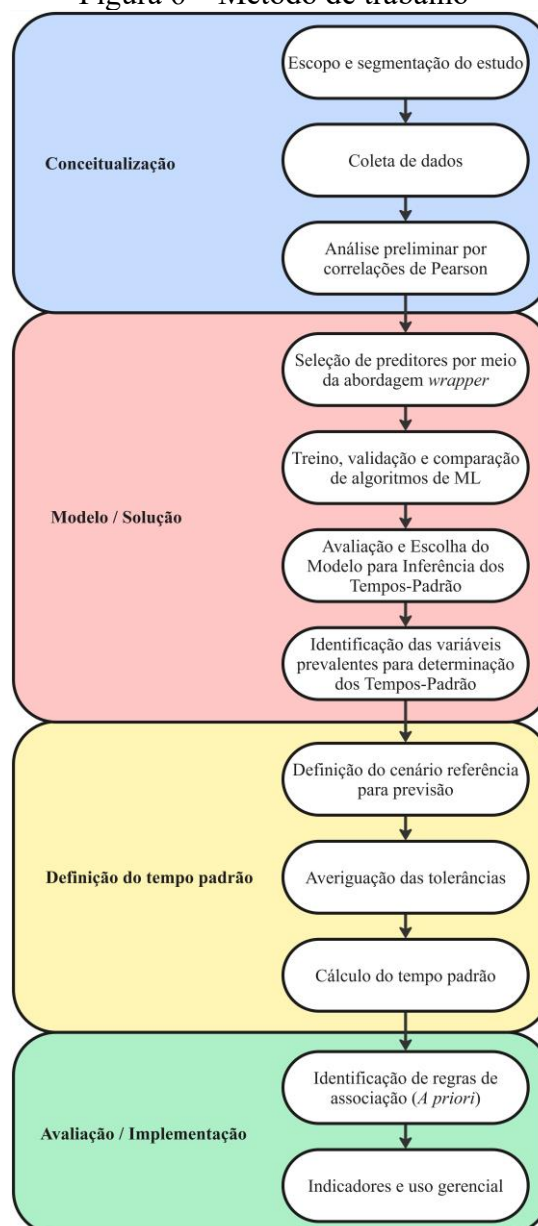
O departamento de metrologia carece de indicadores quantitativos para mensurar sua produtividade. Os indicadores são ligados sobretudo a conformidade e prazos, por exemplo, % de inspeções na meta de *lead-time* de 1 dia no setor de metrologia, e taxa de reprovação. O setor não possui métricas de saída por recurso, eficiência de uso da capacidade ou variabilidade de tempos.

3.2 MÉTODO DE TRABALHO

O método de trabalho define a sequência de etapas que orienta a condução do estudo e deve ser seguido a rigor para assegurar replicabilidade e clareza (Dresch; Lacerda; Antunes Júnior, 2014; Mentzer; Flint, 1997). Esse fluxo está alinhado ao método de pesquisa baseado

em modelos em operações, no qual a investigação percorre as fases descritas na Figura 4. A Figura 6 sintetiza essa correspondência no presente estudo: Conceitualização (escopo, dados, análise exploratória de dados) → Modelo/Solução, seleção de preditores e comparação de algoritmos que compõem o modelo científico → Definição do tempo-padrão, tradução da previsão em parâmetro de engenharia, “resolução” do modelo → Avaliação/Implementação, regras de associação, indicadores e uso gerencial, etapa de validação e retorno de aprendizagens ao processo (Bertrand; Fransoo; Udenio, 2023).

Figura 6 – Método de trabalho



Fonte: Adaptado de Stefano (2020), Gori (2016) e Bertrand, Fransoo e Udenio (2023)

Na Conceitualização, definem-se o escopo e a segmentação do estudo (máquinas, famílias e faixas de quantidade de medidas), consolidam-se os dados utilizados no estudo, e realiza-se uma análise exploratória inicial. Essa etapa inclui limpeza de dados, tratamento de *outliers* e checagens simples de correlação para entender relações e evitar *leakage* — entendido como o uso, direta ou indiretamente, de informações que não estarão disponíveis no momento da previsão ou que são derivadas do próprio alvo, o que inflaria artificialmente o desempenho do modelo e reduziria sua generalização (Kaufman *et al.*, 2012). O resultado é um conjunto de variáveis coerente com a operação e pronto para modelagem.

Na faixa Modelo/Solução, o conjunto de preditores é refinado por uma abordagem *wrapper*, um procedimento que avalia subconjuntos de variáveis usando o próprio modelo como critério, treinando-o e validando-o a cada combinação para reter apenas os preditores que melhoram o desempenho fora da amostra (Kohavi; John, 1997). Com os preditores fixados, diferentes algoritmos de aprendizado são treinados e comparados sob a mesma partição e o mesmo pré-processamento. O modelo adequado é selecionado por desempenho (R^2 , RMSE) e interpretabilidade; em seguida, extraem-se medidas de importância de variáveis para orientar hipóteses e futuras ações operacionais.

A faixa Definição do tempo-padrão traduz a previsão do modelo para um parâmetro de engenharia. Primeiro, fixa-se um cenário de referência para previsão (por exemplo: programa pronto, sem retrabalho e sem *setup* extra), obtendo o tempo normal do processo. Após, determinam-se as tolerâncias (pessoais e fadiga/recuperação, eventualmente diferenciadas por máquina), que transformam o tempo normal em tempo-padrão publicável e comparável entre segmentos.

Por fim, em Avaliação/Implementação, aplicam-se regras de associação (*A priori*) para identificar combinações de condições associadas a tempos elevados (suporte, confiança e *lift*), o que ajuda a priorizar causas e contramedidas. Os resultados alimentam indicadores gerenciais — como percentis P50/P75/P95 do tempo-padrão por máquina/faixa, aderência do realizado ao previsto e impactos por fator — e fecham o ciclo de melhoria contínua: as descobertas retornam à base de dados e à segmentação, permitindo ajustes periódicos no modelo e nas práticas operacionais.

Nas subseções seguintes serão detalhadas cada etapa do método de trabalho.

3.3 COLETA DE DADOS

Inicialmente, realizou-se o acompanhamento e a caracterização dos processos do setor de metrologia. Durante essa fase, foram analisadas documentações pré-existentes sobre as atividades realizadas no setor de metrologia, por meio desses documentos foi elaborado um sumário das atividades apresentado no Quadro 8.

Quadro 8 - Processos do setor de metrologia

Atividade	Descrição
Re-Exame	Aferição contínua dos calibradores em uso na linha de produção, baseado em um sistema de “prazo de validade”, a fim de se manter a confiabilidade dos calibradores por meio de diagnósticos. Calibradores que não passam nesse exame podem ser retrabalhados ou sucateados.
Exame-Novo	Validação de calibradores novos antes do seu emprego nas medições, assegurando conformidade com suas especificações. Definição da metodologia de Re-Exame.
Atendimentos de Produtos	Medição de componentes tanto para a produção quanto para demandas de engenharia, aproveitando a precisão superior das máquinas de metrologia em relação às salas de medição dos galpões. Tarefa crítica quando há algum problema de qualidade na linha de produção.
Pré-Exame	Avaliação preliminar de peças semiacabadas, identificando ajustes necessários antes do exame final. Pré-Exames não são documentados no sistema e são realizados sob demanda.

Fonte: Autor, empresa Z (2025)

Todas as tarefas são sequenciadas em cada estação por *flow racks*. Em regra, o sequenciamento segue *First In First Out* (FIFO), mas em situações críticas — como quando há necessidade de inspeção ligada a um problema em uma linha de produto — pré-exames e atendimentos podem receber prioridade, ultrapassando a fila mediante autorização. O setor dispõe de nove estações, mas opera com três a quatro inspetores por turno. Por isso, os inspetores são polivalentes, aptos a operar várias máquinas (com frequência, três ou mais). O sequenciamento permite trabalhos paralelamente entre máquinas diferentes; entretanto, cada máquina processa apenas uma atividade por vez. Consequentemente, a capacidade do sistema depende da alocação dos inspetores e dos tempos de cada estação.

A Empresa Z não dispõe de dados estruturados para todas as atividades do laboratório: apenas Re-Exame e Exame-Novo estão registrados no registro de atividades do setor. Com isso, as estimativas de demanda (Figura 2) refletem apenas parte do trabalho e tendem a subestimar a carga real do setor. Além disso, os tempos de inspeção não são registrados no sistema, exigindo a cronoanálise manual do processo para mensurar tempos representativos e estimar os tempos-padrão.

Para obter os tempos, adotou-se o procedimento padrão de estudo de tempos com cronômetro: i) definição das condições de execução; ii) segregação da tarefa em elementos; iii)

registro de ciclos por elemento em folha de observação; iv) obtenção do tempo normal e inclusão de tolerâncias para estimar o tempo-padrão (Freivalds; Niebel, 2014; Kanawaty, 1992). Como havia apenas dois inspetores por estação de trabalho, um para cada turno, o índice de desempenho foi desconsiderado, visto que não havia amostras suficientes sobre o trabalho para aplicar o conceito. Operacionalmente, os tempos foram medidos com cronômetros e registrados em planilha que reproduz a folha de estudo de tempos.

Para definir os elementos que compõe o tempo de inspeção, observou-se o processo para diferentes calibradores e listaram-se tarefas como: retirada no *flow rack*, consulta à norma e requisitos de inspeção na base de dados, limpeza do instrumento e montagem de dispositivos de medição. Entretanto, surgiram limitações: i) muitos elementos tinham duração expressivamente curta frente ao horizonte total de inspeção, elevando o erro de medição; ii) havia alta variabilidade entre máquinas e famílias (o mesmo elemento assumia durações distintas conforme tecnologia/acessórios); iii) parte das tarefas era acíclica/condicional (p.ex., troca de dispositivo ao alternar famílias); iv) a necessidade de ajustes dependia da similaridade entre calibradores, o que tornava o tempo imprevisível. Diante disso, agrupou-se as tarefas em elementos de maior abrangência e adotar unidades inteiras de minutos. Essa aproximação, embora incorra em erros de arredondamento aceitáveis, reduz a variação gerada pela fragmentação excessiva das medições e permite a execução da cronoanálise por um único avaliador. As categorias resultantes desse agrupamento estão apresentadas no Quadro 9.

Quadro 9 – Descrição dos elementos de tempo

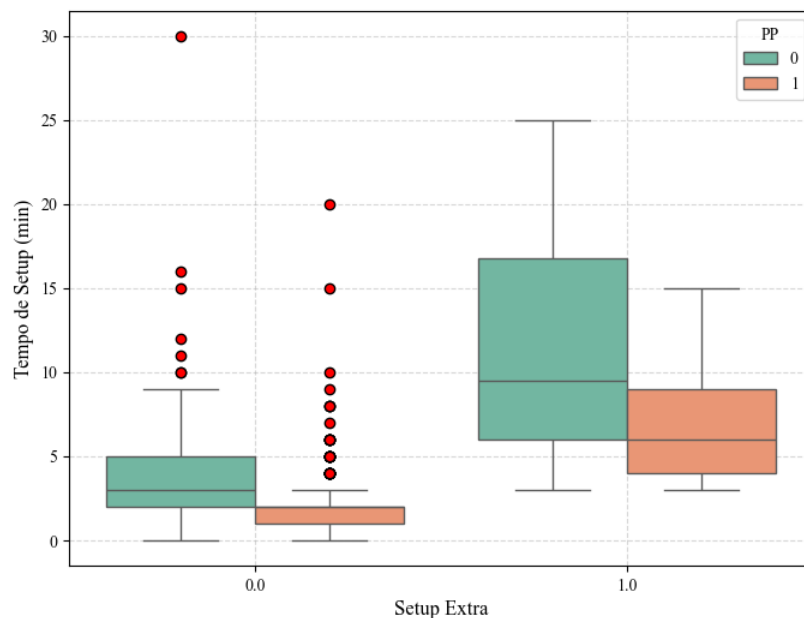
Elemento	Descrição
Tempo de <i>Setup</i>	Qualquer tarefa que o inspetor precisa realizar antes da medição. Não inclui a calibração diária da máquina.
Tempo de Medição	Período em que a peça está efetivamente sendo medida. Pode ser manual ou automatizada.
Tempo de Análise	Fase de comparação dos dados obtidos com os padrões, verificação de não conformidades e registro do resultado no sistema.

Fonte: Autor, empresa Z (2025)

Durante o acompanhamento das tarefas descritas no Quadro 8, observou-se que o tempo dedicado ao *setup* — isto é, à preparação dos equipamentos e dos calibradores e à configuração dos parâmetros de medição — representava parcela expressiva do tempo total de inspeção. Em várias medições, o *setup* consumiu entre 20% e 30% do tempo de ciclo, refletindo etapas como troca de fixações, calibragem inicial, desmonte e limpeza da peça e ajustes de *software*. Outro fator crítico para o aumento do tempo de *setup* é a programação das máquinas automáticas: calibradores sem programa prévio demandavam um tempo superior para serem configurados. Para facilitar a análise, passou-se a classificar como “*setup* extra” toda ação

executada durante o *setup* de calibração que não se referisse à programação das máquinas. Os efeitos desses fatores podem ser observados na Figura 7. No eixo horizontal, *Setup* Extra = 0 representa inspeções sem etapa adicional de *setup* e *Setup* Extra = 1 representa inspeções com etapa extra, que podem ser a preparação de um dispositivo de fixação, ou a limpeza do calibrador a ser medido. Dentro de cada grupo, as cores distinguem a condição do programa de medição, PP = 0 indica que há necessidade de criar um programa do zero, enquanto PP = 1 indica que há programa pronto, ou que a inspeção não precisa de programa (em máquinas manuais).

Figura 7 – Influência de *setup* extra e necessidade de programação no tempo de *setup*



Fonte: Elaborado pelo autor (2025)

Com as atividades principais do setor devidamente mapeadas (Quadro 8), e uma descrição do processo, definiu-se, os dados a serem coletados e as respectivas unidades. Para garantir a representatividade da amostra e apoiar a análise estatística, selecionaram-se variáveis operacionais, como tempo de medição, tempo de *setup* e ocorrências de não conformidade, e atributos descritivos dos calibradores, norma, família, máquina usada e tipo de tarefa. Além disso, foram incluídos indicadores binários para eventos críticos (por exemplo, “*setup* extra” e “problemas de máquina”) e registros de datas. Todas as variáveis coletadas estão descritas na Tabela 1.

Tabela 1 - Dados Coletados

Dado	Sigla	Valor exemplo	Unidade / Tipo
Data	-	23/01/2025	-
Norma da peça	-	4716500318	-
Família da peça	-	165	-
Máquina	-	Máquina 1	-
Tarefa	-	ReExame	-
<i>Setup extra?</i>	Setup Extra	<i>True / False</i>	<i>Boolean</i>
<i>Programa pronto?</i>	PP	<i>True / False</i>	<i>Boolean</i>
<i>Problema na máquina?</i>	MP	<i>True / False</i>	<i>Boolean</i>
<i>Há não conformidade?</i>	NC	<i>True / False</i>	<i>Boolean</i>
Tempo de <i>setup</i>	tsetup	8	minutos
Tempo de medição	tmedicao	4	minutos
Tempo de análise	tanalise	1	minutos
Tempo de inspeção	tfinal	13	minutos
Quantidade de características medidas	NMedidas	12	-

Fonte: Dados coletados pelo autor (2025)

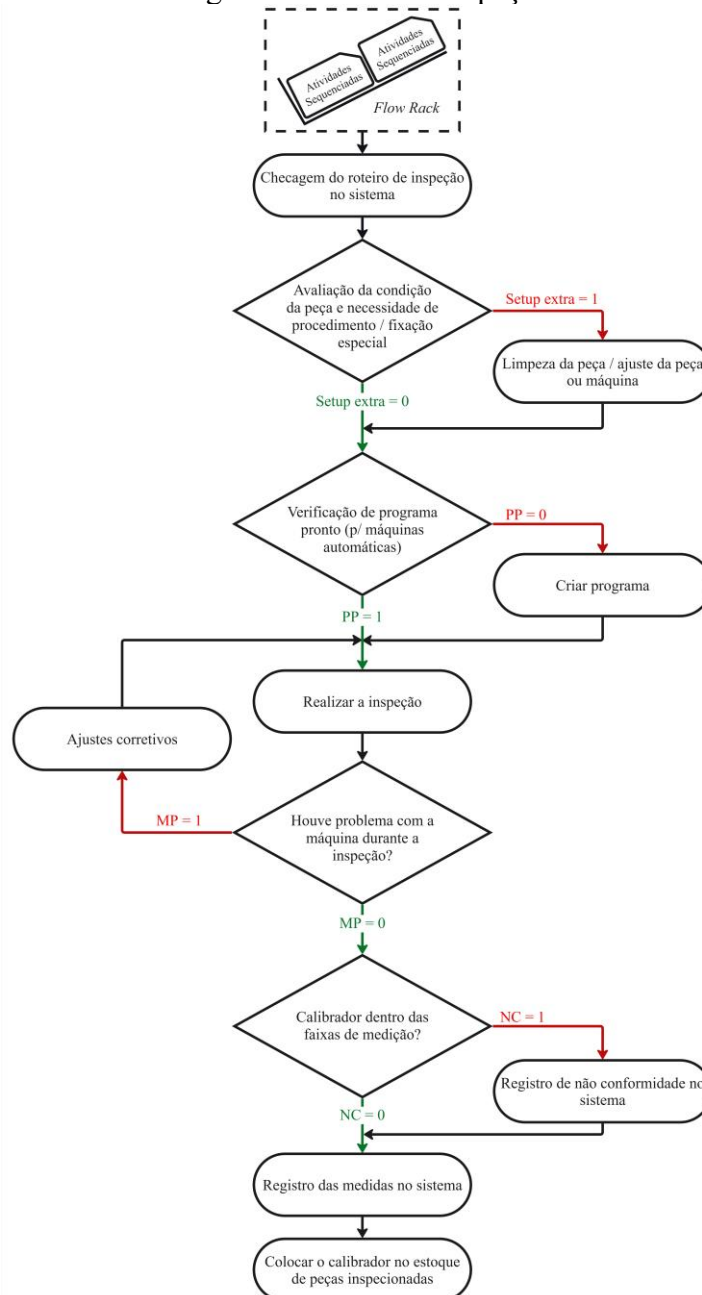
Antes de selecionar os preditores para os modelos de ML, realiza-se uma análise exploratória por meio de um mapa de correlações de Pearson para verificar possíveis dependências bivariadas e colinearidades entre os preditores. Essa etapa não pretende explicar causalmente o tempo de inspeção, mas situar o terreno estatístico em que os modelos serão ajustados, identificando redundâncias potenciais e sinais complementares que justificam a inclusão conjunta das variáveis no pipeline preditivo.

3.4 DEFINIÇÃO DO ESCOPO DA ANÁLISE

A análise contempla operações de inspeção rotineiras realizadas no período definido, cobrindo famílias de calibradores atendidas pelas nove estações. Entram no escopo os ciclos completos de inspeção realizados sob método padrão vigente e condições de operação normais de turno. Foram excluídos da análise: i) interrupções extraordinárias (falhas de utilidades, queda de luz), ii) medições descartadas por erro de procedimento, iii) medições com mudança de método no meio do ciclo.

A unidade de análise é o ciclo de inspeção de um calibrador, detalhado na Figura 8. O tempo observado refere-se ao tempo de processamento do posto (da disponibilização do item para medição ao término da leitura/registro), com distinção entre *setup*/ajustes condicionais e etapas cíclicas. Atividades administrativas fora do posto (ex.: logística interna, aprovação documental externa) não compõem o ciclo. No setor são utilizadas nove estações de trabalho com diferentes máquinas/instrumentos, as distribuições de tempo são evidenciadas na Figura 9. Durante um período de um mês foram coletados tempos de inspeção para todas as estações, totalizando 246 registros.

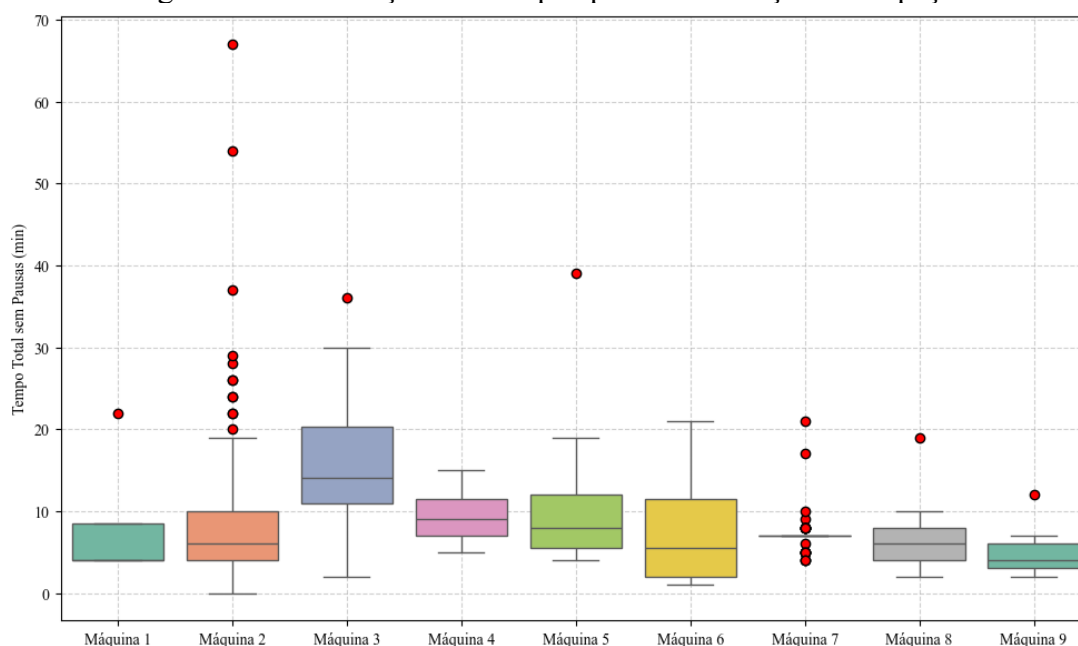
Figura 8 – Fluxo de inspeção



Fonte: Elaborado pelo autor (2025)

As principais limitações do escopo são: i) janela temporal curta (sensível a sazonalidade). O período analisado cobre apenas 2 meses, o que pode não refletir variações sazonais. Isso tende a instabilizar parâmetros e reduzir a generalização do modelo. ii) ausência de tempos registrados automaticamente no sistema, portanto depende-se de cronoanálise e apontamentos manuais, sujeitos a erros de medição e arredondamentos. Isso pode introduzir viés e ruído aos resultados. iii) variabilidade residual entre máquinas e operadores não é totalmente capturada. Pode haver diferenças de estado da máquina, nível de experiência e microcondições (*setup* fino, ferramental, ajustes) que não estão 100% representadas nas variáveis. Isso limita a explicabilidade e a precisão dos modelos. Essas limitações são mitigadas por segmentação, descarte de outliers procedimentais e marcação explícita de eventos condicionais.

Figura 9 – Distribuição dos tempos para cada estação de inspeção



Fonte: Elaborado pelo autor

Com o desdobramento do escopo e das limitações delineadas, a próxima etapa organiza o conjunto bruto em dados analiticamente utilizáveis.

3.5 ANÁLISE DE DADOS

Com a base consolidada para uso analítico, passa-se à análise exploratória de dados (*Exploratory Data Analysis*, EDA) para compreender o comportamento das variáveis e sua relação com o tempo de inspeção, identificar padrões, possíveis redundâncias e eventuais inconsistências. Essa leitura inicial prepara o terreno para as escolhas de modelagem e validação apresentadas a seguir.

Antes de modelar, aplicar correlações de Pearson como parte da EDA é útil por três razões complementares: (i) ajuda a entender a estrutura dos dados e a orientar hipóteses ainda na fase exploratória, reduzindo retrabalho depois; (ii) permite detectar redundâncias entre preditores e sinalizar multicolinearidade, fenômeno que aumenta a variância dos coeficientes e dificulta a interpretação de modelos de regressão; e (iii) integra um pipeline de pré-modelagem recomendado em textos de referência de modelagem preditiva, onde filtros simples (como inspeções de correlação) antecedem seleções *wrapper/embedded* e ajudam a estabilizar o conjunto de variáveis (Guyon; Elisseeff, 2003; Shrestha, 2020; Tukey, 1977).

Para essa análise, adotou-se um procedimento EDA: (i) selecionar apenas variáveis numéricas não constantes; (ii) calcular a correlação de Pearson (r) entre pares de preditores e entre cada preditor e o alvo (tempo total), registrando também os valores- p ; e (iii) visualizar a matriz em *heatmap* para destacar magnitudes e sinais (Lee Rodgers; Nicewander, 1988; Wilkinson; Friendly, 2009). O coeficiente de Pearson é uma medida clássica de associação linear — pode ser visto, entre outras interpretações equivalentes, como o médio produto cruzado de variáveis padronizadas, o que o torna adequado para triagens rápidas em conjuntos de dados industriais.

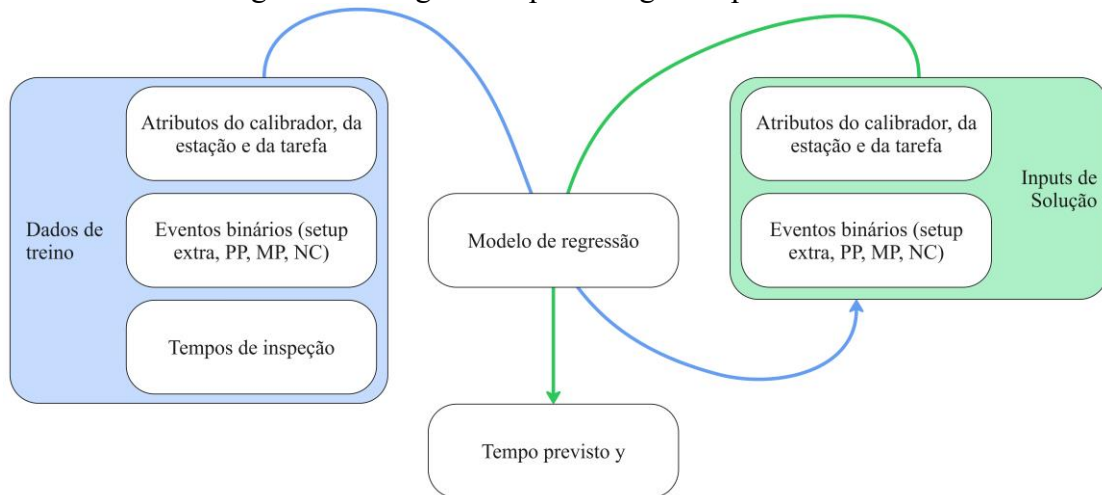
Após delimitar o escopo, registrar a coleta e realizar a análise exploratória, inicia-se a descrição do método de modelagem em ML. Este estudo formula a previsão do tempo total de inspeção como um problema de aprendizado supervisionado por regressão: deseja-se estimar uma variável contínua y (tempo, em minutos) a partir de um vetor de preditores X que descrevem o item, o posto e as condições operacionais.

$$y = f(X) + \varepsilon \quad (1)$$

Na equação, f é uma função desconhecida aprendida a partir dos dados observados e ε representa o erro aleatório. Na literatura de *statistical learning*, essa formulação é clássica para problemas com resposta numérica e permite acomodar relações lineares ou não lineares entre

preditores e resposta (Hastie; Tibshirani; Friedman, 2009). A ênfase é em prever y com desempenho preditivo aceitável em observações futuras a partir de um conjunto fixo de preditores, em linha com a prática de *predictive modeling* (Kuhn; Johnson, 2013) e com a distinção entre objetivos preditivos e explicativos discutida por Shmueli (2010).

Figura 10 – Lógica de aprendizagem supervisionada



Fonte: Elaborado pelo autor

O vetor de preditores X foi definido a partir de variáveis disponíveis antes do início da inspeção, ou seja, ele não considera as variáveis temporais (tsetup, tpausa e tmedição). Essa escolha alinha-se ao princípio de evitar *leakage*¹ — isto é, impedir que informações ilegítimas sobre o alvo entrem na modelagem — prática recomendada em problemas preditivos para não inflar artificialmente o desempenho (Kaufman *et al.*, 2012). As variáveis categóricas foram tratadas como fatores e codificadas por *OneHotEncoder*² em um fluxo único de preparo e modelagem (*pipeline*³) (Buitinck *et al.*, 2013; Pedregosa *et al.*, 2011).

A escolha do subconjunto final de preditores adotou uma abordagem *wrapper*⁴: para combinações candidatas de variáveis numéricas e categóricas, estimou-se um *pipeline* (*one-hot* + regressão via Random Forest) e comparou-se o desempenho, restando a combinação com melhor ajuste. Métodos *wrapper* são clássicos em seleção de variáveis porque avaliam diretamente o valor preditivo do conjunto de características em pareamento com o algoritmo de

¹ Uso, direto ou indireto, de informação indisponível no momento da previsão (ou derivada do próprio alvo) na etapa de treino; isso infla artificialmente o desempenho e reduz a generalização (Kaufman *et al.*, 2012).

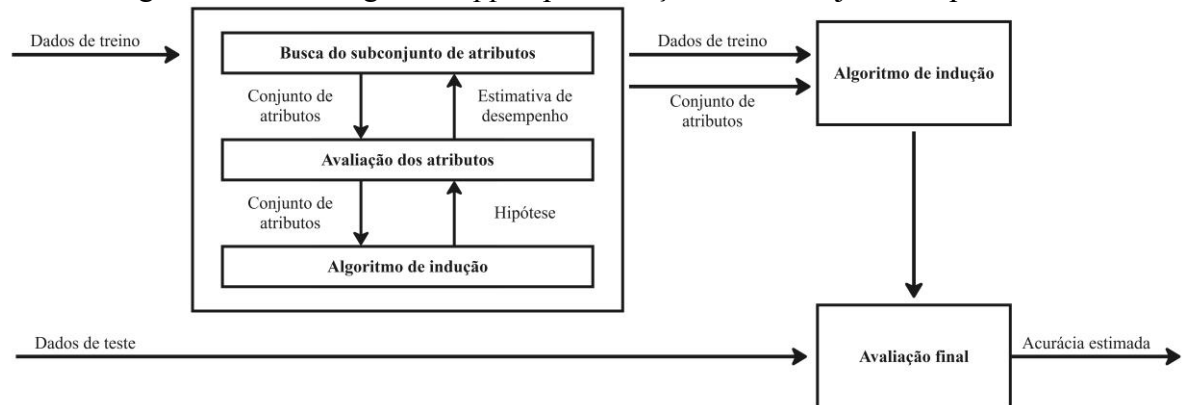
² Codificação 1-de-K: cada categoria vira uma coluna binária; evita ordenação espúria de rótulos em modelos numéricos (Buitinck *et al.*, 2013; Pedregosa *et al.*, 2011).

³ Objeto que encadeia pré-processamentos e modelo, garantindo que as transformações sejam ajustadas apenas nos dados de treino, reduzindo risco de vazamento (Buitinck *et al.*, 2013).

⁴ Estratégia de seleção de variáveis que avalia subconjuntos usando o próprio modelo via validação, restando os que melhoram o desempenho fora da amostra (Guyon; Elisseeff, 2003; Kohavi; John, 1997).

aprendizado, ao contrário de filtros univariados; a formulação e as vantagens desse desenho são discutidas por Kohavi e John (1997). O uso da Random Forest como estimador no *wrapper* justifica-se por sua capacidade de capturar não linearidades e interações entre preditoras com poucas suposições paramétricas (Breiman, 2001). Assim, a seleção privilegia conjuntos de variáveis que, em bloco, explicam melhor o tempo de inspeção, sem depender de engenharia extensiva de características.

Figura 11 – Abordagem *wrapper* para seleção de subconjunto de preditores



Fonte: Elaborado pelo autor, adaptado de Kohavi e John (1997)

A busca percorre subconjuntos que atendem a um limite mínimo de variáveis e, para cada combinação, o desempenho é estimado por validação cruzada repetida (*Repeated K-Fold*, 5×10), utilizando a média do R^2 como critério de seleção, procedimento recomendado para obter estimativas mais estáveis com amostras moderadas e para comparar modelos sob o mesmo protocolo (Kuhn; Johnson, 2013). Após selecionar o melhor subconjunto, o modelo é reajustado nos dados de treino completos e avaliado uma única vez no conjunto de teste externo, evitando o viés otimista decorrente de usar o teste durante a seleção (Varma; Simon, 2006). Todo o fluxo (codificação *one-hot* das categóricas, padronização quando aplicável e estimador) é encapsulado em *pipelines*, de modo que as mesmas transformações ajustadas no treino são aplicadas nas partições de validação e no teste, reforçando a comparabilidade e prevenindo *data leakage* (Kuhn; Johnson, 2013).

Após selecionar as variáveis, os algoritmos de ML são definidos baseado em outros estudos de aplicação de ML em processos (Çaklı; Dağdeviren, 2023; Cheng; Lin; Hu, 2025; Onaran; Yanık, 2020; Urban; Koblasa; Mendřický, 2024; Weller *et al.*, 2023). Para capturar relações não lineares entre os preditores e o tempo de inspeção, foram considerados modelos baseados em árvores de decisão e seus *ensembles*, além de métodos de margem/vizinhança, rede neural rasa e *boosting*. A Árvore de Decisão (CART) oferece regras “se-então”

interpretáveis e lida naturalmente com interações, mas apresenta alta variância quando muito profunda, exigindo poda e controle de complexidade (Çaklıt; Dağdeviren, 2023; Retzlaff *et al.*, 2024). Em contraste, as Random Forests reduzem a variância por agregação de múltiplas árvores, preservando a capacidade de modelar não linearidades e interações sem engenharia extensiva de atributos, sendo amplamente reconhecidas como uma alternativa robusta em dados tabulares, embora menos interpretáveis que uma árvore única (Breiman, 2001).

Como referência não paramétrica baseada em proximidade, o *k-Nearest Neighbors* (kNN) foi incluído pela simplicidade e desempenho aceitável em bases pequenas/médias; contudo, demanda padronização e tende a degradar com alta dimensionalidade após a codificação *one-hot*, fenômeno relacionado ao comportamento de métricas de distância em alta dimensão (Arora; Khanduja; Bansal, 2022).

Modelos de margem como a *Support Vector Regression* (SVR) oferecem uma abordagem eficaz para relações suaves e não lineares, mas são sensíveis ao ajuste de hiperparâmetros (C , ϵ , γ) e ao escalonamento dos preditores (Smola; Schölkopf, 2004). C é o parâmetro de regularização que define o trade-off entre maximizar a margem e penalizar desvios no ajuste — valores maiores punem mais o erro e tendem a margens menores, enquanto valores menores toleram mais erro para soluções mais suaves (Smola; Schölkopf, 2004). ϵ determina a largura do “tubo” ϵ -insensitive: desvios dentro de ϵ não geram perda, tornando a solução mais esparsa à medida que ϵ aumenta (Smola; Schölkopf, 2004). γ é o coeficiente do *kernel* Gaussiano que controla o “alcance” da influência de cada amostra: γ baixo produz funções mais suaves e de maior alcance; γ alto concentra influência local e pode aumentar a complexidade/*overfitting* (Hastie; Tibshirani; Friedman, 2009; Smola; Schölkopf, 2004). A título de sensibilidade adicional, uma rede neural rasa (MLP) pode aproximar funções complexas, sua inclusão visa comparar um classificador universal simples com métodos baseados em árvores.

Por fim, foram contemplados métodos de *gradient boosting* sobre árvores — XGBoost, LightGBM e CatBoost — dada sua reconhecida expressiva acurácia em dados tabulares e capacidade de capturar interações complexas. Estudos comparativos recentes mostram o desempenho competitivo dessas variantes e discutem compromissos de velocidade, generalização e *tuning* (Bentéjac; Csörgő; Martínez-Muñoz, 2021). CatBoost, em particular, propõe estratégias para mitigar o *prediction shift* e trata categorias nativamente, o que o torna adequado quando há muitas variáveis categóricas, mesmo quando se opta por *one-hot*, sua formulação reforça os benefícios do *boosting* em ambientes com categóricas (Prokhorenkova *et al.*, 2019).

Para assegurar comparabilidade entre algoritmos, as estimativas adotam a mesma formulação do problema: a variável-alvo é o tempo total de inspeção e o vetor de preditoras X permanece fixo, composto pelos mesmos atributos do item e do posto, eventos binários críticos e indicador de complexidade. O preparo dos dados é idêntico para os modelos, por meio de um único fluxo de pré-processamento com codificação *one-hot* das variáveis categóricas via *ColumnTransformer*, mantendo configuração e parâmetros de transformação constantes ao longo das estimações. O nível de agregação é global (*pooled*), incorporando Máquina e Família como variáveis de grupo no próprio X . Para evitar vieses de configuração, utilizam-se hiperparâmetros de base padronizados por família de algoritmo e a mesma convenção de controle de aleatoriedade para garantir reprodutibilidade. O código de implementação segue a lógica descritas nos parágrafos subsequentes.

Inicialmente, definiram-se a matriz de preditoras X e a variável-alvo y . O conjunto de dados foi particionado em treino (80%) e teste (20%), por amostragem aleatória (*random_state=42*). Adota-se um fluxo único de pré-processamento: as variáveis categóricas são codificadas por *one-hot encoding* via *ColumnTransformer*, e as numéricas são mantidas. Esse arranjo assegura consistência entre ajuste e predição, evita vazamento de informação e preserva a reprodutibilidade via *Pipeline* (Buitinck *et al.*, 2013; Kaufman *et al.*, 2012).

Quatro famílias de modelos são estimadas exclusivamente sobre o conjunto de treino, com seleção de hiperparâmetros por validação cruzada (*GridSearchCV*, 5 *folds*, métrica de seleção = *mean squared error*): Árvore de Regressão (CART), Random Forest, SVR e kNN. Para CART, a grade explora profundidade máxima e restrições mínimas de divisão/folha; para Random Forest, combinações de número de árvores, profundidade, fração de atributos e tamanho de folha; para SVR, grades separadas por *kernel* com variações de C , ϵ e γ ; e, para kNN, quantidade de vizinhos, esquema de pesos e métrica de distância. Nos métodos sensíveis à escala e à distância (SVR e kNN), aplicou-se *StandardScaler* apenas às variáveis numéricas. A padronização melhora a estabilidade desses algoritmos porque suas decisões dependem de distâncias/margens, sendo prática recomendada em *Support Vector Machines* (SVMs) e vizinhança no guia do *scikit-learn* e em textos de modelagem preditiva; as variáveis categóricas codificadas por *one-hot* não precisam ser reescaladas, sob pena de distorcer a representação binária (Buitinck *et al.*, 2013; Kuhn; Johnson, 2013).

Para a rede neural (ANN/MLP), adotou-se um *Pipeline* com *one-hot* denso para variáveis categóricas e padronização nas variáveis numéricas, seguido do *MLPRegressor*. A necessidade de padronização decorre da sensibilidade do MLP à escala dos preditores, recomendação explícita na documentação do *scikit-learn* para redes neurais supervisionadas

(Buitinck *et al.*, 2013). O uso de *one-hot* com saída densa assegura representação binária adequada das categorias e compatibilidade com estimadores que esperam matrizes densas. A justificativa para incluir uma MLP no conjunto de candidatos é que *perceptrons* multicamadas são aproximadores universais: com volume suficiente de neurônios e uma não linearidade adequada, podem aproximar funções contínuas arbitrárias, conforme os resultados de Cybenko (1989) e Hornik (1991).

Os métodos de *boosting* baseados em árvores foram integrados via API compatível com *scikit-learn* e operam diretamente sobre a matriz produzida pelo *ColumnTransformer*, aceitando entradas esparsas resultantes do *one-hot encoding*; a configuração dos modelos seguiu grades que combinam número de árvores, profundidade, learning rate e frações de amostragem/atributos, conforme documentação oficial dessas bibliotecas (Buitinck *et al.*, 2013). Para LightGBM, dadas bases pequenas com muitas *dummies*, ajustaram-se hiperparâmetros que reduzem barreiras de divisão e aumentam a resolução de histogramas — por exemplo, *min_data_in_leaf* baixo, *min_gain_to_split* = 0 e *max_bin* elevado — de acordo com as diretrizes do manual de parâmetros e de *tuning* (Shi *et al.*, 2025). O CatBoost foi estimado com *one-hot* denso, e a busca contemplou iterações, profundidade, *learning rate* e regularização L2, parâmetros descritos na referência oficial (Prokhorenkova *et al.*, 2019). Para contextualizar o uso conjunto dessas variantes, revisões comparativas na literatura sintetizam suas diferenças de desempenho, velocidade e sensibilidade a hiperparâmetros em dados tabulares (Bentéjac; Csörgő; Martínez-Muñoz, 2021).

Após a seleção via *GridSearchCV*, cada estimador foi reajustado no conjunto de treino completo e avaliado no conjunto de teste comum, reportando-se R^2 e RMSE, métricas discutidas em textos de referência de Hastie *et al.* (2009) e James *et al.* (2014). Para o melhor algoritmo, as importâncias de atributos são agregadas do nível das *dummies* para o nível das variáveis originais, somando-se as contribuições das colunas derivadas de um mesmo preditor; essa prática evita viés a favor de variáveis com muitos níveis e está alinhada às recomendações de literatura sobre vieses de importância com diferentes escalas/número de categorias (Strobl *et al.*, 2007). Por fim, são gerados gráficos de $RMSE \times R^2$ e diagramas Previsto vs. Real, conforme sugerido em guias de modelagem preditiva para avaliação de modelos de regressão (Kuhn; Johnson, 2013).

Depois de estimar os tempos normais de inspeção utilizando ML, busca-se entender quais fatores elevam os tempos de inspeção. Para isso, emprega-se o *A priori*, algoritmo clássico para mineração de regras de associação: ele descobre *itemsets* frequentes por meio do princípio antimonotônico (se um conjunto é frequente, todos os seus subconjuntos também são)

e, a partir deles, gera regras $A \Rightarrow B$ que relacionam coocorrências em dados transacionais (Agrawal; Srikant, 1994). A análise de associação é apresentada como técnica não supervisionada para extrair relações entre itens/atributos, com definições formais para suporte, confiança e *lift*, e discussão de como escolher limiares e interpretar regras (Agrawal; Srikant, 1994; Han; Kamber; Pei, 2011).

Cada regra é apresentada com suporte, confiança e *lift*. Suporte é a frequência com que um padrão aparece nos dados: quantas observações exibem simultaneamente as condições do antecedente e o desfecho da regra. Trabalhar com um limiar mínimo de suporte é prática canônica porque filtra padrões raros e garante relevância operacional (Agrawal; Srikant, 1994). Confiança responde à pergunta: “quando o antecedente acontece, com que frequência o consequente também ocorre?”. Trata-se de uma medida condicional de acerto da regra no subconjunto em que as condições do antecedente estão presentes (Han; Kamber; Pei, 2011). *Lift* qualifica a força da associação, se o *lift* é maior que 1, a presença do antecedente está associada a um aumento da chance do consequente acima do que se esperaria ao acaso; valores próximos de 1 sugerem nenhuma associação relevante e valores menores que 1 indicam associação negativa (Agrawal; Srikant, 1994; Han; Kamber; Pei, 2011).

No conjunto analisado, o objetivo é identificar combinações de condições associadas a diferentes faixas de duração. Para isso, definiram-se três alvos: tempo de inspeção $\geq P50$, tempo de inspeção $\geq P75$ e tempo de inspeção $\geq P95$. A variável NMedidas foi discretizada em bandas (“1–3”, “4–6”, “7–9” e “Maiores que 9”) para reduzir cardinalidade e estabilizar coocorrências. Em cada alvo, mineraram-se *itemsets* frequentes com tamanho máximo de 3 itens e limiares específicos, suporte mínimo de 20% e confiança $\geq 65\%$ para $\geq P50$; suporte mínimo de 4% e confiança $\geq 65\%$ para $\geq P75$; e suporte mínimo de 1% e confiança $\geq 50\%$ para $\geq P95$. As regras foram ranqueadas por *lift* e passaram por filtragem de redundância (remoção de superconjuntos que não agregam confiança). Os resultados obtidos são apresentados na seção 4.3.

Após avaliar os fatores responsáveis pelos tempos normais mais elevados, transformase o tempo normal obtido para cada inspeção em um tempo-padrão utilizável pelo PCP. Conceitualmente, o tempo normal (t_n) é o tempo observado para executar a tarefa ajustado pelo fator de ritmo do operador; isto é, o tempo cronometrado (ou previsto) multiplicado por uma avaliação de desempenho (Barnes, 1977). No presente estudo, como não havia inspetores suficiente por máquina para estimar avaliações de ritmo confiáveis (menos de 1 inspetor por máquina), e porque os tempos normais usados provêm de um modelo preditivo treinado sobre histórico, desconsiderou-se o ajuste de ritmo do operador. A partir do tempo normal, obtém-se

o tempo-padrão ($t_{padrão}$) adicionando-se as tolerâncias para necessidades pessoais (T_p), fadiga (T_f) e esperas inevitáveis (T_{ei}), prática clássica na literatura de métodos e medição do trabalho (Freivalds; Niebel, 2014; Silva, A V; Coimbra, 1980).

$$t_{padrão} = t_n + (t_n \times T_{total}) \quad (2)$$

Para estimar as tolerâncias adotadas no cálculo dos tempos-padrão, parte-se do modelo de *allowances* da engenharia de métodos, descrito pela equação a seguir:

$$T_{total} = T_p + T_{ei} + T_f \quad (3)$$

Adotou-se tolerância pessoal de 5% conforme indicado por Freivalds e Niebel (2014) e tolerância por fadiga $T_f = (A \times B) + M + CA$, com $A = EF + EM$, onde EF é esforço físico, EM esforço mental, M monotonia e CA fadiga (Silva, A V; Coimbra, 1980).

Tabela 2 – Fatores de esforço físico e mental

ESFORÇO FÍSICO		ESFORÇO MENTAL	
GRAU	%	GRAU	%
Muito Leve (ML)	1,8%	Leve (L)	0,6%
Leve (L)	3,6%	Médio (M)	1,8%
Médio (M)	5,4%		
Pesado (P)	7,2%	Pesado (P)	3,0%
Muito Pesado (MP)	9,0%		

Fonte: adaptado de Silva, Coimbra (1980) e (Gawlik *et al.*, 2023)

A Tabela 2 apresenta os percentuais de tolerância por fadiga associados ao esforço exigido na tarefa, separando componentes físico e mental. Para cada dimensão, escolhe-se o grau observado (p. ex., “Médio” ou “Pesado”). Valores maiores indicam maior exigência e, portanto, maior tolerância necessária. No esforço físico, os percentuais vão de 1,8% (Muito Leve) a 9,0% (Muito Pesado); no esforço mental, de 0,6% (Leve) a 3,0% (Pesado).

Tabela 3 – Fatores de recuperação

RECUPERAÇÃO	
% TEMPO RECUPERADO	FATOR "B"
0 - 5	1,00
6 - 10	0,90
11 - 15	0,80
16 - 20	0,71
21 - 25	0,62
26 - 30	0,54
31 - 35	0,46
36 - 40	0,39
41 - 45	0,32
46 - 50	0,26
51 - 55	0,20
56 - 60	0,15

Fonte: adaptado de Silva e Coimbra (1980)

A Tabela 3 relaciona a parcela do turno destinada à recuperação/pausas (% de tempo recuperado) ao fator “B”, que corrige a capacidade efetiva de trabalho. Quanto maior o tempo reservado à recuperação, menor o fator B, refletindo redução do tempo produtivo.

Tabela 4 – Fatores de monotonia

MONOTONIA	
DURAÇÃO DO CICLO (MIN.)	%
de 0 a 0,05	7,8%
de 0,06 a 0,25	5,4%
de 0,26 a 0,50	3,6%
de 0,51 a 1,00	2,1%
de 1,01 a 2,00	1,0%
de 2,01 a 3,00	0,5%
de 3,01 a 4,00	0,2%

Fonte: adaptado de Silva e Coimbra (1980)

A Tabela 4 descreve os acréscimos por monotonia aplicados conforme a duração do ciclo. Quanto mais curto é o ciclo, maior a parcela adicionada ao tempo normal para compensar os efeitos de repetição e atenção contínua. Assim, ciclos entre 0 – 0,05 min recebem 7,8%, enquanto ciclos mais longos, como 3,01 – 4,00 min, recebem 0,2%.

Tabela 5 – Fatores para condições ambientais

CONDIÇÕES AMBIENTAIS						
TÉRMICAS		ATMOSFÉRICAS		OUTRAS		
TEMPERATURA	%	DESCRIÇÃO	%	TIPO	DESCRIÇÃO	%
Gelada, de 0 a 7 °C	3,6%	Boas, local bem ventilado ou ar fresco	0,0%	RUÍDO	Baixo nível	0,0%
Baixa, de 7 a 15 °C	1,8%	Razoáveis, local mal ventilado, presença de mau cheiro ou fumaça não-tóxica	2,4%	RUÍDO	Excessivo, obrigando ao uso de protetor auricular	1,8%
Normal, de 15 a 26 °C	0,0%			UMIDADE	Ambiente seco e agradável	0,0%
Alta, de 26 a 34 °C	1,8%	Más, alta concentração de pós. Presença de fumaça ou pó tóxicos. Uso obrigatório de máscara facial.	5,6%	UMIDADE	Excessivo até 25 °C	1,8%
Excelente, de 34 a 40 °C	3,6%			UMIDADE	Excessivo até 40 °C	3,6%
				VIBRAÇÃO	Vibração do solo ou máquina	1,8%

Fonte: adaptado de Silva e Coimbra (1980)

A Tabela 5 define acréscimos por condições ambientais a serem somados ao tempo normal conforme três frentes: térmicas (faixas de temperatura), atmosféricas (qualidade do ar/ventilação) e outras (ruído, umidade e vibração). Na prática, identifica-se o cenário de trabalho e aplica-se o percentual correspondente em cada frente relevante. Valores de referência “normais” não acrescentam tolerância.

No laboratório de metrologia da Empresa Z, os calibradores manuseados pelo inspetor são leves (em geral ≤ 500 g e, nos casos mais pesados, ≈ 1 kg) e o deslocamento dessas peças é pontual e muito mais curto que o próprio tempo de medição. Em termos ergonômicos, tarefas com levantamento e transporte sistemático bem inferiores a $\sim 4,5$ kg se enquadram em níveis “leves/sedentários” de exigência, o que sustenta a decisão de considerar desprezível o componente de esforço físico para fins de tolerância (Gawlik *et al.*, 2023). Por isso, o esforço físico foi classificado como muito leve.

Em contrapartida, o esforço mental é elevado. Nas máquinas automáticas, o inspetor precisa programar a rotina de medição e acompanhar a execução para evitar contatos indevidos entre apalpador e peça; nas máquinas manuais, há alinhamentos finos e operação cuidadosa de instrumentos, pontos reconhecidos na literatura como fontes de fadiga cognitiva (atenção contínua, tensão visual, tomada de decisão) (Freivalds; Niebel, 2014; Silva, A V; Coimbra, 1980). Assim, adotou-se esforço mental pesado (P). Como os ciclos são relativamente longos e

variados, não se caracteriza repetição monótona de microtarefas, portanto, a monotonia é considerada nula (Silva, A V; Coimbra, 1980). Finalmente, o ambiente do laboratório é rigidamente controlado (temperatura, umidade, ruído e vibração dentro de faixas ideais), prática exigida por normas de acreditação como a ISO/IEC 17.025, razão pela qual as condições ambientais também são tratadas como nulas (ISO, 2017).

No presente estudo não foi atribuída tolerância de “esperas inevitáveis” porque, ao longo do período analisado, não foram observadas paradas aleatórias fora do controle do inspetor (p. ex., quebras, falta de insumos, interrupções externas) (Das, 1990; Silva, A V; Coimbra, 1980). Em paralelo, há evidência de que laboratórios que implementam Lean reduzem filas, esperas internas e *turnaround time* (TAT), reduzindo a incidência de atrasos sistêmicos e converge com a realidade deste laboratório metrológico (fluxo estabilizado e poucas interrupções) (Inal *et al.*, 2018). Quando tais atrasos existem, a recomendação é quantificá-los objetivamente (por exemplo, por *work sampling*) e incorporá-los como tolerância apenas mediante evidência, em vez de assumir valores a priori (Das, 1990).

O fator de recuperação (B) representa a fração do ciclo em que o operador dispõe de “tempo de descanso” natural porque a máquina executa autonomamente; por isso ele escala a fadiga: quanto menor a recuperação, maior o B (até $B = 1$) quando praticamente não há janelas de descanso) (Silva, A V; Coimbra, 1980). Para atribuir o fator por equipamento, primeiro cronometrou-se, por ciclo, o trecho em que a máquina executa a medição sem intervenção ativa do inspetor (tempo de recuperação) e comparou-se com o tempo total de medição (tempo normal), obtendo o percentual de recuperação por máquina. Em seguida, cada percentual é classificado nas faixas de recuperação (0–5%, 6–10%, 11–15%, 16–20%... 56–60%) e convertido no respectivo fator B, conforme a Tabela 3.

Em operações assistidas, nas quais o inspetor precisa permanecer atento mesmo durante rotinas automáticas para prevenir contatos indevidos e intervir em anomalias, as atividades são essencialmente de supervisão/vigilância, de alta carga cognitiva e com recuperação mínima; nesses casos, adotou-se $B = 1$ (Máquinas 1, 2, 6, 7, 8 e 9). Quando há blocos automáticos longos sem necessidade de vigilância direta (p.ex., varreduras/traçados programados), ocorrem janelas claras de recuperação e $B < 1$; isso é típico de inspeções ópticas/automatizadas e de rotinas para CMMs configuradas para operação autônoma. Assim, as Máquinas 5, 4 e 3 receberam, respectivamente, $B = 0,71$, $B = 0,46$ e $B = 0,39$, refletindo percentuais crescentes de execução automática sem acompanhamento contínuo.

Quadro 10 – Síntese dos fatores para Cálculo das Tolerâncias

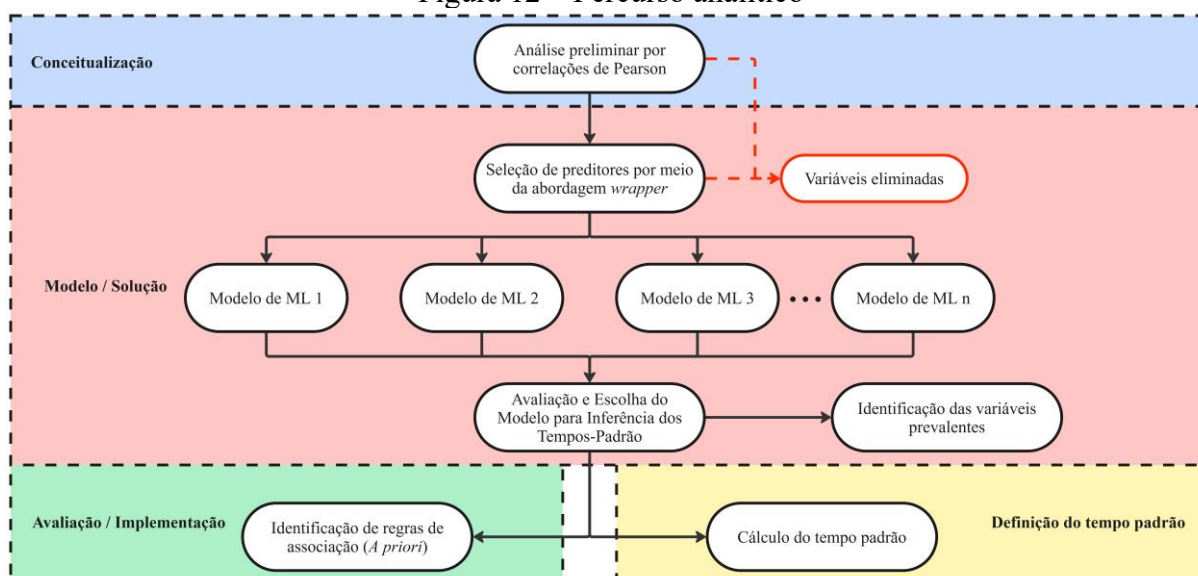
Fator	Valor	Justificativa
EF	Muito Leve (1,8%)	Peças ≤ 1 kg e deslocamentos breves;
EM	Pesado (3%)	Programação e monitoramento contínuo nas máquinas automáticas, alinhamentos finos nas máquinas manuais.
Monotonia	Nulo (0%)	Ciclos longos e variados; sem repetição cíclica curta.
CA	Nulo (0%)	Laboratório climatizado segundo ISO/IEC 17.025.
Espera Inevitável	Nulo (0%)	Não evidenciada durante observações.
Fator B	B = 1 (Máquinas 1, 2, 6, 7, 8 e 9); B = 0,71 (máquina 5); B = 0,46 (máquina 4) e B = 0,39 (máquina 3).	Proporção de tempo de recuperação, obtido por cronometragem.

Fonte: Elaborado pelo autor

Após definir as tolerâncias e o modelo adequado, gera-se previsões para o cenário de referência, para então converter o tempo normal previsto em tempo-padrão estimado. Com o modelo adequado treinado novamente sobre o conjunto integral, geram-se previsões para o histórico. Em seguida organiza-se um *dataframe* dedicado às predições, contendo variáveis de processo (Máquina, Família, NMedidas etc.). Para refletir as condições padrão de programação e operação, filtra-se o conjunto para PP = 1, MP = 0, NConformidade = 0 e Setup Extra = 0. Essa etapa almeja garantir que os tempos derivados representem o cenário de referência que será adotado em PCP/ERP, evitando que condições eventuais inflacionem os parâmetros. As predições do modelo são tratadas como tempos normais. O tempo-padrão é calculado a partir da (2). Para uso gerencial, os tempos-padrão são sumarizados por Máquina, e por Máquina $\times$ faixas de NMedidas para os percentis P50 (mediana), P75 e P95.

A Figura 12 sintetiza o percurso metodológico desta etapa: dos dados preparados, passa-se a uma triagem exploratória por correlações, seguida da seleção de preditores via abordagem *wrapper* e da comparação entre modelos de ML sob o mesmo pré-processamento. O modelo escolhido é então fixado e suas variáveis prevalentes são registradas. Assim, encerra-se a modelagem com um pipeline claro, reproduzível e pronto para gerar previsões consistentes. A Seção 4 parte desse resultado: converte as previsões em tempo-padrão e explora padrões operacionais por meio de regras de associação.

Figura 12 – Percurso analítico



Fonte: Elaborado pelo autor

Para avaliar a aderência prática e a credibilidade dos resultados quantitativos do estudo, foi conduzida uma entrevista individual com dois profissionais diretamente envolvidos na gestão e na operação do processo de metrologia da Empresa Z, com o objetivo exclusivo de avaliar a utilidade prática e a aplicabilidade dos resultados do estudo. Participaram uma planejadora de qualidade, com formação em Engenharia Mecânica, e um gestor de qualidade, formado em Direito e com pós-graduação em Logística. As entrevistas ocorreram virtualmente, com consentimento prévio para registro das respostas, preservação de anonimato no relatório e uso exclusivamente acadêmico e interno das informações. A adoção de entrevistas se justifica por sua adequação para captar percepções, raciocínios e justificativas daqueles que vivenciam o processo, contribuindo para explicar, qualificar e ampliar a interpretação dos achados numéricos — um uso clássico de integração em métodos mistos (Creswell, 2009).

Os participantes foram selecionados por amostragem intencional, a fim de reunir informantes-chave com visão técnica e gerencial diretamente relacionada ao fenômeno estudado. O tamanho do grupo foi definido por suficiência informacional combinando a noção de saturação — ausência de novos elementos relevantes após sucessivas entrevistas — com o princípio de “*information power*”, segundo o qual quanto mais específico e informativo for o escopo/participante, menor pode ser o número de entrevistas necessário (Guest; Bunce; Johnson, 2006). As conversas foram registradas em áudio e transformadas em material textual para análise posterior.

Quadro 11 – Questões para avaliação qualitativa

Questões	Justificativa
Como o estudo contribuiu para a empresa?	Esclarece utilidade prática e “valor de uso” dos resultados. Permite verificar se tempos-padrão e achados quantitativos apoiam decisões de planejamento, custos e programação, finalidades clássicas de estudos de tempos e de mensuração do trabalho (Kanawaty, 1992).
Quais foram os benefícios de obter os tempos-padrão?	Perguntar por melhoria em métricas conecta o estudo ao desenho e uso de sistemas de medição de desempenho (Neely; Gregory; Platts, 1995).
Como foi a implementação dos resultados da pesquisa?	Implementação efetiva de métricas exige desenho, adoção e rotinas de uso; perguntar pelo “funcionamento” e pelo conjunto de indicadores verifica aderência ao ciclo de medição-gestão (Neely; Gregory; Platts, 1995).
Quais foram as principais dificuldades na implementação?	A pergunta serve para identificar determinantes/barreiras que afetam o sucesso da adoção em múltiplos domínios, como recursos, cultura/clima organizacional, complexidade percebida do método, engajamento de pessoas e condução do processo (Damschroder <i>et al.</i> , 2022).
A publicação dos tempos-padrão influenciou novas decisões no setor?	Tempos-padrão suportam balanceamento de recursos e previsão de capacidade; verificar impacto em dimensionamento e <i>throughput</i> confirma uso típico desses tempos (Kanawaty, 1992).

Fonte: Elaborado pelo autor

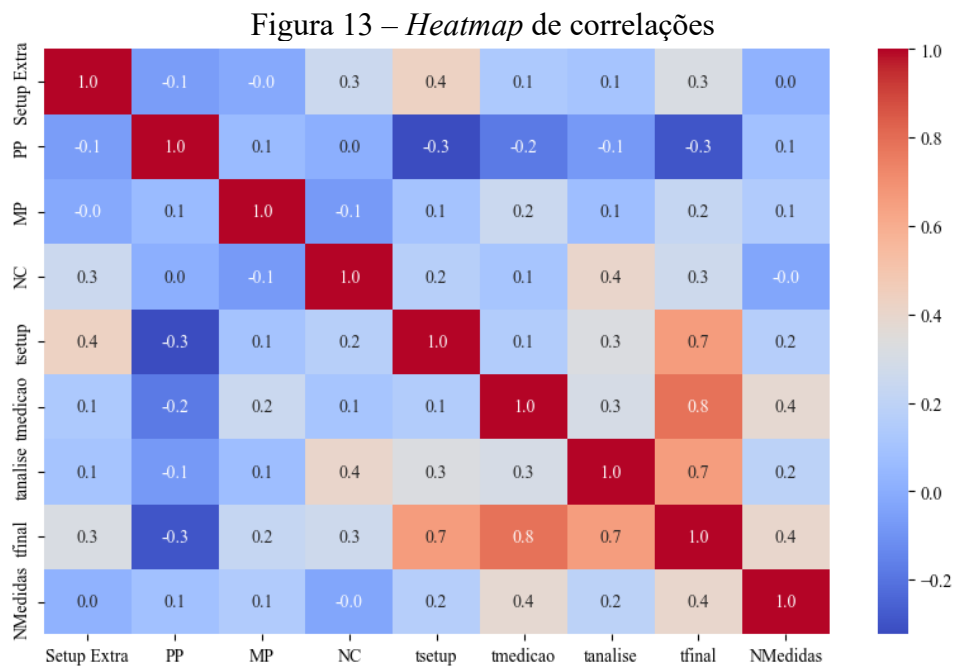
Definidos o desenho metodológico, o *pipeline* de dados, os critérios de formação do tempo-padrão e as tolerâncias estabelecidos, além dos procedimentos de validação e consolidação para publicação, a Seção 4 apresenta os achados empíricos do estudo.

4 ANÁLISE DOS RESULTADOS

Esta seção organiza os resultados em três frentes complementares. Primeiro, compara-se o desempenho dos algoritmos de ML para identificar o estimador com melhor capacidade preditiva e estabilidade, além de interpretar os preditores mais relevantes para o tempo de inspeção. Em seguida, aplicam-se regras de associação para descobrir combinações de condições operacionais que elevam a probabilidade de inspeções longas, reportando suporte, confiança e *lift* para priorização de ações. Por fim, convertem-se as previsões geradas no cenário de referência em tempos-padrão por máquina e por faixas de NMedidas, apresentados por percentis (P50, P75, P95) para apoiar decisão de planejamento e capacidade.

4.1 AVALIAÇÃO DOS ALGORITMOS DE MACHINE LEARNING

Realizou-se uma análise exploratória de dados antes para verificar e potencialmente eliminar algum dos preditores. Os resultados da análise são apresentados por meio de um *Heatmap* de correlações onde valores em vermelho indicam correlação positiva ($r = 1$), valores em azul correlação negativa ($r = -1$), e magnitudes próximas de $r = 0$ sugerem associação fraca.



Fonte: Elaborado pelo autor (2025)

O mapa de calor sintetiza correlações de Pearson entre variáveis de processo e tempos. Observa-se que não há pares com $|r| \approx 1$, portanto não existe colinearidade perfeita no conjunto.

Isso é desejável porque colinearidade elevada/estrutural pode tornar estimativas instáveis e pouco interpretáveis, enquanto níveis moderados permitem combinar sinais sem inflar incerteza (Dormann *et al.*, 2013). Ao mesmo tempo, é importante lembrar que Pearson mede apenas associação linear bivariada; ausência de correlação forte não implica ausência de dependência, que relações monotônicas não lineares ou efeitos com pontos influentes podem ser subestimados pela correlação (Schober; Boer; Schwarte, 2018).

Do ponto de vista preditivo, a inexistência de correlações perfeitas sugere complementaridade de informação entre preditores — algo que modelos modernos exploram ao captar não linearidades e interações entre variáveis (além do que a matriz de correlação identifica). Essa capacidade é intrínseca aos métodos de árvores e seus *ensembles* (*Random Forests*) e aos *boostings* de gradiente, que constroem funções aditivas de árvores para modelar efeitos complexos (Breiman, 2001; Natekin; Knoll, 2013). Na seleção/avaliação de variáveis, a literatura também alerta que relevância pode emergir apenas em conjunto (interações) e não aparecer em medidas marginais como a correlação; logo, convém avaliar utilidade no contexto do modelo e não só por filtros univariados (Guyon; Elisseeff, 2003).

Note-se que as variáveis *tsetup*, *tmedicao* e *tanalise* são componentes do próprio tempo final. Portanto, não estão disponíveis no momento de previsão e, se incluídas como preditores, introduziriam vazamento de informação (*data leakage*), ou seja, o modelo passaria a “antecipar” o alvo usando partes do próprio alvo, produzindo métricas artificialmente otimistas e inadequada generalização fora da amostra (Kaufman *et al.*, 2012; Shmueli, 2010). A partir dessa análise, essas variáveis foram excluídas do treinamento, que emprega preditores de contexto e planejamento (Máquina, Família, NMedidas, PP, MP, NC, Setup Extra), alinhados ao cenário de uso real do modelo.

Após realizar a EDA, é realizada uma análise *wrapper* para selecionar os preditores. O melhor modelo de busca de subconjuntos obteve $R^2 = 0,455$ para o conjunto de teste, indicando que o conjunto de preditoras selecionado explica cerca de 45,5% da variabilidade observada no tempo de inspeção. O vetor *X* selecionado inclui, entre as numéricas/binárias, *Setup Extra*, PP, MP, NC e NMedidas; e, entre as categóricas, Máquina e Família.

Quadro 12 – Preditores selecionados pelo método *wrapper* ($R^2 = 0,455$)

Variável	Tipo	Hipótese
Setup Extra	Numérica/binária	Evento condicional de preparação, causa aumento no tempo de inspeção.
PP	Numérica/binária	Programa pronto, se verdadeiro, diminui o tempo de inspeção.
MP	Numérica/binária	Problema de máquina, se verdadeiro, aumenta o tempo de inspeção.

Variável	Tipo	Hipótese
NC	Numérica/binária	Ocorrência de não conformidade, se verdadeiro, aumenta o tempo de inspeção.
NMedidas	Numérica	Nº de características medidas cresce proporcionalmente ao tempo de medição.
Máquina	Categórica	Máquinas diferentes realizam a inspeção em tempos diferentes.
Família	Categórica	Calibradores de famílias diferentes possuem tempos de inspeção diferentes.

Fonte: Elaborado pelo autor

Depois de selecionar os preditores, verifica-se o quão bem cada algoritmo consegue capturar as relações e estimar os tempos. Reporta-se R^2 e RMSE em minutos para comparar acurácia e generalização. A Tabela 6 resume os resultados.

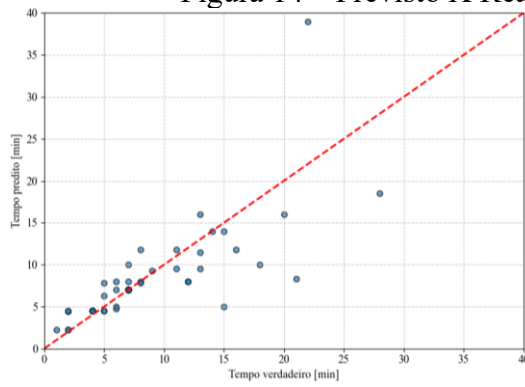
Tabela 6 – Resultados para cada algoritmo de ML

Algoritmo	R^2	RMSE
Decision Tree	0,477	4,132
Random Forest	0,572	3,734
SVR	0,442	4,266
kNN	0,505	4,017
ANN	0,477	4,132
XGBoost	0,508	4,006
LightGBM	0,281	4,844
CatBoost	0,592	3,647

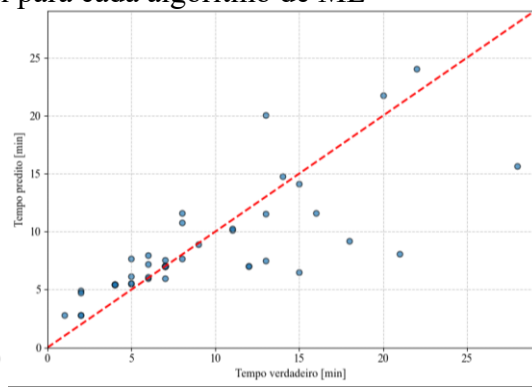
Fonte: Elaborado pelo autor

A Tabela 6 resume o desempenho dos algoritmos testados usando duas métricas de erro e ajuste, R^2 (quanto maior, melhor explicação da variabilidade) e RMSE (quanto menor, melhor precisão). Essa comparação serve para escolher o modelo a ser utilizado nas análises subsequentes.

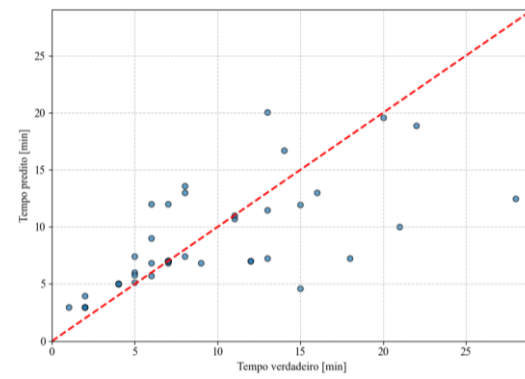
Figura 14 – Previsto X Real para cada algoritmo de ML



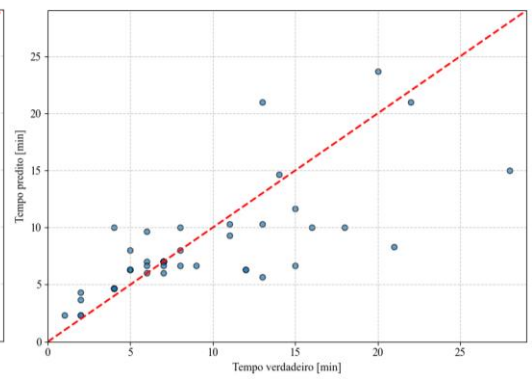
(a) – Decision Tree



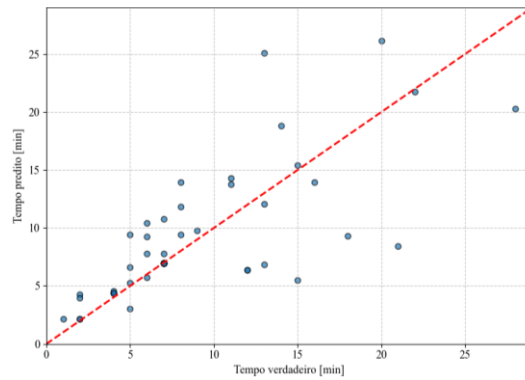
(b) – Random Forest



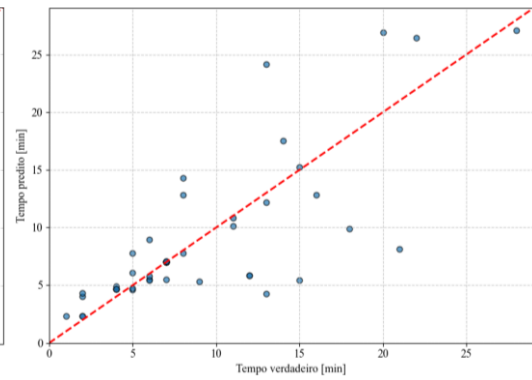
(c) – SVR



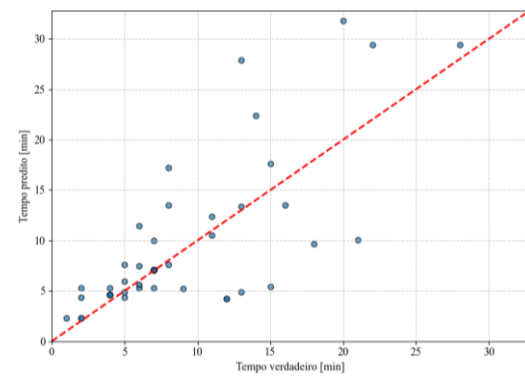
(d) - kNN



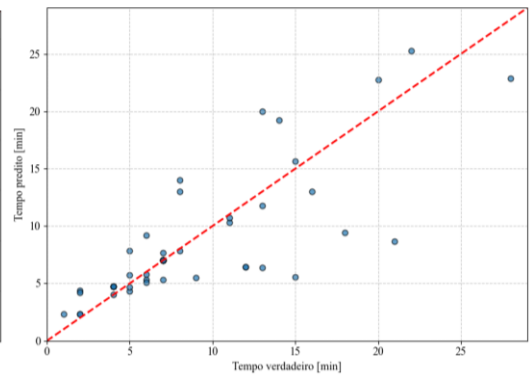
(e) – ANN



(f) - XGBoost



(g) – Light GBM



(h) - CatBoost

No geral, os resultados evidenciam vantagem consistente dos *ensembles* baseados em árvores, com CatBoost no topo ($R^2 = 0,592$; RMSE = 3,647 min), seguido de Random Forest ($R^2 = 0,572$; RMSE = 3,734 min) e XGBoost ($R^2 = 0,508$; RMSE = 4,006 min). A diferença entre CatBoost e Random Forest é reduzida ($\approx 2,3\%$ de redução no RMSE), mas suficiente para selecionar como modelo adequado, sobretudo pela adequada tratativa de não linearidades, interações e categóricas. Modelos de referência como SVR, ANN e Decision Tree ficaram atrás (RMSE $\approx 4,1$ – $4,3$ min), o que é compatível com a natureza mista dos dados e o porte da amostra. O LightGBM teve desempenho inferior nesta base ($R^2 = 0,281$), possivelmente por sensibilidade maior à configuração/volume de dados após o *one-hot*. Com base nisso, CatBoost é adotado para gerar as previsões operacionais, com Random Forest como alternativa de referência.

Para interpretar quais preditores mais contribuíram para explicar o tempo de inspeção, calculou-se a importância das variáveis a partir do melhor algoritmo (CatBoost). As variáveis de prevalentes foram explicitadas na Tabela 7.

Variável	Importância
Quantidade de características medidas (NMedidas)	30,62%
Máquina	19,28%
Programa pronto? (PP)	17,63%
Há não conformidade? (NC)	11,52%
Família	11,42%
Setup Extra	5,00%
Problema na máquina? (MP)	4,54%

Fonte: Elaborado pelo autor

Os resultados indicam que a quantidade de características medidas (NMedidas) é o principal determinante do tempo (30,62%). O resultado é coerente com a lógica de carga de trabalho, quanto mais características a medir, maior a quantidade de posicionamentos, trocas de apalpador/rotina e registros, elevando o esforço total. Operacionalmente, recomenda-se segmentar por faixas de NMedidas e usar esse preditor como eixo de planejamento de capacidade.

A Máquina aparece com importância elevada (19,28%), sinalizando diferenças sistêmicas entre equipamentos (cinemática, estratégias de medição, *setups* e estabilidade). Isso justifica a parametrização por máquina na etapa de tempos-padrão e o uso do fator de recuperação B específico por equipamento. O Programa Pronto (PP) se destaca (17,63%),

indicando que ter rotina disponível é um fator importante para obter tempos menores. O resultado sustenta governança de biblioteca de programas, atualização e padronização de rotinas como ações diretas de redução de tempo.

Os indicadores de NC (11,52%) e Família (11,42%) mantêm papel relevante. NC reflete o impacto típico de retrabalho (re-medições, registros e segregação), enquanto Família captura traços geométricos/tecnológicos compartilhados que influenciam a rota e a complexidade da medição. Mesmo que, para definir tempos-padrão operacionais, o cenário de referência exclua NC, a sua importância ajuda a explicar a variabilidade observada no histórico.

Por fim, *Setup* Extra (5,00%) e MP (4,54%) têm contribuição menor, porém não desprezível: eventos extraordinários e instabilidades de equipamento elevam tempo quando ocorrem, o que recomenda monitoramento e *checklists* de *setup*/manutenção para reduzir incidência.

4.2 INFERÊNCIA DOS TEMPOS-PADRÃO

Esta subseção apresenta os tempos-padrão por máquina obtidos a partir das previsões do modelo no cenário de referência descrito na seção 3.4.

Tabela 8 – Tempos-padrão médios por máquina

Máquina	$\bar{t}_n$	Tolerâncias (T_{total})	$\bar{t}_{padrão}$
Máquina 1	4,08	9,80%	4,48
Máquina 2	7,21	9,80%	7,92
Máquina 3	13,25	6,87%	14,16
Máquina 4	7,27	7,21%	7,79
Máquina 5	7,61	8,41%	8,25
Máquina 6	4,10	9,80%	4,50
Máquina 7	6,00	9,80%	6,59
Máquina 8	5,83	9,80%	6,41
Máquina 9	3,68	9,80%	4,04

Fonte: Elaborado pelo autor, tempos em minutos

A Tabela 8 apresenta, para cada máquina, o tempo normal médio previsto ($\bar{t}_n$), o acréscimo percentual de tolerâncias (T_{total}) e o tempo-padrão médio resultante ($\bar{t}_{padrão}$). Valores próximos de 9,8% indicam recuperação mínima (vigilância contínua; $B \approx 1$), ao passo que acréscimos menores, evidenciam trechos autônomos sem necessidade de acompanhamento direto.

O maior tempo médio ocorre na Máquina 3, com valor substancialmente superior ao das demais. A principal razão é o escopo mais amplo de aplicação: é uma máquina capaz de, em uma única inspeção, avaliar mais de 10 características, o que eleva o tempo típico e amplia a complexidade dos ciclos. Em contraste, as Máquinas 1, 7 e 8, por serem manuais, são preferencialmente empregadas em peças com menor número de características, pois itens mais complexos tendem a ser encaminhados às máquinas automáticas, que executam rotinas mais eficientes para cargas maiores de medição. Por fim, a Máquina 9, embora automática, é especializada: costuma medir um único tipo de característica, operando em faixas de tempo menores e com variabilidade reduzida.

Tabela 9 – Tempos-padrão por máquina e características medidas

Máquina	NMedidas	$\bar{t}_{padr\tilde{a}o}$ (P50)	$\bar{t}_{padr\tilde{a}o}$ (P75)	$\bar{t}_{padr\tilde{a}o}$ (P95)
1	1-3	4,48	4,48	4,48
2	1-3	8,48	8,49	8,74
2	4-6	12,50	12,50	12,50
3	1-3	8,89	17,15	19,61
3	4-6	11,55	12,76	16,39
3	7-9	15,68	15,68	15,68
3	≥ 10	25,02	25,02	25,02
4	1-3	6,96	10,11	10,11
5	1-3	8,44	8,87	8,87
5	4-6	5,84	5,97	6,07
5	7-9	6,05	12,72	18,05
6	1-3	4,13	5,10	5,67
6	4-6	11,93	11,93	11,93
7	1-3	6,74	7,72	7,72
8	1-3	6,93	8,17	8,65
9	1-3	4,04	4,04	4,04

Fonte: Elaborado pelo autor, tempos em minutos

A Tabela 9 sintetiza os tempos-padrão por máquina e por faixa de NMedidas, reportando três níveis de serviço: P50, P75 e P95 (patamar conservador). Observa-se o comportamento esperado de elevação do tempo conforme cresce NMedidas em várias máquinas — sobretudo nas Máquinas 3, 6 e 7 —, sinal de que a carga de características é um direcionador efetivo do tempo de inspeção. Em máquinas com menor variabilidade ($P50 \approx P75 \approx P95$), os padrões são estáveis e previsíveis.

Além disso, algumas máquinas exibem dispersão relevante, indicando heterogeneidade entre peças ou rotinas mesmo no cenário de referência. Destacam-se: Máquina 2, em que o P95 praticamente dobra o P50; Máquina 3 (1–3 e 4–6), que combina patamar alto com cauda longa; Máquina 4 (1–3), com casos esparsos mais complexos; e Máquina 5 (4–6), onde o P95 é substancialmente superior ao P50. Em termos de publicação do padrão, o uso do P75 por célula oferece cobertura próxima de 75% das ordens, enquanto o P95 pode ser adotado para itens críticos. A diferença entre P95 e P50 fornece, ainda, um indicador direto de variabilidade operacional a ser priorizada em ações de melhoria contínua.

No nível tático do PCP, os percentis orientam a formulação de prazos e a escolha de margens. Ordens cujo perfil cai em combinações historicamente dispersas (P95 distante) pedem planos com folga explícita e validações prévias de preparação. Ordens em combinações estáveis tendem a cumprir prazos mesmo em janelas mais curtas. O resultado prático aparece na redução de retrabalho de agenda e na estabilização do ritmo diário.

Quanto à produtividade, os tempos-padrão por percentil podem atuar como denominador comum para comparar máquinas sem confundir efeito de mix com eficiência. Ao confrontar a produtividade observada com o percentil de referência da combinação Máquina $\times$ NMedidas, distingue-se o que decorre de características dos calibradores do que resulta de método e execução. Com isso, a análise direciona esforços para frentes em que a variabilidade é estrutural, evidenciada pela distância entre P95 e P50, e evita conclusões incoerentes entre máquinas de naturezas diferentes.

A inferência dos tempos-padrão atende diretamente à pergunta de pesquisa ao estimar, com precisão operacionalmente relevante, não só quanto tempo cada máquina tende a demandar no cenário de referência, mas também como esse tempo se desloca conforme o perfil das peças (faixas de características medidas). Ao produzir tempos-padrão publicáveis por máquina e estratificados por complexidade em percentis (P50, P75, P95), as saídas do modelo deixam de ser somente um inventário descritivo e podem passar a embasar decisões, como dimensionamento de capacidade, sequenciamento, preparação e gestão de riscos. Lidas no contexto da Empresa Z, elas mostram onde a operação é previsível ($P50 \approx P75 \approx P95$), onde há cauda longa que requer preparação e agrupamento de lotes e onde percentis mais conservadores (P75/P95) reduzem reprogramações e perdas de produtividade.

4.3 FATORES EXPLICATIVOS DO TEMPO DE INSPEÇÃO

Nesta etapa, busca-se identificar as condições operacionais que, quando presentes, tornam mais provável a ocorrência de inspeções demoradas. Com os tempos previstos e consolidados, aplicou-se regras de associação (*A priori*) para identificar combinações de condições associadas a inspeções mais longas. Utilizou-se os limiares definidos na seção 3.4, suporte mínimo 20%, confiança $\geq 65\%$ para o percentil 50, suporte mínimo 4%, confiança $\geq 65\%$ para o percentil 75, e suporte mínimo 1%, confiança $\geq 50\%$ para o percentil 95. As regras foram ordenadas por *lift* para destacar padrões com ganho real de probabilidade em relação ao acaso.

Tabela 10 – Fatores que elevam o tempo para o percentil 50

Regras	Suporte	Confiança	<i>Lift</i>
Família = 142, Máquina = Máquina 7	23,17%	100%	1,49
Família = 142, PP = 1	25,61%	95,45%	1,42
Família = 142	26,42%	91,55%	1,36
Máquina = Máquina 7, <i>Setup</i> Extra = 1	22,76%	90,32%	1,35
Máquina = Máquina 7, NConformidade = 0	22,36%	90,16%	1,34
Máquina = Máquina 7	25,20%	89,86%	1,34

Fonte: Elaborado pelo autor

Para o alvo $\geq P50$, os padrões apontam um “perfil típico” mais do que gargalos: os *lifts* são moderados ($\approx 1,3$ – $1,5$), o que indica aumento modesto da chance de se posicionar acima da mediana. O principal vetor é a Família 142 na Máquina 7, a regra Família 142 + Máquina 7 tem 100% de confiança, suporte 23,17% e *lift* 1,49, sinalizando que, quando essa combinação ocorre, o tempo tende a ficar ao menos na faixa mediana — não por anomalia, mas por característica do processo manual dessa máquina. Ainda, na Família 142, PP = 1 (programa pronto) aparece com *lift* 1,42 e confiança 95,45%, sugerindo que, mesmo com roteiro preparado, a execução mantém tempos em torno ou acima do P50. A Máquina 7 isolada eleva a probabilidade de $\geq P50$ (confiança 89,66%, *lift* 1,34), e o acréscimo de *Setup* Extra = 1 intensifica levemente o efeito (*lift* 1,35), o que é coerente com tarefas manuais que exigem manipulações adicionais.

Tabela 11 – Fatores que elevam o tempo para o percentil 75

Regras	Suporte	Confiança	<i>Lift</i>
Família = 165, Máquina = Máquina 3	7,72%	90,48%	3,23
Família = 165, NMedidas = 4-6	5,69%	82,35%	2,94
Máquina = Máquina 3	8,13%	74,07%	2,64
NMedidas = 4-6	7,32%	69,23%	2,47
Máquina = Máquina 2, PP = 0	5,69%	66,67%	2,38

Fonte: Elaborado pelo autor

Os resultados convergem para dois vetores principais: Máquina 3 e a faixa de 4 a 6 características medidas. A presença isolada de Máquina 3 eleva a chance de tempo $\geq P75$ (confiança 74,07%, *lift* 2,64), e as combinações com a Família 165 intensificam o efeito: Família = 165 + Máquina 3 (confiança 90,48%, *lift* 3,23) e Família = 165 + NMedidas = 4–6 (confiança 82,35%, *lift* 2,94), sugerindo que, nessa família, peças com carga “média” tendem a alongar a inspeção. Do lado de condições de processo, uma regra recebe destaque: Máquina 2 sem programa pronto (PP = 0) (confiança 66,67%, *lift* 2,38). Reforçando que ausência de programa pronto é gatilho de alongamento, em linha com a interpretação de importância de variáveis da seção anterior. Em termos de cobertura, os suportes entre ~5% e 9% indicam padrões não raros, logo úteis para ação gerencial.

Tabela 12 – Fatores que elevam o tempo para o percentil 95

Regras	Suporte	Confiança	<i>Lift</i>
Máquina = Máquina 3, <i>Setup</i> Extra = 1	1,22%	75,00%	14,19
Família = 165, MP = 1	1,22%	60,00%	11,35

Fonte: Elaborado pelo autor

A Tabela 12 mostra dois “gatilhos de cauda” que explicam os eventos mais extremos, ambos com suporte baixo (~1,22%) porém *lifts* muito altos: Máquina 3 + *Setup* Extra = 1 (confiança 75%, *lift* 14,19) e Família 165 + MP = 1 (confiança 60%, *lift* 11,35). Em termos práticos, são condições raras, mas de alto impacto, que multiplicam em ordem de grandeza a chance de cair no topo da distribuição.

Para interpretar as regras, é preciso cautela: associação não implica causalidade — suporte, confiança e *lift* expressam coocorrências úteis para triagem, mas não estabelecem causa; portanto, as hipóteses extraídas devem ser validadas no processo real (Agrawal; Srikant, 1994). Além disso, redundância e colinearidade são frequentes: um mesmo fenômeno pode surgir em várias regras parcialmente sobrepostas, razão pela qual se recomenda filtragem por

medidas de interesse antes da tomada de decisão (Han; Kamber; Pei, 2011). Outro ponto é que a discretização de variáveis contínuas (como NMedidas em faixas) influencia a descoberta de padrões; testar limiares alternativos é boa prática para checar a robustez das associações (Dougherty; Kohavi; Sahami, 1995).

A leitura conjunta das regras em P50, P75 e P95 permite traduzir os resultados em decisões operacionais concretas. Para o percentil 50, o padrão recorrente da Família 142 na Máquina 7 mostra que a execução tende a posicionar os tempos ao redor ou ligeiramente acima da mediana. Para gestão, isso significa planejar capacidade e mix levando em conta que a carga dessa família desloca a mediana da Máquina 7. Em termos práticos, vale reservar janelas específicas para essa família, e há potenciais benefícios em melhorias do procedimento padrão para essa família nessa máquina, de modo a reduzir pequenas perdas que se acumulam ao longo do turno. Também é recomendável que a avaliação de produtividade dessa máquina seja acompanhada por família, pois a composição do mix explica parte relevante da variabilidade observada.

No percentil 75, os sinais são mais prescritivos. A Máquina 3 figura como eixo central e o risco de alongamento aumenta quando combinada à Família 165 e à faixa de 4–6 características; adicionalmente, a ausência de PP na Máquina 2 eleva a probabilidade de tempos longos. Em termos de gestão, recomenda-se tratar esses cenários como fluxos sensíveis: i) adotar janelas de processamento dedicadas para ordens da Família 165 com 4–6 características; ii) usar as regras de P75 como critérios de liberação no agendamento, priorizando sequência que evite encadeamento de setups e preserve capacidade para lotes mais previsíveis; e iii) Na política de programação, considerar a necessidade de criar programas, mantendo ordens sem programa pronto fora de janelas críticas da Máquina 2. Essas ordens podem entrar numa fila de preparação, com metas explícitas de cobertura de PP para as famílias mais frequentes.

No percentil 95, a utilidade para a decisão é de gestão de risco e *Service Level Agreement* (SLA). No caso Máquina 3 + *Setup* Extra, recomenda-se utilizar um critério de liberação antes do agendamento (pré-verificação de roteiro, fixações e recursos), preferindo janelas dedicadas e preparação antecipada quando o setup adicional for inevitável. MP = 1 é um marcador de incidente identificado durante a execução; por isso, ele orienta pós-ação e aprendizado (registro de caso, 5 Porquês, Pareto de falhas) e deve acionar, antecipadamente, marcadores de risco para as próximas ordens da Família 165, estado das aplicações de *Total Productive Maintenance* (TPM), pendências de manutenção, histórico de falhas recentes, alarmes/sintomas de condição, *backlog* de setups atípicos ou desvios de tempo de preparação. Se qualquer marcador de risco estiver ativo, a ordem deve ser tratada como fluxo sensível; se não, segue o fluxo regular.

Essas recomendações devem ser priorizadas com critérios objetivos. Para decisões de curto prazo, o que combina *lift* elevado com suporte razoável deve ser prioridade de melhoria; o que tem *lift* altíssimo e suporte reduzido entra no catálogo de contenções e alertas (como as regras de P95); o que tem suporte elevado em P50 serve para balancear capacidade e promessas de prazo, e pode ser alvo de melhoria para ganhos menores. Em médio prazo, os próprios achados podem alimentar a rotina gerencial: um painel de risco por ordem que sinaliza padrões de P75/P95 no momento do agendamento, e um quadro de status de PP por família e máquina, que impede a liberação para a Máquina 2 sem roteiro pronto. Por fim, o acompanhamento de indicadores encerra o ciclo de decisão. A proporção de ordens com risco $\geq$ P75 processadas com PP = 1, o tempo médio de preparação e sua queda após padronizações, a incidência mensal de eventos $\geq$ P95 e a produtividade da Máquina 7 estratificada por família funcionam como medidas de monitoramento da execução das ações e de evidência da sustentação das melhorias ao longo do tempo.

Apresentado o desempenho dos modelos, os tempos-padrão obtidos e as evidências sobre os fatores que mais elevam o tempo de inspeção, a Seção 5 avança para a interpretação desses achados. Nela, os resultados são confrontados com a literatura e com a realidade operacional da Empresa Z, destacando implicações práticas, benefícios e limites do estudo, além de apontar recomendações para aplicação e melhoria contínua do processo.

5 DISCUSSÃO DOS RESULTADOS

Comparando diretamente com a literatura, este estudo confirma a vantagem de *ensembles* baseados em árvores quando há variáveis mistas (numéricas e categóricas) e interações — razão pela qual o CatBoost, com seu esquema de “*ordered boosting*” que reduz *prediction shift* em categorias, prevaleceu nesse conjunto. Ainda assim, o R^2 obtido ($\approx 0,59$) fica abaixo de alguns relatos com objetivos próximos, como Çaklıt e Dağdeviren (2023), que reportam desempenhos mais elevados em cenários específicos. Essa diferença é compatível com três fatores estruturais: i) escopo e ruído do domínio — tempos de inspeção em metrologia tendem a incorporar variância operacional difícil de observar (ex.: qualidade/estado da peça, microparadas, interpretação dos resultados pelo inspetor), elevando o erro irreduzível frente a contextos de produção mais estáveis; ii) estratégia de validação e controle de vazamento — adotou-se particionamento e engenharia de atributos conservadores (e.g., segregação por máquina/família, uso de percentis e variáveis “limpas” como NMedidas, PP, Setup Extra), o que tende a reduzir *overfitting* e “ R^2 inflado”, mas aumenta a validade externa (Varma; Simon, 2006); e iii) riqueza e granularidade dos preditores — estudos com R^2 superiores frequentemente dispõem de telemetria/*process mining* e atributos de processo mais densos (sequência de operações, trajetórias de ferramenta, metadados de programa), enquanto o *dataset* deste estudo é limitado a variáveis operacionais essenciais e generalizáveis em quantidades pequenas (Schmitt *et al.*, 2024).

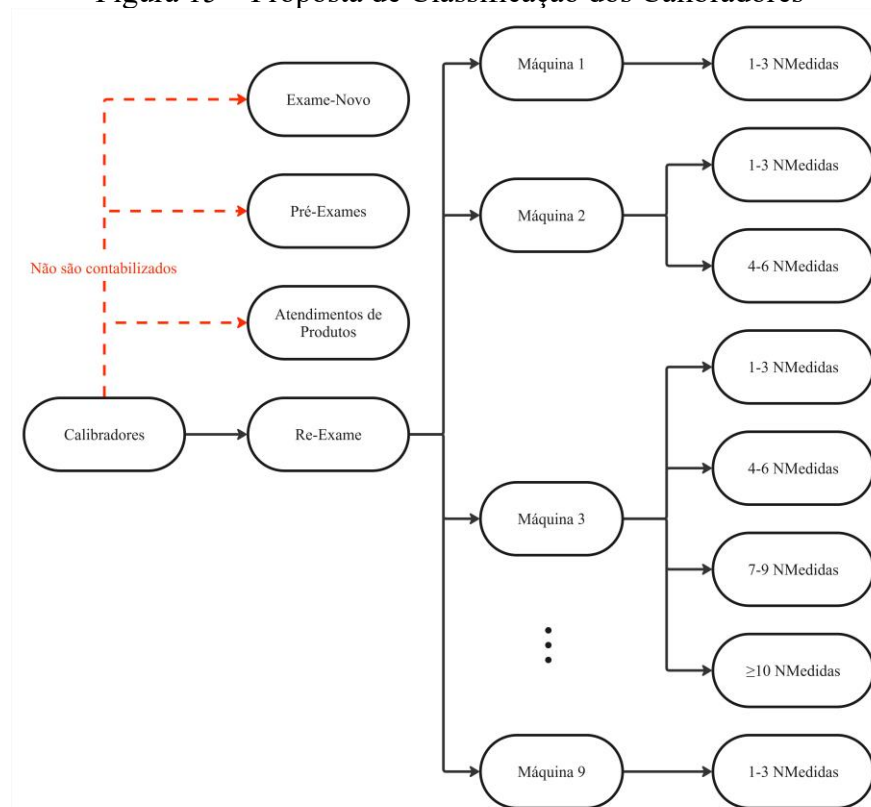
Em contraste, resultados como os de Weller *et al.* (2023) e Urban, Koblasa e Mendřický (2024) reforçam que, mantidas as boas práticas de segmentação e de seleção de atributos, *ensembles* entregam o melhor compromisso entre acurácia e robustez; métodos como k-NN podem superar quando há forte similaridade local e base densa por vizinhança — condição que não caracteriza integralmente o cenário deste estudo. Em síntese, o desempenho obtido é consistente com a literatura quando se pondera domínio, desenho de validação e informatividade dos preditores, mais do que “perseguir R^2 ”, a contribuição está em reduzir o erro confiavelmente (RMSE na ordem de minutos úteis) e em traduzir as previsões em tempos-padrão publicáveis por percentis, o que fornece a base operacional para a gestão diária.

Com essa base, do lado gerencial, adota-se a métrica peça-homem-hora (PHH) como principal indicador de produtividade, calculada a partir dos tempos-padrão publicados. Essa métrica foi escolhida por ser alinhada ao padrão gerencial da Empresa Z. A publicação de valores por percentis (P50/P75/P95) por máquina $\times$ faixa de NMedidas permite derivar três trilhas de gestão: (i) planejamento de prazo ancorado no P75 (cobertura típica), ou no P95 para

itens críticos; (ii) meta de produtividade por turno/célula baseada em PHH calculado com o percentil de referência; e (iii) proposição e avaliação de melhorias via monitoramento da diferença entre os percentis P50–P95, que sinaliza cauda longa e prioriza ações em *setup*, programa pronto (PP) e padronização.

Por fim, a opção por agrupar, em vez de publicar tempos-padrão para cada calibrador, é coerente com os erros observados e com a necessidade de uma classificação operacional: reduz o risco de subestimar itens complexos e superestimar itens simples, alinha metas de PHH ao esforço real e preserva generalização — isto é, a métrica continua estável mesmo quando o mix muda. Como salvaguardas, devem ser definidos: i) um canal de exceções para peças atípicas; ii) um gatilho de revisão quando o erro absoluto médio de uma classe excede um limiar, bem como revisões periódicas dos valores preditos; iii) auditorias amostrais de PP, MP, NC e *setup* extra para evitar que condições eventuais virem “novo normal” na base de cálculo. Isso mantém a métrica útil para decisão diária, sem perder rastreabilidade metodológica. Como o fluxo de processamento é agrupado em um *pipeline*, o método é replicável, e à medida que novos dados são incorporados e o modelo é reestimado periodicamente, espera-se redução do erro (RMSE) e maior estabilidade das previsões de tempo de inspeção.

Figura 15 – Proposta de Classificação dos Calibradores



Fonte: Elaborado pelo autor

A classificação nova responde a dois fatos empíricos dos resultados: o erro (e a dispersão) cresce com a complexidade da peça e varia entre máquinas; logo, agrupar por esses eixos reduz heterogeneidade residual, melhora a previsibilidade do PHH e simplifica a comunicação entre PCP e o setor. Na prática, como a demanda das tarefas de Re-Exame é regida pelo tempo de validade dos calibradores, trata-se de um componente de carga determinístico (previsível com alta precisão), o que permite planejar essa parcela da produção com antecedência. A partir disso, recomenda-se estimar um *buffer* diário para as demais atividades com base na sua frequência observada.

A avaliação qualitativa permitiu avaliar o estado de implementação do estudo na Empresa Z e os entrevistados apontaram quatro efeitos centrais. Primeiro a quantificação dos problemas percebidos no setor, convertendo percepções dispersas em medidas para decisão. Esse é o papel de um sistema de medição, ligar diagnóstico à ação. No *Balanced Scorecard* (BSC), medidas financeiras são complementadas por indicadores operacionais que impulsionam o desempenho futuro, orientando decisões e evitando vieses de controle estritamente contábil (Kaplan; Norton, 1992). Revisões da literatura reforçam que sistemas de medição devem derivar da estratégia e que o ato de medir molda comportamentos e o próprio ciclo de planejamento e controle (Bourne *et al.*, 2003).

Segundo, *insights* de programação e preparação sobre como criar rotinas de programação para a Máquina 2 e reduzir setups adicionais, em linha com o *Single Minute Exchange of Die* (SMED), que propõe separar tarefas de *setup* internas e externas e converter o máximo possível do interno em externo para reduzir paradas, combinando padronização de fixações, pré-ajustes fora da máquina e checagens rápidas antes da liberação (Shingo, 1985). Terceiro, o entendimento do caminho de melhoria, evidências sobre recursos críticos ajudam a priorizar onde atuar primeiro, coerente com a lógica da Teoria das Restrições, na qual a restrição comanda o fluxo e orienta esforço de melhoria (Şimşit; Günay; Vayvay, 2014). Por fim, a incorporação de variáveis ignoradas nos cálculos de produtividade e a agenda de qualificação da base de dados, reforçando a maturidade analítica e aumentando a utilidade operacional dos modelos, consistente com a literatura de gestão orientada por dados (Davenport; Harris, 2007; Szukits; Moricz, 2023).

O estudo também contribuiu para priorizar frentes de melhoria com maior efeito sobre tempo e *output*, sobretudo ao promover uma revisão crítica do número e do escopo das características medidas. A verificação sistemática do que é realmente necessário para atender ao objetivo de controle e aos requisitos do cliente leva à eliminação de medições redundantes e de checagens em nível de detalhe além do necessário — casos típicos de superprocessamento,

classificado no Lean como desperdício por não agregar valor do ponto de vista do cliente (Ohno, 1988; Shingo, 1985; Womack; Jones, 2003).

Quanto à classificação, o setor por ora permanece com a classificação A–B–C descrita anteriormente (Figura 5), embora os tempos usados nessa classe tenham sido recalculados com base no modelo revisado. A manutenção desse arranjo tem duas razões operacionais: i) ausência do campo “quantidade de características medidas” no banco de dados, o que impede estimar previamente o perfil de cada calibrador; e ii) conveniência de somar tempos pela rota completa do calibrador (muitos passam por mais de uma máquina) na falta de um PCP/sequenciamento formal. Como trajetória de evolução, a recomendação é incluir NMedidas na base e adotar um PCP leve com heurísticas simples de sequenciamento; com isso, torna-se viável gradualmente migrar do A–B–C para a classificação proposta pelo estudo.

A principal dificuldade para implementar as melhorias tem sido capacidade de execução, sobretudo recursos humanos. Iniciativas como desenvolver ou revisar programas de medição exigem dedicar inspetores exclusivamente a essa tarefa por um período, o que, dado o volume atual de ordens, não é factível sem comprometer o atendimento. Adicionalmente, há uma limitação de dados e telemetria, como as máquinas não registram automaticamente o tempo total de inspeção, não é possível coletar continuamente e a manter uma base atualizada para alimentar o modelo ao longo do tempo. Os entrevistados destacaram como ponto positivo a transparência do trabalho com dados, os testes foram compreensíveis e o alinhamento “um-a-um” entre as previsões do modelo e o comportamento observado em operação foi percebido como claro, o que reforça a confiabilidade prática das estimativas.

Como passo futuro, propõe-se elevar eventos operacionais a parâmetros de planejamento, em especial *setup* extra, programa pronto (PP) e não conformidade (NC). Para cada classe (Máquina $\times$ faixa de NMedidas), estima-se a taxa de ocorrência e o impacto médio em minutos desses eventos a partir do histórico recente e, com isso, definem-se fatores de segurança coerentes com o nível de serviço. Esses fatores seriam aplicados fora do tempo-padrão — isto é, como camadas de ajuste no plano de capacidade e nas janelas de sequenciamento por classe — permitindo reservar carga para preparação, retrabalhos pontuais e variações previsíveis sem “contaminar” o padrão. O resultado esperado é um plano mais aderente à realidade, mantendo o tempo-padrão limpo para comparação de desempenho e para comunicação com as áreas clientes.

Em termos de generalização, o núcleo do estudo é uma abordagem que transforma dados operacionais em tempos-padrão por recortes de complexidade (P50, P75, P95). A Figura 12 sintetiza essa lógica — EDA por correlação, seleção de preditores por *wrapper*, comparação e

escolha do modelo para inferir tempos, e conversão do tempo normal no tempo-padrão sob um cenário de referência. A transposição prática consiste em trocar “calibradores” pelo objeto de trabalho local (ex.: ordens de serviço de manutenção, lote de inspeção de qualidade de recebimento, *ticket* de engenharia de processos, *milk-run* de logística interna), manter a estratificação por famílias/complexidade e publicar os tempos por percentis para gestão de promessa e capacidade. Em cada contexto, uma análise prévia por correlação (*heatmap*) e a seleção de variáveis do tipo *wrapper* podem revelar direcionadores específicos, sugerir ajustes de faixas de complexidade e eliminar colinearidades redundantes. Ainda assim, a ideia central do estudo permanece: preparar e curar dados, treinar e validar o modelo, projetar o cenário de referência e publicar tempos por percentis. A etapa de EDA + *wrapper* apenas customiza os preditores ao domínio, enquanto o formato de entrega (P50/P75/P95) segue estável e comparável entre empresas e setores.

Além dos tempos-padrão, o módulo de regras de associação se transfere como um “radar de risco” de alongamento. Ao minerar coocorrências entre condições operacionais e tempos elevados, gera-se listas de fluxos sensíveis para preparar, sequenciar, ou submeter a verificação prévia. A replicação é mais simples quando a organização dispõe de dados retirados diretamente dos processos (*logs* de ERP/MES, registros de máquina, *timestamps* de início/fim, status de programa, apontamentos de *setups* e não conformidades), pois a coleta é contínua e padronizada. Ainda assim, os pré-requisitos são modestos, registrar variáveis de processo e distinguir o cenário de referência de eventos extraordinários. Onde não há telemetria, é viável começar com dados administrativos/operacionais curados, evoluindo a granularidade ao longo do tempo sem mudar o método, apenas enriquecendo o que está sendo medido.

No âmbito acadêmico, o trabalho contribui em duas frentes complementares. Primeiro, formaliza um problema pouco tratado — a carência de tempos-padrão confiáveis em rotinas de inspeção/metrologia — e o operacionaliza com a aplicação de ML orientado a uso. Isso desloca o foco típico da literatura de manufatura direta para um setor de apoio crítico, ampliando o campo empírico de estudos de tempos e de aplicações de ML.

Segundo, descreve uma abordagem reprodutível que documenta as decisões de pesquisa que tornam o método auditável e transferível: i) formulação do problema e objetivos; ii) extração/curadoria dos dados; iii) EDA (correlação/*heatmap*) e seleção de preditores por *wrapper*; iv) comparação sistemática de algoritmos com validação e ajuste de hiperparâmetros, seleção do modelo adequado, e avaliação por R^2 /RMSE com leitura de importância; v) mineração de regras para identificação dos fatores que alongam o tempo; e vi) tradução do tempo normal em tempos-padrão por percentis sob cenário de referência e aplicação de

tolerâncias por máquina. O resultado é uma abordagem replicável que integra modelagem, validação e publicação operacional de tempos, útil para futuras comparações entre domínios e para generalização analítica em outras operações.

6 CONCLUSÕES

O presente estudo teve por objetivo estimar tempos-padrão de inspeção em metrologia e identificar os fatores que mais influenciam sua variação. Para isso, estruturou-se um *pipeline* reprodutível de aprendizado de máquina aplicado a dados coletados do setor, com preparação e curadoria das variáveis, comparação sistemática de algoritmos e seleção do modelo vencedor. Entre as alternativas testadas, os *ensembles* baseados em árvores — em especial o CatBoost — apresentaram o melhor compromisso entre acurácia e robustez para dados mistos (numéricos e categóricos), atingindo desempenho consistente ($R^2 \approx 0,59$) e confirmando a viabilidade de produzir previsões confiáveis de tempo normal e, a partir delas, tempos-padrão publicáveis por percentis.

Os resultados empíricos apontaram como direcionadores centrais o número de características medidas (NMedidas) e diferenças estruturais entre máquinas, além de efeitos operacionais como programa pronto e *setup* extra. Com base nisso, foi proposta e implementada uma classificação operacional por Máquina $\times$ faixa de NMedidas, que reduz heterogeneidade residual e facilita a publicação de tempos-padrão por P50/P75/P95. Essa base permitiu adotar a métrica peça-homem-hora como indicador gerencial, desdobrando três usos práticos: planejamento de prazos ancorado em percentis, metas de produtividade por classe e priorização de melhorias a partir da diferença entre percentis.

Os resultados apresentados se alinham diretamente aos objetivos propostos. O objetivo geral — estimar tempos-padrão de inspeção para avaliar a produtividade a partir de machine learning — foi atendido ao se treinar e selecionar um modelo vencedor, converter as previsões de tempo normal em tempos-padrão por percentis sob cenário de referência e desdobrá-los em PHH para uso gerencial. Quanto às contribuições colaterais: a) a análise comparativa de algoritmos foi conduzida com validação e ajuste de hiperparâmetros, selecionando CatBoost como o melhor modelo, evidenciando a superioridade relativa de *ensembles* de árvores nesta base; b) a classificação dos calibradores em grupos por semelhança materializou-se na clusterização Máquina $\times$ faixa de NMedidas, que agrupa itens por esforço de medição comparável e reduz a heterogeneidade residual; c) a averiguação dos fatores que elevam o tempo foi realizada combinando importância de variáveis (destaque para NMedidas e diferenças entre máquinas) e regras de associação que apontam gatilhos operacionais (PP=0, *setup* extra); e d) o modelo para averiguação de tempos-padrão em setores de apoio foi proposto e demonstrado como um fluxo metodológico reprodutível, deixando diretrizes para monitoramento, reestimativa e expansão de atributos em aplicações futuras. Em conjunto, esses

entregáveis fecham o ciclo entre problema, método e uso prático, atingindo o objetivo central e cada uma das metas específicas definidas no escopo do estudo.

Em termos acadêmicos, este trabalho formaliza e operacionaliza um problema pouco tratado — a falta de tempos-padrão confiáveis em inspeção/metrologia — ao traduzi-lo em alvos mensuráveis por ML e deslocar o foco para um setor de apoio de alta variabilidade e mix elevado; e apresenta um fluxo metodológico reproduzível de ponta a ponta, que integra curadoria de dados, EDA com seleção por *wrapper*, comparação e validação de algoritmos, mineração de regras para identificação dos fatores de alongamento e publicação operacional por percentis sob cenário de referência, com tolerâncias por máquina. Assim, oferece um *framework* auditável e transferível que conecta modelagem, validação e uso gerencial direto, ampliando o escopo dos estudos de tempos e facilitando sua replicação em outros contextos.

Entretanto, os métodos apresentados nesse estudo possuem limitações que exigem cuidados. A variabilidade individual dos operadores pode gerar desvios mesmo após o ajuste de desempenho; tarefas não repetitivas ou de elevado conteúdo cognitivo podem reduzir a precisão dos estudos; o estudo de tempos está sujeito a vieses de cronometragem e de avaliação de ritmo pelo observador; a implementação de sistemas de dados padronizados envolve custos de treinamento, aquisição de bases e adaptação de processos; além disso, o acompanhamento rigoroso de tempos pré-estabelecidos pode induzir respostas comportamentais ao tempo/pressão, fazendo com que trabalhadores alterem o ritmo (privilegiem velocidade em detrimento de qualidade) (Heitz, 2014). Em contextos de observação direta, mudanças de comportamento também podem ocorrer pelo chamado “efeito Hawthorne”, isto é, a reatividade por saber que está sendo observado, o que pode enviesar as medições do estudo (Goodwin *et al.*, 2017). Outras limitações destacadas são a granularidade dos dados, e o tamanho da base em algumas classes, fatores que restringem ganhos adicionais de acurácia.

Como passos futuros, sugere-se i) ampliar o conjunto de atributos com metadados de processo e logs; ii) monitorar *drift* e reestimar periodicamente os percentis por classe; e iii) explorar, em caráter piloto, fatores de segurança derivados das taxas de ocorrência e impactos médios de eventos como PP = 0, *setup* extra e NC para construção de cenários de capacidade. Em síntese, o estudo entrega um método replicável, evidência empírica sobre os principais direcionadores de tempo de inspeção e um arranjo gerencial que conecta tempos por percentis à gestão diária por PHH, oferecendo ganhos imediatos de previsibilidade e alinhamento entre engenharia de tempos e operação.

No âmbito acadêmico, abrem-se frentes para novos estudos. Validar externamente o método em outros contextos de metrologia ou em outros setores de apoio; explorar modelos

hierárquicos/multinível que decomponham variabilidade entre máquinas, famílias e operadores, permitindo inferência mais estável com classes pequenas; quantificar incerteza preditiva e comparar políticas de decisão para diferentes cenários; incorporar atributos temporais/sequenciais (ordem de inspeções, reprogramações, *toolpaths*/estratégias de medição) e avaliar ganhos de acurácia; e, por fim, avaliar protocolos de confiabilidade da medição de tempo (consistência interobservador e precisão de *timestamp*), reduzindo ruído e melhorando a reprodutibilidade científica. Essas linhas complementam a contribuição aplicada do trabalho e podem ampliar a generalização, a interpretabilidade e a robustez do método em diferentes domínios industriais.

REFERÊNCIAS

AGGARWAL, C. C.; HINNEBURG, A.; KEIM, D. A. On the surprising behavior of distance metrics in high dimensional space. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, [s. l.], vol. 1973, p. 420–434, 2001. https://doi.org/10.1007/3-540-44503-x_27.

AGRAWAL, R.; SRIKANT, R. Fast Algorithms for Mining Association Rules in Large Databases. 1994. **Proceedings of the 20th International Conference on Very Large Data Bases [...]**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1994. p. 487–499.

ALTMEIER, M.; KUNKEL, K.; MOYA-QUIROGA, R.; WIESNER, F.; MÜLLER, N. Indirect manufacturing costs: An overlooked source for clear savings. **McKinsey & Company**, [s. l.], no. January, 2023. .

ARORA, I.; KHANDUJA, N.; BANSAL, M. Effect of Distance Metric and Feature Scaling on KNN Algorithm while Classifying X-rays. **CEUR Workshop Proceedings**, [s. l.], vol. 3176, p. 61–75, 2022. .

BARNES, R. M. **Estudo de Movimentos e de Tempos: Projeto e Medida do Trabalho**. São Paulo: Edgard Blücher, 1977.

BENTÉJAC, C.; CSÖRGÖ, A.; MARTÍNEZ-MUÑOZ, G. A comparative analysis of gradient boosting algorithms. **Artificial Intelligence Review**, [s. l.], vol. 54, no. 3, p. 1937–1967, 2021. DOI 10.1007/s10462-020-09896-5. Available at: <https://doi.org/10.1007/s10462-020-09896-5>.

BERTRAND, J. W. M.; FRANSOO, J. C.; UDENIO, M. Model-based research. In: KARLSSON, C. (ed.). **Research Methods for Operations and Supply Chain Management**. 3rd ed. Abingdon; New York: Routledge, 2023. p. 274–311. DOI 10.4324/9781003315001. Available at: <https://doi.org/10.4324/9781003315001>.

BOURNE, M.; NEELY, A.; MILLS, J.; PLATTS, K. Implementing Performance Measurement Systems: A Literature Review. **Int. J. Business Performance Management Int. J. Business Performance Management**, [s. l.], vol. 5, p. 1–24, 2003. <https://doi.org/10.1504/IJBPM.2003.002097>.

BREIMAN, L. Random Forests. **Machine Learning**, [s. l.], vol. 45, no. 1, p. 5–32, 2001. DOI 10.1023/A:1010933404324. Available at: <https://doi.org/10.1023/A:1010933404324>.

BREZNIK, M.; BUCHMEISTER, B.; VUJICA HERZOG, N. Assembly Line Optimization Using MTM Time Standard and Simulation Modeling—A Case Study. **Applied Sciences (Switzerland)**, [s. l.], vol. 13, no. 10, 2023. <https://doi.org/10.3390/app13106265>.

BUCHMEISTER, B.; HERZOG, N. V. Advancements in Data Analysis for the Work-Sampling Method. **Algorithms**, [s. l.], vol. 17, no. 5, 2024. <https://doi.org/10.3390/a17050183>.

BUITINCK, L.; LOUPPE, G.; BLONDEL, M.; PEDREGOSA, F.; MUELLER, A.; GRISEL, O.; NICULAE, V.; PRETTENHOFER, P.; GRAMFORT, A.; GROBLER, J.; LAYTON, R.; VANDERPLAS, J.; JOLY, A.; HOLT, B.; VAROQUAUX, G. {API} design for machine learning software: experiences from the scikit-learn project. 2013. **ECML PKDD Workshop: Languages for Data Mining and Machine Learning [...]**. [S. l.: s. n.], 2013. p. 108–122.

ÇAKLT, E.; DAĞDEVİREN, M. Comparative analysis of machine learning algorithms for predicting standard time in a manufacturing environment. **Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM**, [s. l.], vol. 37, 2023. <https://doi.org/10.1017/S0890060422000245>.

CHEHADE, M.; SYLLA, A.; DIALLO, ABDOUL RAHIME DOREMUS, Y. Manufacturing time estimation for offer pricing: A machine learning application in a French metallurgy industry. **Engineering Applications of Artificial Intelligence**, [s. l.], vol. 137, 2024. .

CHEN, T.; GUESTRIN, C. XGBoost: A scalable tree boosting system. **Association for Computing Machinery**, [s. l.], vol. 13-17-Aug, p. 785–794, 2016. <https://doi.org/10.1145/2939672.2939785>.

CHENG, C. F.; LIN, C. J.; HU, Q. X. Machine learning models for sequential motion recognition in human activity for logistics operations efficiency improvement. **Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM**, [s. l.], vol. 39, 2025. <https://doi.org/10.1017/S0890060424000313>.

CHOUDHARY, A.; HARDING, J.; TIWARI, M. Data mining in manufacturing: A review based on the kind of knowledge. **Journal of Intelligent Manufacturing**, [s. l.], vol. 20, p. 501–521, 2009. <https://doi.org/10.1007/s10845-008-0145-x>.

CRESWELL, J. W. Research design: Qualitative, quantitative, and mixed methods approaches, 3rd ed. **Research design: Qualitative, quantitative, and mixed methods approaches, 3rd ed.**, Thousand Oaks, CA, US, , p. xxix, 260–xxix, 260, 2009. .

CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of Control, Signals, and Systems**, [s. l.], vol. 2, no. 4, p. 303–314, 1989.

<https://doi.org/10.1007/BF02551274>.

DA SILVA, A. G.; GOMES, M. V. M.; WINKLER, I. Virtual Reality and Digital Human Modeling for Ergonomic Assessment in Industrial Product Development: A Patent and Literature Review. **Applied Sciences (Switzerland)**, [s. l.], vol. 12, no. 3, 2022. <https://doi.org/10.3390/app12031084>.

DAMSCHRODER, L. J.; REARDON, C. M.; WIDERQUIST, M. A. O.; LOWERY, J. The updated Consolidated Framework for Implementation Research based on user feedback. **Implementation Science**, [s. l.], vol. 17, no. 1, p. 75, 2022. DOI 10.1186/s13012-022-01245-0. Available at: <https://doi.org/10.1186/s13012-022-01245-0>.

DAS, B. Personal, Unavoidable Delay, and Fatigue Allowances for Setting Performance Standards. **International Journal of Operations & Production Management**, [s. l.], vol. 10, no. 5, p. 62–70, 1990. DOI 10.1108/01443579010143477. Available at: <https://doi.org/10.1108/01443579010143477>.

DAVENPORT, T. H.; HARRIS, J. G. **Competing on Analytics: The New Science of Winning**. 1st ed. USA: Harvard Business School Press, 2007.

DORMANN, C. F.; ELITH, J.; BACHER, S.; BUCHMANN, C.; CARL, G.; CARRÉ, G.; MARQUÉZ, J. R. G.; GRUBER, B.; LAFOURCADE, B.; LEITÃO, P. J.; MÜNKEMÜLLER, T.; MCCLEAN, C.; OSBORNE, P. E.; REINEKING, B.; SCHRÖDER, B.; SKIDMORE, A. K.; ZURELL, D.; LAUTENBACH, S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. **Ecography**, [s. l.], vol. 36, no. 1, p. 27–46, 2013. DOI <https://doi.org/10.1111/j.1600-0587.2012.07348.x>. Available at: <https://nsojournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0587.2012.07348.x>.

DOUGHERTY, J.; KOHAVI, R.; SAHAMI, M. Supervised and Unsupervised Discretization of Continuous Features. In: PRIEDITIS, A.; RUSSELL, S. (eds.). **Machine Learning Proceedings 1995**. San Francisco (CA): Morgan Kaufmann, 1995. p. 194–202. DOI <https://doi.org/10.1016/B978-1-55860-377-6.50032-3>. Available at: <https://www.sciencedirect.com/science/article/pii/B9781558603776500323>.

DRESCH, A.; LACERDA, D.; ANTUNES JÚNIOR, J. A. V. Desing Science Research: A Method for Science and Technology Advancement. [s. l.], 2014. https://doi.org/10.1007/978-3-319-07374-3_4.

DREYFUS, P. A.; PSAROMMATIS, F.; MAY, G.; KIRITSIS, D. Virtual metrology as an approach for product quality estimation in Industry 4.0: a systematic review and integrative conceptual framework. **International Journal of Production Research**, [s. l.], vol. 60, no. 2, p. 742–765, 2022. DOI 10.1080/00207543.2021.1976433. Available at:

<https://doi.org/10.1080/00207543.2021.1976433>.

EISENHARDT, K.; GRAEBNER, M. Theory Building From Cases: Opportunities And Challenges. **The Academy of Management Journal**, [s. l.], vol. 50, p. 25–32, 2007. <https://doi.org/10.5465/AMJ.2007.24160888>.

EMMERT-STREIB, F.; DEHMER, M. High-Dimensional LASSO-Based Computational Regression Models: Regularization, Shrinkage, and Selection. **Machine Learning and Knowledge Extraction**, [s. l.], vol. 1, no. 1, p. 359–383, 2019. <https://doi.org/10.3390/make1010021>.

ERMEL, A. P. C. **LITERATURE GROUNDED THEORY: MÉTODO DE PESQUISA PARA INVESTIGAÇÃO SOBRE O CONHECIMENTO CIENTÍFICO E TECNOLÓGICO**. 2021. Universidade do Vale do Rio dos Sinos, 2021. Available at: <https://www.repositorio.jesuita.org.br/handle/UNISINOS/9270>.

FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Mag.**, USA, vol. 17, no. 3, p. 37–54, 1996. DOI 10.1609/aimag.v17i3.1230. Available at: <https://doi.org/10.1609/aimag.v17i3.1230>.

FLORES-HUAMÁN, K.-J.; LORENZO-ESPEJO, A.; MUÑOZ-DÍAZ, M.-L.; ESCUDERO-SANTANA, A. Enhancing process lead time forecasting with machine learning and upstream process data: A case study in wind tower manufacturing. **Computers & Industrial Engineering**, [s. l.], vol. 209, no. July, p. 111410, 2025. DOI 10.1016/j.cie.2025.111410. Available at: <https://doi.org/10.1016/j.cie.2025.111410>.

FLYVBJERG, B. Five Misunderstandings About Case-Study Research. **Qualitative Inquiry**, [s. l.], vol. 12, p. 219–245, 2006. <https://doi.org/10.1177/1077800405284363>.

FREJEIRO-GONZÁLEZ, L.; FEBRERO-BANDE, M.; GONZÁLEZ-MANTEIGA, W. A Critical Review of LASSO and Its Derivatives for Variable Selection Under Dependence Among Covariates. **International Statistical Review**, [s. l.], vol. 90, no. 1, p. 118–145, 2022. DOI <https://doi.org/10.1111/insr.12469>. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12469>.

FREVALDS, A.; NIEBEL, B. W. **Niebel's methods, standards, and work design**. [S. l.: s. n.], 2014.

GAWLIK, A.; LÜDEMANN, J.; NEUHAUSEN, A.; ZEPP, C.; VITINIUS, F.; KLEINERT, J. A Systematic Review of Workplace Physical Activity Coaching. **Journal of occupational rehabilitation**, Netherlands, vol. 33, no. 3, p. 550–569, Sep. 2023. <https://doi.org/10.1007/s10926-023-10093-8>.

GÖDRI, I. Improving Delivery Performance in High-Mix Low-Volume

Manufacturing by Model-Based and Data-Driven Methods. **Applied Sciences (Switzerland)**, [s. l.], vol. 12, no. 11, 2022. <https://doi.org/10.3390/app12115618>.

GOMES RIBEIRO, T.; CARLA DE SOUZA GOMES DOS SANTOS, A.; LOPES HERMSDORFF, W.; AUGUSTO FARIA SALLES, S.; DA CUNHA REIS, A. Controle Estatístico Do Processo E Engenharia De Métodos: Uma Aplicação Conjunta Na Produção De Picolés Statistical Process Control and Methods Engineering: a Joint Application in the Production of Popsicles. [s. l.], 2019. .

GOODWIN, M. A.; STANGE, K. C.; ZYZANSKI, S. J.; CRABTREE, B. F.; BORAWSKI, E. A.; FLOCKE, S. A. The Hawthorne Effect in Direct Observation Research with Physicians and Patients. **Journal of Evaluation in Clinical Practice**, [s. l.], 2017. <https://doi.org/10.1111/jep.12781>.The.

GORI, R. S. L. **Modelo de distribuição de recursos orçamentários baseado em indicadores de desempenho para um Instituto Federal de Educação, Ciência e Tecnologia**. 2016. Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, RS, 2016.

GUEST, G.; BUNCE, A.; JOHNSON, L. How Many Interviews Are Enough? **Field Methods - FIELD METHOD**, [s. l.], vol. 18, p. 59–82, 2006. <https://doi.org/10.1177/1525822X05279903>.

GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. **J. Mach. Learn. Res.**, [s. l.], vol. 3, no. null, p. 1157–1182, 2003. .

HAN, J.; KAMBER, M.; PEI, J. **Data Mining: Concepts and Techniques**. 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.

HARDING, J. A.; SHAHBAZ, M.; SRINIVAS; KUSIAK, A. Data Mining in Manufacturing: A Review. **Journal of Manufacturing Science and Engineering**, [s. l.], vol. 128, no. 4, p. 969–976, 2005. DOI 10.1115/1.2194554. Available at: <https://doi.org/10.1115/1.2194554>.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. [S. l.]: Springer, 2009(Springer series in statistics). Available at: <https://books.google.com.br/books?id=eBSgoAEACAAJ>.

HEITZ, R. P. The speed-accuracy tradeoff: history, physiology, methodology, and behavior. **Frontiers in neuroscience**, Switzerland, vol. 8, p. 150, 2014. <https://doi.org/10.3389/fnins.2014.00150>.

HO, F. K.; COLE, T. J. Non-linear predictor outcome associations. **BMJ Medicine**, [s. l.], vol. 2, no. 1, p. e000396, 2023. <https://doi.org/10.1136/bmjmed-2022-000396>.

HORNIK, K. Approximation capabilities of multilayer feedforward networks. **Neural**

Networks, [s. l.], vol. 4, no. 2, p. 251–257, 1991. [https://doi.org/10.1016/0893-6080\(91\)90009-T](https://doi.org/10.1016/0893-6080(91)90009-T).

INAL, T. C.; GORUROGLU OZTURK, O.; KIBAR, F.; CETINER, S.; MATYAR, S.; DAGLIOGLU, G.; YAMAN, A. Lean six sigma methodologies improve clinical laboratory efficiency and reduce turnaround times. **Journal of clinical laboratory analysis**, United States, vol. 32, no. 1, Jan. 2018. <https://doi.org/10.1002/jcla.22180>.

ISO. **ISO/IEC 17025:2017 — General requirements for the competence of testing and calibration laboratories**. [S. l.: s. n.], 2017.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning: with Applications in R**. [S. l.]: Springer Publishing Company, Incorporated, 2014.

JI, J.; PANNAKKONG, W.; BUDDHAKULSOMSIRI, J. A Computer Vision-Based Model for Automatic Motion Time Study. **Computers, Materials and Continua**, [s. l.], vol. 73, no. 2, p. 3557–3574, 2022. <https://doi.org/10.32604/cmc.2022.030418>.

KALNE, P. S.; MEHENDALE, A. M. The Purpose of Time-Motion Studies (TMSs) in Healthcare: A Literature Review. **Cureus**, [s. l.], vol. 14, no. 10, 2022. <https://doi.org/10.7759/cureus.29869>.

KANAWATY, G. **Introduction to Work Study**. 4th ed. Geneva: International Labour Office, 1992. Available at: <https://hcmindonesia.files.wordpress.com/2012/12/introduction-to-work-study.pdf>.

KAPLAN, R.; NORTON, D. The Balanced Scorecard: Measures That Drive Performance. **Harvard Business Review**, [s. l.], , p. 71–79, 1992. .

KAUFMAN, S.; ROSSET, S.; PERLICH, C.; STITELMAN, O. Leakage in data mining: Formulation, detection, and avoidance. **ACM Trans. Knowl. Discov. Data**, New York, NY, USA, vol. 6, no. 4, 2012. DOI 10.1145/2382577.2382579. Available at: <https://doi.org/10.1145/2382577.2382579>.

KI, I.; SONG, H.; RYU, J.; JEONG, J. Production Improvement Rate with Time Series Data on Standard Time at Manufacturing Sites. **Applied Sciences (Switzerland)**, [s. l.], vol. 13, no. 19, 2023. <https://doi.org/10.3390/app131910937>.

KOHAVI, R.; JOHN, G. H. Wrappers for feature subset selection. **Artificial Intelligence**, [s. l.], vol. 97, no. 1, p. 273–324, 1997. DOI [https://doi.org/10.1016/S0004-3702\(97\)00043-X](https://doi.org/10.1016/S0004-3702(97)00043-X). Available at: <https://www.sciencedirect.com/science/article/pii/S000437029700043X>.

KUHN, M.; JOHNSON, K. **Applied predictive modeling**. [S. l.: s. n.], 2013.

<https://doi.org/10.1007/978-1-4614-6849-3>.

LAST, M.; MAIMON, O.; MINKOV, E. Improving stability of decision trees. **International Journal of Pattern Recognition and Artificial Intelligence**, [s. l.], vol. 16, no. 2, p. 145–159, 2002. <https://doi.org/10.1142/S0218001402001599>.

LEE RODGERS, J.; NICEWANDER, W. A. Thirteen Ways to Look at the Correlation Coefficient. **The American Statistician**, [s. l.], vol. 42, no. 1, p. 59–66, 1 Feb. 1988. DOI 10.1080/00031305.1988.10475524. Available at: <https://doi.org/10.1080/00031305.1988.10475524>.

LIN, H.; LI, M. Deep Learning. **Practitioner's Guide to Data Science**, [s. l.], , p. 259–336, 2023. <https://doi.org/10.1201/9781351132916-12>.

LIU, Y.-J.; NI, D.; SHAO, X.; GONG, D.-L.; LI, J.-J. A hierarchical model-based method for wafer level virtual metrology under process information deficiency. **Quality Engineering**, [s. l.], vol. 36, p. 534 – 547, 2024. .

LUCIANO, E. L.; DA SILVA, J. H.; RIBEIRO, R. B.; SIMOES, E. A.; FORMIGONI, A.; DOS SANTOS, E. F. Analysis of Time Measurement Strategies in the Automotive Components Industry Using Design Science Research. **International Journal of Business Administration**, [s. l.], vol. 13, no. 6, p. 14, 2022. <https://doi.org/10.5430/ijba.v13n6p14>.

MAGDZIAK, M. Estimating Time of Coordinate Measurements Based on the Adopted Measurement Strategy. **Sensors**, [s. l.], vol. 22, no. 19, 2022. DOI 10.3390/s22197310. Available at: <https://www.mdpi.com/1424-8220/22/19/7310>.

MAGKLARAS, A.; GOGOS, C.; ALEFRAGIS, P.; BIRBAS, A. Enhancing Parameters Tuning of Overlay Models with Ridge Regression: Addressing Multicollinearity in High-Dimensional Data. **Mathematics**, [s. l.], vol. 12, no. 20, 2024. <https://doi.org/10.3390/math12203179>.

MAZAREINEZHAD, F.; SEKKAY, F.; IMBEAU, D. Development of a framework to implement time analysis in digital human modeling systems using predetermined motion time systems. **Discover Applied Sciences**, [s. l.], vol. 7, no. 5, 2025. DOI 10.1007/s42452-025-06813-9. Available at: <https://doi.org/10.1007/s42452-025-06813-9>.

MENTZER, J. T.; FLINT, D. J. Validity in logistics research. [s. l.], vol. 18, no. 1, p. 199–216, 1997. .

MONTESINOS LÓPEZ, O. A.; MONTESINOS LÓPEZ, A.; CROSSA, J. Fundamentals of Artificial Neural Networks and Deep Learning. **Multivariate Statistical Machine Learning Methods for Genomic Prediction**. Cham: Springer, 2022. DOI 10.1007/978-3-030-89010-0_10. Available at:

<https://www.ncbi.nlm.nih.gov/books/NBK583971/>.

NATEKIN, A.; KNOLL, A. Gradient boosting machines, a tutorial. **Frontiers in neurorobotics**, Switzerland, vol. 7, p. 21, 2013. <https://doi.org/10.3389/fnbot.2013.00021>.

NEELY, A.; GREGORY, M.; PLATTS, K. Performance measurement system design: A literature review and research agenda. **International Journal of Operations and Production Management**, [s. l.], vol. 25, no. 12, p. 1228–1263, 1995. <https://doi.org/10.1108/01443570510633639>.

NUR, A. R.; JAYA, A. K.; SISWANTO, S. Comparative Analysis of Ridge, LASSO, and Elastic Net Regularization Approaches in Handling Multicollinearity for Infant Mortality Data in South Sulawesi. **Jurnal Matematika, Statistika dan Komputasi**, [s. l.], vol. 20, no. 2, p. 311–319, 2023. <https://doi.org/10.20956/j.v20i2.31632>.

OHNO, T. **Toyota Production System: Beyond Large-Scale Production**. Portland, OR: Productivity Press, 1988.

ONARAN, E.; YANIK, S. Predicting cycle times in textile manufacturing using artificial neural network. **Advances in Intelligent Systems and Computing**, [s. l.], vol. 1029, p. 305–312, 2020. https://doi.org/10.1007/978-3-030-23756-1_38.

ÖRDEK, B.; BORGIANNI, Y.; COATANEA, E. Machine learning-supported manufacturing: a review and directions for future research. **Production and Manufacturing Research**, [s. l.], vol. 12, no. 1, p. 1–50, 2024. DOI 10.1080/21693277.2024.2326526. Available at: <https://doi.org/10.1080/21693277.2024.2326526>.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COUNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine Learning in {P}ython. **Journal of Machine Learning Research**, [s. l.], vol. 12, p. 2825–2830, 2011. .

PROKHORENKOVA, L.; GUSEV, G.; VOROBEOV, A.; DOROGUSH, A. V.; GULIN, A. CatBoost: unbiased boosting with categorical features. [s. l.], 2019. Available at: <https://arxiv.org/abs/1706.09516>.

RETZLAFF, C. O.; ANGERSCHMID, A.; SARANTI, A.; SCHNEEBERGER, D.; RÖTTGER, R.; MÜLLER, H.; HOLZINGER, A. Post-hoc vs ante-hoc explanations: xAI design guidelines for data scientists. **Cognitive Systems Research**, [s. l.], vol. 86, no. April, p. 101243, 2024. DOI 10.1016/j.cogsys.2024.101243. Available at: <https://doi.org/10.1016/j.cogsys.2024.101243>.

ROSA, R. C.; GONÇALVES, M. C.; BARBALHO, S. C. M. Machine learning applied

to forecasting the manufacturing time of new products prototypes and ETO products: An exploratory study. **International Journal of Production Economics**, [s. l.], vol. 287, no. August 2024, 2025. <https://doi.org/10.1016/j.ijpe.2025.109688>.

ROUSTAEI, N. Application and interpretation of linear-regression analysis. **Medical Hypothesis, Discovery, and Innovation in Ophthalmology**, [s. l.], vol. 13, no. 3, p. 151–159, 2024. <https://doi.org/10.51329/mehdiophthal1506>.

SCHALL MARK C., J.; FETHKE, N. B.; ROEMIG, V. Digital Human Modeling in the Occupational Safety and Health Process: An Application in Manufacturing. **IIE Transactions on Occupational Ergonomics and Human Factors**, [s. l.], vol. 6, no. 2, p. 64–75, 2018. DOI 10.1080/24725838.2018.1491430. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6455952/>.

SCHMITT, A.-M.; MILLER, E.; ENGELMANN, B.; BATRES, R.; SCHMITT, J. G-code evaluation in CNC milling to predict energy consumption through Machine Learning. **Advances in Industrial and Manufacturing Engineering**, [s. l.], vol. 8, p. 100140, 2024. DOI <https://doi.org/10.1016/j.aime.2024.100140>. Available at: <https://www.sciencedirect.com/science/article/pii/S2666912924000059>.

SCHOBBER, P.; BOER, C.; SCHWARTE, L. A. Correlation Coefficients: Appropriate Use and Interpretation. **Anesthesia and analgesia**, United States, vol. 126, no. 5, p. 1763–1768, May 2018. <https://doi.org/10.1213/ANE.0000000000002864>.

SHI, Y.; KE, G.; SOUKHAVONG, D.; LAMB, J.; MENG, Q.; FINLEY, T.; WANG, T.; CHEN, W.; MA, W.; YE, Q.; LIU, T.-Y.; TITOV, N.; CORTES, D. lightgbm: Light Gradient Boosting Machine. [s. l.], 2025. Available at: <https://github.com/Microsoft/LightGBM>.

SHINGO, S. **A Revolution in Manufacturing: The SMED System**. Portland, OR: Productivity Press, 1985.

SHMUELI, G. To explain or to predict? **Statistical Science**, [s. l.], vol. 25, no. 3, p. 289–310, 2010. <https://doi.org/10.1214/10-STS330>.

SHRESTHA, N. Detecting Multicollinearity in Regression Analysis. **American Journal of Applied Mathematics and Statistics**, [s. l.], vol. 8, p. 39–42, 2020. <https://doi.org/10.12691/ajams-8-2-1>.

SILVA, A. V.; COIMBRA, R. R. C. **Manual de Tempos e Métodos**. São Paulo: Hemus, 1980.

ŞİMŞİT, Z. T.; GÜNAY, N. S.; VAYVAY, Ö. Theory of Constraints: A Literature Review. **Procedia - Social and Behavioral Sciences**, [s. l.], vol. 150, p. 930–936, 2014. DOI

<https://doi.org/10.1016/j.sbspro.2014.09.104>. Available at:
<https://www.sciencedirect.com/science/article/pii/S1877042814051532>.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, [s. l.], vol. 14, no. 3, p. 199–222, 2004. DOI 10.1023/B:STCO.0000035301.49549.88. Available at:
<https://doi.org/10.1023/B:STCO.0000035301.49549.88>.

STEFANO, G. da S. DOES THE THEORY OF CONSTRAINTS IN SUPPLY CHAIN MANAGEMENT REALLY MATTER? An Assessment of the Impactsof the TOC in the Redesignof a Supply Chain. [s. l.], , p. 1–219, 2020. .

STROBL, C.; BOULESTEIX, A.-L.; ZEILEIS, A.; HOTHORN, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. **BMC Bioinformatics**, [s. l.], vol. 8, no. 1, p. 25, 2007. DOI 10.1186/1471-2105-8-25. Available at:
<https://doi.org/10.1186/1471-2105-8-25>.

SUN, C.; DOMINGUEZ-CABALLERO, J.; WARD, R.; AYVAR-SOBERANIS, S.; CURTIS, D. Machining cycle time prediction: Data-driven modelling of machine tool feedrate behavior with neural networks. **Robotics and Computer-Integrated Manufacturing**, [s. l.], vol. 73, 2022. .

SZUKITS, Á.; MORICZ, P. Towards data-driven decision making: the role of analytical culture and centralization efforts. **Review of Managerial Science**, [s. l.], vol. 18, p. 2849–2887, 2023. <https://doi.org/10.1007/s11846-023-00694-1>.

TIBSHIRANI, R. Regression Shrinkage and Selection via the Lasso. **Journal of the Royal Statistical Society. Series B (Methodological)**, [s. l.], vol. 58, no. 1, p. 267–288, 1996. Available at: [jstor.org/stable/2346178](https://www.jstor.org/stable/2346178).

TUKEY, J. W. **Exploratory data analysis**. Reading, Mass: Addison-Wesley Pub. Co., 1977(Addison-Wesley series in behavioral science).

URBAN, M.; KOBLASA, F.; MENDŘICKÝ, R. Machine Learning in Small and Medium-Sized Enterprises, Methodology for the Estimation of the Production Time. **Applied Sciences**, [s. l.], vol. 14, no. 19, 2024. <https://doi.org/10.3390/app14198608>.

VARMA, S.; SIMON, R. Bias in error estimation when using cross-validation for model selection. **BMC Bioinformatics**, [s. l.], vol. 7, no. 1, p. 91, 2006. DOI 10.1186/1471-2105-7-91. Available at: <https://doi.org/10.1186/1471-2105-7-91>.

VOSS, C.; TSIKRIKTSIS, N.; FROHLICH, M. Case Research in Operations Management. **International Journal of Operations & Production Management**, [s. l.], vol. 22, p. 195–219, 2002. <https://doi.org/10.1108/01443570210414329>.

WASHINGTON STATE DEPARTMENT OF AGRICULTURE (WSDA). Metrology Laboratory. 2025. Available at: <https://agr.wa.gov/departments/laboratories/metrology-lab>. Accessed on: 10 Aug. 2025.

WELLER, J.; ROESMANN, D.; EGGERT, S.; VON ENZBERG, S.; GRÄSSLER, I.; DUMITRESCU, R. Identification and prediction of standard times in machining for precision steel tubes through the usage of data analytics. **Procedia CIRP**, [s. l.], vol. 119, p. 514–520, 2023. DOI 10.1016/j.procir.2023.01.011. Available at: <https://doi.org/10.1016/j.procir.2023.01.011>.

WILKINSON, L.; FRIENDLY, M. The History of the Cluster Heat Map. **The American Statistician**, [s. l.], vol. 63, p. 179–184, 2009. <https://doi.org/10.1198/tas.2009.0033>.

WOMACK, J. P.; JONES, D. T. **Lean Thinking, Banish Waste and Create Wealth in Your Corporation**. 2nd ed. [S. l.: s. n.], 2003.

YIN, R. K. **Case study research and applications : design and methods**. 6th edition. Los Angeles SE - XXX, 319: Sage, 2018. <https://doi.org/LK> - <https://worldcat.org/title/1015881135>.

ZHOU, D. X. On grouping effect of elastic net. **Statistics and Probability Letters**, [s. l.], vol. 83, no. 9, p. 2108–2112, 2013. DOI 10.1016/j.spl.2013.05.014. Available at: <http://dx.doi.org/10.1016/j.spl.2013.05.014>.

ZHU, L.; JOHNSON, C.; MEJVIK, J.; VARISCO, M.; SCHIRALDI, M. Key performance indicators for manufacturing operations management in the process industry. **IEEE International Conference on Industrial Engineering and Engineering Management**, [s. l.], vol. 2017-Decem, no. December, p. 969–973, 2017. <https://doi.org/10.1109/IEEM.2017.8290036>.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society. Series B: Statistical Methodology**, [s. l.], vol. 67, no. 2, p. 301–320, 2005. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>.

AGGARWAL, C. C.; HINNEBURG, A.; KEIM, D. A. On the surprising behavior of distance metrics in high dimensional space. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, [s. l.], vol. 1973, p. 420–434, 2001. https://doi.org/10.1007/3-540-44503-x_27.

AGRAWAL, R.; SRIKANT, R. Fast Algorithms for Mining Association Rules in Large Databases. 1994. **Proceedings of the 20th International Conference on Very Large Data Bases** [...]. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1994. p. 487–

499.

ALTMEIER, M.; KUNKEL, K.; MOYA-QUIROGA, R.; WIESNER, F.; MÜLLER, N. Indirect manufacturing costs: An overlooked source for clear savings. **McKinsey & Company**, [s. l.], no. January, 2023. .

ARORA, I.; KHANDUJA, N.; BANSAL, M. Effect of Distance Metric and Feature Scaling on KNN Algorithm while Classifying X-rays. **CEUR Workshop Proceedings**, [s. l.], vol. 3176, p. 61–75, 2022. .

BARNES, R. M. **Estudo de Movimentos e de Tempos: Projeto e Medida do Trabalho**. São Paulo: Edgard Blücher, 1977.

BENTÉJAC, C.; CSÖRGŐ, A.; MARTÍNEZ-MUÑOZ, G. A comparative analysis of gradient boosting algorithms. **Artificial Intelligence Review**, [s. l.], vol. 54, no. 3, p. 1937–1967, 2021. DOI 10.1007/s10462-020-09896-5. Available at: <https://doi.org/10.1007/s10462-020-09896-5>.

BERTRAND, J. W. M.; FRANSOO, J. C.; UDENIO, M. Model-based research. In: KARLSSON, C. (ed.). **Research Methods for Operations and Supply Chain Management**. 3rd ed. Abingdon; New York: Routledge, 2023. p. 274–311. DOI 10.4324/9781003315001. Available at: <https://doi.org/10.4324/9781003315001>.

BOURNE, M.; NEELY, A.; MILLS, J.; PLATTS, K. Implementing Performance Measurement Systems: A Literature Review. **Int. J. Business Performance Management Int. J. Business Performance Management**, [s. l.], vol. 5, p. 1–24, 2003. <https://doi.org/10.1504/IJBPM.2003.002097>.

BREIMAN, L. Random Forests. **Machine Learning**, [s. l.], vol. 45, no. 1, p. 5–32, 2001. DOI 10.1023/A:1010933404324. Available at: <https://doi.org/10.1023/A:1010933404324>.

BREZNIK, M.; BUCHMEISTER, B.; VUJICA HERZOG, N. Assembly Line Optimization Using MTM Time Standard and Simulation Modeling—A Case Study. **Applied Sciences (Switzerland)**, [s. l.], vol. 13, no. 10, 2023. <https://doi.org/10.3390/app13106265>.

BUCHMEISTER, B.; HERZOG, N. V. Advancements in Data Analysis for the Work-Sampling Method. **Algorithms**, [s. l.], vol. 17, no. 5, 2024. <https://doi.org/10.3390/a17050183>.

BUITINCK, L.; LOUPPE, G.; BLONDEL, M.; PEDREGOSA, F.; MUELLER, A.; GRISEL, O.; NICULAE, V.; PRETTENHOFER, P.; GRAMFORT, A.; GROBLER, J.; LAYTON, R.; VANDERPLAS, J.; JOLY, A.; HOLT, B.; VAROQUAUX, G. {API} design for machine learning software: experiences from the scikit-learn project. 2013. **ECML PKDD Workshop: Languages for Data Mining and Machine Learning [...]**. [S. l.: s. n.], 2013. p.

108–122.

ÇAKLT, E.; DAĞDEVIREN, M. Comparative analysis of machine learning algorithms for predicting standard time in a manufacturing environment. **Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM**, [s. l.], vol. 37, 2023. <https://doi.org/10.1017/S0890060422000245>.

CHEHADE, M.; SYLLA, A.; DIALLO, ABDOUL RAHIME DOREMUS, Y. Manufacturing time estimation for offer pricing: A machine learning application in a French metallurgy industry. **Engineering Applications of Artificial Intelligence**, [s. l.], vol. 137, 2024. .

CHEN, T.; GUESTRIN, C. XGBoost: A scalable tree boosting system. **Association for Computing Machinery**, [s. l.], vol. 13-17-Aug, p. 785–794, 2016. <https://doi.org/10.1145/2939672.2939785>.

CHENG, C. F.; LIN, C. J.; HU, Q. X. Machine learning models for sequential motion recognition in human activity for logistics operations efficiency improvement. **Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM**, [s. l.], vol. 39, 2025. <https://doi.org/10.1017/S0890060424000313>.

CHOUDHARY, A.; HARDING, J.; TIWARI, M. Data mining in manufacturing: A review based on the kind of knowledge. **Journal of Intelligent Manufacturing**, [s. l.], vol. 20, p. 501–521, 2009. <https://doi.org/10.1007/s10845-008-0145-x>.

CRESWELL, J. W. Research design: Qualitative, quantitative, and mixed methods approaches, 3rd ed. **Research design: Qualitative, quantitative, and mixed methods approaches, 3rd ed.**, Thousand Oaks, CA, US, , p. xxix, 260–xxix, 260, 2009. .

CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of Control, Signals, and Systems**, [s. l.], vol. 2, no. 4, p. 303–314, 1989. <https://doi.org/10.1007/BF02551274>.

DA SILVA, A. G.; GOMES, M. V. M.; WINKLER, I. Virtual Reality and Digital Human Modeling for Ergonomic Assessment in Industrial Product Development: A Patent and Literature Review. **Applied Sciences (Switzerland)**, [s. l.], vol. 12, no. 3, 2022. <https://doi.org/10.3390/app12031084>.

DAMSCHRODER, L. J.; REARDON, C. M.; WIDERQUIST, M. A. O.; LOWERY, J. The updated Consolidated Framework for Implementation Research based on user feedback. **Implementation Science**, [s. l.], vol. 17, no. 1, p. 75, 2022. DOI 10.1186/s13012-022-01245-0. Available at: <https://doi.org/10.1186/s13012-022-01245-0>.

DAS, B. Personal, Unavoidable Delay, and Fatigue Allowances for Setting

Performance Standards. **International Journal of Operations & Production Management**, [s. l.], vol. 10, no. 5, p. 62–70, 1990. DOI 10.1108/01443579010143477. Available at: <https://doi.org/10.1108/01443579010143477>.

DAVENPORT, T. H.; HARRIS, J. G. **Competing on Analytics: The New Science of Winning**. 1st ed. USA: Harvard Business School Press, 2007.

DORMANN, C. F.; ELITH, J.; BACHER, S.; BUCHMANN, C.; CARL, G.; CARRÉ, G.; MARQUÉZ, J. R. G.; GRUBER, B.; LAFOURCADE, B.; LEITÃO, P. J.; MÜNKEMÜLLER, T.; MCCLEAN, C.; OSBORNE, P. E.; REINEKING, B.; SCHRÖDER, B.; SKIDMORE, A. K.; ZURELL, D.; LAUTENBACH, S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. **Ecography**, [s. l.], vol. 36, no. 1, p. 27–46, 2013. DOI <https://doi.org/10.1111/j.1600-0587.2012.07348.x>. Available at: <https://nsojournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0587.2012.07348.x>.

DOUGHERTY, J.; KOHAVI, R.; SAHAMI, M. Supervised and Unsupervised Discretization of Continuous Features. In: PRIEDITIS, A.; RUSSELL, S. (eds.). **Machine Learning Proceedings 1995**. San Francisco (CA): Morgan Kaufmann, 1995. p. 194–202. DOI <https://doi.org/10.1016/B978-1-55860-377-6.50032-3>. Available at: <https://www.sciencedirect.com/science/article/pii/B9781558603776500323>.

DRESCH, A.; LACERDA, D.; ANTUNES JÚNIOR, J. A. V. Desing Science Research: A Method for Science and Technology Advancement. [s. l.], 2014. https://doi.org/10.1007/978-3-319-07374-3_4.

DREYFUS, P. A.; PSAROMMATIS, F.; MAY, G.; KIRITSIS, D. Virtual metrology as an approach for product quality estimation in Industry 4.0: a systematic review and integrative conceptual framework. **International Journal of Production Research**, [s. l.], vol. 60, no. 2, p. 742–765, 2022. DOI 10.1080/00207543.2021.1976433. Available at: <https://doi.org/10.1080/00207543.2021.1976433>.

EISENHARDT, K.; GRAEBNER, M. Theory Building From Cases: Opportunities And Challenges. **The Academy of Management Journal**, [s. l.], vol. 50, p. 25–32, 2007. <https://doi.org/10.5465/AMJ.2007.24160888>.

EMMERT-STREIB, F.; DEHMER, M. High-Dimensional LASSO-Based Computational Regression Models: Regularization, Shrinkage, and Selection. **Machine Learning and Knowledge Extraction**, [s. l.], vol. 1, no. 1, p. 359–383, 2019. <https://doi.org/10.3390/make1010021>.

ERMEL, A. P. C. **LITERATURE GROUNDED THEORY: MÉTODO DE PESQUISA PARA INVESTIGAÇÃO SOBRE O CONHECIMENTO CIENTÍFICO E**

TECNOLÓGICO. 2021. Universidade do Vale do Rio dos Sinos, 2021. Available at: <https://www.repositorio.jesuita.org.br/handle/UNISINOS/9270>.

FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Mag.**, USA, vol. 17, no. 3, p. 37–54, 1996. DOI 10.1609/aimag.v17i3.1230. Available at: <https://doi.org/10.1609/aimag.v17i3.1230>.

FLORES-HUAMÁN, K.-J.; LORENZO-ESPEJO, A.; MUÑOZ-DÍAZ, M.-L.; ESCUDERO-SANTANA, A. Enhancing process lead time forecasting with machine learning and upstream process data: A case study in wind tower manufacturing. **Computers & Industrial Engineering**, [s. l.], vol. 209, no. July, p. 111410, 2025. DOI 10.1016/j.cie.2025.111410. Available at: <https://doi.org/10.1016/j.cie.2025.111410>.

FLYVBJERG, B. Five Misunderstandings About Case-Study Research. **Qualitative Inquiry**, [s. l.], vol. 12, p. 219–245, 2006. <https://doi.org/10.1177/1077800405284363>.

FREIJEIRO-GONZÁLEZ, L.; FEBRERO-BANDE, M.; GONZÁLEZ-MANTEIGA, W. A Critical Review of LASSO and Its Derivatives for Variable Selection Under Dependence Among Covariates. **International Statistical Review**, [s. l.], vol. 90, no. 1, p. 118–145, 2022. DOI <https://doi.org/10.1111/insr.12469>. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12469>.

FREVALDS, A.; NIEBEL, B. W. **Niebel's methods, standards, and work design**. [S. l.: s. n.], 2014.

GAWLIK, A.; LÜDEMANN, J.; NEUHAUSEN, A.; ZEPP, C.; VITINIUS, F.; KLEINERT, J. A Systematic Review of Workplace Physical Activity Coaching. **Journal of occupational rehabilitation**, Netherlands, vol. 33, no. 3, p. 550–569, Sep. 2023. <https://doi.org/10.1007/s10926-023-10093-8>.

GÖDRI, I. Improving Delivery Performance in High-Mix Low-Volume Manufacturing by Model-Based and Data-Driven Methods. **Applied Sciences (Switzerland)**, [s. l.], vol. 12, no. 11, 2022. <https://doi.org/10.3390/app12115618>.

GOMES RIBEIRO, T.; CARLA DE SOUZA GOMES DOS SANTOS, A.; LOPES HERMSDORFF, W.; AUGUSTO FARIA SALLES, S.; DA CUNHA REIS, A. Controle Estatístico Do Processo E Engenharia De Métodos: Uma Aplicação Conjunta Na Produção De Picolés Statistical Process Control and Methods Engineering: a Joint Application in the Production of Popsicles. [s. l.], 2019. .

GOODWIN, M. A.; STANGE, K. C.; ZYZANSKI, S. J.; CRABTREE, B. F.; BORAWSKI, E. A.; FLOCKE, S. A. The Hawthorne Effect in Direct Observation Research with Physicians and Patients. **Journal of Evaluation in Clinical Practice**, [s. l.], 2017.

<https://doi.org/10.1111/jep.12781>.The.

GORI, R. S. L. **Modelo de distribuição de recursos orçamentários baseado em indicadores de desempenho para um Instituto Federal de Educação, Ciência e Tecnologia**. 2016. Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, RS, 2016.

GUEST, G.; BUNCE, A.; JOHNSON, L. How Many Interviews Are Enough? **Field Methods - FIELD METHOD**, [s. l.], vol. 18, p. 59–82, 2006. <https://doi.org/10.1177/1525822X05279903>.

GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. **J. Mach. Learn. Res.**, [s. l.], vol. 3, no. null, p. 1157–1182, 2003. .

HAN, J.; KAMBER, M.; PEI, J. **Data Mining: Concepts and Techniques**. 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.

HARDING, J. A.; SHAHBAZ, M.; SRINIVAS; KUSIAK, A. Data Mining in Manufacturing: A Review. **Journal of Manufacturing Science and Engineering**, [s. l.], vol. 128, no. 4, p. 969–976, 2005. DOI 10.1115/1.2194554. Available at: <https://doi.org/10.1115/1.2194554>.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. [S. l.]: Springer, 2009(Springer series in statistics). Available at: <https://books.google.com.br/books?id=eBSgoAEACAAJ>.

HEITZ, R. P. The speed-accuracy tradeoff: history, physiology, methodology, and behavior. **Frontiers in neuroscience**, Switzerland, vol. 8, p. 150, 2014. <https://doi.org/10.3389/fnins.2014.00150>.

HO, F. K.; COLE, T. J. Non-linear predictor outcome associations. **BMJ Medicine**, [s. l.], vol. 2, no. 1, p. e000396, 2023. <https://doi.org/10.1136/bmjmed-2022-000396>.

HORNIK, K. Approximation capabilities of multilayer feedforward networks. **Neural Networks**, [s. l.], vol. 4, no. 2, p. 251–257, 1991. [https://doi.org/10.1016/0893-6080\(91\)90009-T](https://doi.org/10.1016/0893-6080(91)90009-T).

INAL, T. C.; GORUROGLU OZTURK, O.; KIBAR, F.; CETINER, S.; MATYAR, S.; DAGLIOGLU, G.; YAMAN, A. Lean six sigma methodologies improve clinical laboratory efficiency and reduce turnaround times. **Journal of clinical laboratory analysis**, United States, vol. 32, no. 1, Jan. 2018. <https://doi.org/10.1002/jcla.22180>.

ISO. **ISO/IEC 17025:2017 — General requirements for the competence of testing and calibration laboratories**. [S. l.: s. n.], 2017.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning: with Applications in R**. [S. l.]: Springer Publishing Company,

Incorporated, 2014.

JI, J.; PANNAKONG, W.; BUDDHAKULSOMSIRI, J. A Computer Vision-Based Model for Automatic Motion Time Study. **Computers, Materials and Continua**, [s. l.], vol. 73, no. 2, p. 3557–3574, 2022. <https://doi.org/10.32604/cmc.2022.030418>.

KALNE, P. S.; MEHENDALE, A. M. The Purpose of Time-Motion Studies (TMSs) in Healthcare: A Literature Review. **Cureus**, [s. l.], vol. 14, no. 10, 2022. <https://doi.org/10.7759/cureus.29869>.

KANAWATY, G. **Introduction to Work Study**. 4th ed. Geneva: International Labour Office, 1992. Available at: <https://hcmindonesia.files.wordpress.com/2012/12/introduction-to-work-study.pdf>.

KAPLAN, R.; NORTON, D. The Balanced Scorecard: Measures That Drive Performance. **Harvard Business Review**, [s. l.], , p. 71–79, 1992. .

KAUFMAN, S.; ROSSET, S.; PERLICH, C.; STITELMAN, O. Leakage in data mining: Formulation, detection, and avoidance. **ACM Trans. Knowl. Discov. Data**, New York, NY, USA, vol. 6, no. 4, 2012. DOI 10.1145/2382577.2382579. Available at: <https://doi.org/10.1145/2382577.2382579>.

KI, I.; SONG, H.; RYU, J.; JEONG, J. Production Improvement Rate with Time Series Data on Standard Time at Manufacturing Sites. **Applied Sciences (Switzerland)**, [s. l.], vol. 13, no. 19, 2023. <https://doi.org/10.3390/app131910937>.

KOHAVI, R.; JOHN, G. H. Wrappers for feature subset selection. **Artificial Intelligence**, [s. l.], vol. 97, no. 1, p. 273–324, 1997. DOI [https://doi.org/10.1016/S0004-3702\(97\)00043-X](https://doi.org/10.1016/S0004-3702(97)00043-X). Available at: <https://www.sciencedirect.com/science/article/pii/S000437029700043X>.

KUHN, M.; JOHNSON, K. **Applied predictive modeling**. [S. l.: s. n.], 2013. <https://doi.org/10.1007/978-1-4614-6849-3>.

LAST, M.; MAIMON, O.; MINKOV, E. Improving stability of decision trees. **International Journal of Pattern Recognition and Artificial Intelligence**, [s. l.], vol. 16, no. 2, p. 145–159, 2002. <https://doi.org/10.1142/S0218001402001599>.

LEE RODGERS, J.; NICEWANDER, W. A. Thirteen Ways to Look at the Correlation Coefficient. **The American Statistician**, [s. l.], vol. 42, no. 1, p. 59–66, 1 Feb. 1988. DOI 10.1080/00031305.1988.10475524. Available at: <https://doi.org/10.1080/00031305.1988.10475524>.

LIN, H.; LI, M. Deep Learning. **Practitioner's Guide to Data Science**, [s. l.], , p. 259–336, 2023. <https://doi.org/10.1201/9781351132916-12>.

LIU, Y.-J.; NI, D.; SHAO, X.; GONG, D.-L.; LI, J.-J. A hierarchical model-based method for wafer level virtual metrology under process information deficiency. **Quality Engineering**, [s. l.], vol. 36, p. 534 – 547, 2024. .

LUCIANO, E. L.; DA SILVA, J. H.; RIBEIRO, R. B.; SIMOES, E. A.; FORMIGONI, A.; DOS SANTOS, E. F. Analysis of Time Measurement Strategies in the Automotive Components Industry Using Design Science Research. **International Journal of Business Administration**, [s. l.], vol. 13, no. 6, p. 14, 2022. <https://doi.org/10.5430/ijba.v13n6p14>.

MAGDZIAK, M. Estimating Time of Coordinate Measurements Based on the Adopted Measurement Strategy. **Sensors**, [s. l.], vol. 22, no. 19, 2022. DOI 10.3390/s22197310. Available at: <https://www.mdpi.com/1424-8220/22/19/7310>.

MAGKLARAS, A.; GOGOS, C.; ALEFRAGIS, P.; BIRBAS, A. Enhancing Parameters Tuning of Overlay Models with Ridge Regression: Addressing Multicollinearity in High-Dimensional Data. **Mathematics**, [s. l.], vol. 12, no. 20, 2024. <https://doi.org/10.3390/math12203179>.

MAZAREINEZHAD, F.; SEKKAY, F.; IMBEAU, D. Development of a framework to implement time analysis in digital human modeling systems using predetermined motion time systems. **Discover Applied Sciences**, [s. l.], vol. 7, no. 5, 2025. DOI 10.1007/s42452-025-06813-9. Available at: <https://doi.org/10.1007/s42452-025-06813-9>.

MENTZER, J. T.; FLINT, D. J. Validity in logistics research. [s. l.], vol. 18, no. 1, p. 199–216, 1997. .

MONTESINOS LÓPEZ, O. A.; MONTESINOS LÓPEZ, A.; CROSSA, J. Fundamentals of Artificial Neural Networks and Deep Learning. **Multivariate Statistical Machine Learning Methods for Genomic Prediction**. Cham: Springer, 2022. DOI 10.1007/978-3-030-89010-0_10. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK583971/>.

NATEKIN, A.; KNOLL, A. Gradient boosting machines, a tutorial. **Frontiers in neurorobotics**, Switzerland, vol. 7, p. 21, 2013. <https://doi.org/10.3389/fnbot.2013.00021>.

NEELY, A.; GREGORY, M.; PLATTS, K. Performance measurement system design: A literature review and research agenda. **International Journal of Operations and Production Management**, [s. l.], vol. 25, no. 12, p. 1228–1263, 1995. <https://doi.org/10.1108/01443570510633639>.

NUR, A. R.; JAYA, A. K.; SISWANTO, S. Comparative Analysis of Ridge, LASSO, and Elastic Net Regularization Approaches in Handling Multicollinearity for Infant Mortality Data in South Sulawesi. **Jurnal Matematika, Statistika dan Komputasi**, [s. l.], vol. 20, no.

2, p. 311–319, 2023. <https://doi.org/10.20956/j.v20i2.31632>.

OHNO, T. **Toyota Production System: Beyond Large-Scale Production**. Portland, OR: Productivity Press, 1988.

ONARAN, E.; YANIK, S. Predicting cycle times in textile manufacturing using artificial neural network. **Advances in Intelligent Systems and Computing**, [s. l.], vol. 1029, p. 305–312, 2020. https://doi.org/10.1007/978-3-030-23756-1_38.

ÖRDEK, B.; BORGIANNI, Y.; COATANEA, E. Machine learning-supported manufacturing: a review and directions for future research. **Production and Manufacturing Research**, [s. l.], vol. 12, no. 1, p. 1–50, 2024. DOI 10.1080/21693277.2024.2326526. Available at: <https://doi.org/10.1080/21693277.2024.2326526>.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine Learning in {P}ython. **Journal of Machine Learning Research**, [s. l.], vol. 12, p. 2825–2830, 2011. .

PROKHORENKOVA, L.; GUSEV, G.; VOROBEOV, A.; DOROGUSH, A. V.; GULIN, A. CatBoost: unbiased boosting with categorical features. [s. l.], 2019. Available at: <https://arxiv.org/abs/1706.09516>.

RETZLAFF, C. O.; ANGERSCHMID, A.; SARANTI, A.; SCHNEEBERGER, D.; RÖTTGER, R.; MÜLLER, H.; HOLZINGER, A. Post-hoc vs ante-hoc explanations: xAI design guidelines for data scientists. **Cognitive Systems Research**, [s. l.], vol. 86, no. April, p. 101243, 2024. DOI 10.1016/j.cogsys.2024.101243. Available at: <https://doi.org/10.1016/j.cogsys.2024.101243>.

ROSA, R. C.; GONÇALVES, M. C.; BARBALHO, S. C. M. Machine learning applied to forecasting the manufacturing time of new products prototypes and ETO products: An exploratory study. **International Journal of Production Economics**, [s. l.], vol. 287, no. August 2024, 2025. <https://doi.org/10.1016/j.ijpe.2025.109688>.

ROUSTAEI, N. Application and interpretation of linear-regression analysis. **Medical Hypothesis, Discovery, and Innovation in Ophthalmology**, [s. l.], vol. 13, no. 3, p. 151–159, 2024. <https://doi.org/10.51329/mehdiophthal1506>.

SCHALL MARK C., J.; FETHKE, N. B.; ROEMIG, V. Digital Human Modeling in the Occupational Safety and Health Process: An Application in Manufacturing. **IIE Transactions on Occupational Ergonomics and Human Factors**, [s. l.], vol. 6, no. 2, p. 64–75, 2018. DOI 10.1080/24725838.2018.1491430. Available at:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC6455952/>.

SCHMITT, A.-M.; MILLER, E.; ENGELMANN, B.; BATRES, R.; SCHMITT, J. G-code evaluation in CNC milling to predict energy consumption through Machine Learning. **Advances in Industrial and Manufacturing Engineering**, [s. l.], vol. 8, p. 100140, 2024. DOI <https://doi.org/10.1016/j.aime.2024.100140>. Available at:

<https://www.sciencedirect.com/science/article/pii/S2666912924000059>.

SCHOBER, P.; BOER, C.; SCHWARTE, L. A. Correlation Coefficients: Appropriate Use and Interpretation. **Anesthesia and analgesia**, United States, vol. 126, no. 5, p. 1763–1768, May 2018. <https://doi.org/10.1213/ANE.0000000000002864>.

SHI, Y.; KE, G.; SOUKHAVONG, D.; LAMB, J.; MENG, Q.; FINLEY, T.; WANG, T.; CHEN, W.; MA, W.; YE, Q.; LIU, T.-Y.; TITOV, N.; CORTES, D. lightgbm: Light Gradient Boosting Machine. [s. l.], 2025. Available at: <https://github.com/Microsoft/LightGBM>.

SHINGO, S. **A Revolution in Manufacturing: The SMED System**. Portland, OR: Productivity Press, 1985.

SHMUELI, G. To explain or to predict? **Statistical Science**, [s. l.], vol. 25, no. 3, p. 289–310, 2010. <https://doi.org/10.1214/10-STS330>.

SHRESTHA, N. Detecting Multicollinearity in Regression Analysis. **American Journal of Applied Mathematics and Statistics**, [s. l.], vol. 8, p. 39–42, 2020. <https://doi.org/10.12691/ajams-8-2-1>.

SILVA, A. V.; COIMBRA, R. R. C. **Manual de Tempos e Métodos**. São Paulo: Hemus, 1980.

ŞİMŞİT, Z. T.; GÜNAY, N. S.; VAYVAY, Ö. Theory of Constraints: A Literature Review. **Procedia - Social and Behavioral Sciences**, [s. l.], vol. 150, p. 930–936, 2014. DOI <https://doi.org/10.1016/j.sbspro.2014.09.104>. Available at: <https://www.sciencedirect.com/science/article/pii/S1877042814051532>.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, [s. l.], vol. 14, no. 3, p. 199–222, 2004. DOI [10.1023/B:STCO.0000035301.49549.88](https://doi.org/10.1023/B:STCO.0000035301.49549.88). Available at: <https://doi.org/10.1023/B:STCO.0000035301.49549.88>.

STEFANO, G. da S. DOES THE THEORY OF CONSTRAINTS IN SUPPLY CHAIN MANAGEMENT REALLY MATTER? An Assessment of the Impactsof the TOC in the Redesignof a Supply Chain. [s. l.], , p. 1–219, 2020. .

STROBL, C.; BOULESTEIX, A.-L.; ZEILEIS, A.; HOTHORN, T. Bias in random

forest variable importance measures: Illustrations, sources and a solution. **BMC Bioinformatics**, [s. l.], vol. 8, no. 1, p. 25, 2007. DOI 10.1186/1471-2105-8-25. Available at: <https://doi.org/10.1186/1471-2105-8-25>.

SUN, C.; DOMINGUEZ-CABALLERO, J.; WARD, R.; AYVAR-SOBERANIS, S.; CURTIS, D. Machining cycle time prediction: Data-driven modelling of machine tool feedrate behavior with neural networks. **Robotics and Computer-Integrated Manufacturing**, [s. l.], vol. 73, 2022. .

SZUKITS, Á.; MORICZ, P. Towards data-driven decision making: the role of analytical culture and centralization efforts. **Review of Managerial Science**, [s. l.], vol. 18, p. 2849–2887, 2023. <https://doi.org/10.1007/s11846-023-00694-1>.

TIBSHIRANI, R. Regression Shrinkage and Selection via the Lasso. **Journal of the Royal Statistical Society. Series B (Methodological)**, [s. l.], vol. 58, no. 1, p. 267–288, 1996. Available at: [jstor.org/stable/2346178](https://www.jstor.org/stable/2346178).

TUKEY, J. W. **Exploratory data analysis**. Reading, Mass: Addison-Wesley Pub. Co., 1977(Addison-Wesley series in behavioral science).

URBAN, M.; KOBLASA, F.; MENDŘICKÝ, R. Machine Learning in Small and Medium-Sized Enterprises, Methodology for the Estimation of the Production Time. **Applied Sciences**, [s. l.], vol. 14, no. 19, 2024. <https://doi.org/10.3390/app14198608>.

VARMA, S.; SIMON, R. Bias in error estimation when using cross-validation for model selection. **BMC Bioinformatics**, [s. l.], vol. 7, no. 1, p. 91, 2006. DOI 10.1186/1471-2105-7-91. Available at: <https://doi.org/10.1186/1471-2105-7-91>.

VOSS, C.; TSIKRIKTSIS, N.; FROHLICH, M. Case Research in Operations Management. **International Journal of Operations & Production Management**, [s. l.], vol. 22, p. 195–219, 2002. <https://doi.org/10.1108/01443570210414329>.

WASHINGTON STATE DEPARTMENT OF AGRICULTURE (WSDA). Metrology Laboratory. 2025. Available at: <https://agr.wa.gov/departments/laboratories/metrology-lab>. Accessed on: 10 Aug. 2025.

WELLER, J.; ROESMANN, D.; EGGERT, S.; VON ENZBERG, S.; GRÄSSLER, I.; DUMITRESCU, R. Identification and prediction of standard times in machining for precision steel tubes through the usage of data analytics. **Procedia CIRP**, [s. l.], vol. 119, p. 514–520, 2023. DOI 10.1016/j.procir.2023.01.011. Available at: <https://doi.org/10.1016/j.procir.2023.01.011>.

WILKINSON, L.; FRIENDLY, M. The History of the Cluster Heat Map. **The American Statistician**, [s. l.], vol. 63, p. 179–184, 2009.

<https://doi.org/10.1198/tas.2009.0033>.

WOMACK, J. P.; JONES, D. T. **Lean Thinking, Banish Waste and Create Wealth in Your Corporation**. 2nd ed. [S. l.: s. n.], 2003.

YIN, R. K. **Case study research and applications : design and methods**. 6th editio. Los Angeles SE - XXX, 319: Sage, 2018. <https://doi.org/LK> - <https://worldcat.org/title/1015881135>.

ZHOU, D. X. On grouping effect of elastic net. **Statistics and Probability Letters**, [s. l.], vol. 83, no. 9, p. 2108–2112, 2013. DOI 10.1016/j.spl.2013.05.014. Available at: <http://dx.doi.org/10.1016/j.spl.2013.05.014>.

ZHU, L.; JOHNSSON, C.; MEJVIK, J.; VARISCO, M.; SCHIRALDI, M. Key performance indicators for manufacturing operations management in the process industry. **IEEE International Conference on Industrial Engineering and Engineering Management**, [s. l.], vol. 2017-Decem, no. December, p. 969–973, 2017. <https://doi.org/10.1109/IEEM.2017.8290036>.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society. Series B: Statistical Methodology**, [s. l.], vol. 67, no. 2, p. 301–320, 2005. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>.