

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE GUARATINGUETÁ

ELEN YANINA AGUIRRE RODRÍGUEZ

**TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA PREDIÇÃO DO CUSTO DA
LOGÍSTICA DE TRANSPORTE: UMA APLICAÇÃO EM EMPRESA DO SEGMENTO DE
AUTOPEÇAS**

Guaratinguetá

2020

ELEN YANINA AGUIRRE RODRÍGUEZ

**TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA PREDIÇÃO DO CUSTO DA
LOGÍSTICA DE TRANSPORTE: UMA APLICAÇÃO EM EMPRESA DO SEGMENTO DE
AUTOPEÇAS**

Dissertação apresentada à Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista para obtenção do título de Mestre em Engenharia de Produção, na área de Gestão e Otimização.

Linha de Pesquisa: Modelagem Organizacional

Orientador:

Prof. Dr. Fernando Augusto Silva Marins

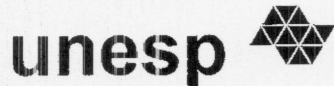
Coorientador:

Prof. Dr. Aneirson Francisco da Silva

Guaratinguetá

2020

R696t	<p>Rodríguez, Elen Yanina Aguirre</p> <p>Técnicas de aprendizado de máquina para predição do custo da logística de transporte: uma aplicação em empresa do segmento de autopeças / Elen Yanina Aguirre Rodríguez – Guaratinguetá, 2020. 120 f : il. Bibliografia: f. 86-94</p> <p>Dissertação (Mestrado) – Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, 2020. Orientador: Prof. Dr. Fernando Augusto Silva Marins Coorientador: Prof. Dr. Aneirson Francisco da Silva</p> <p>1. Processo decisório. 2. Otimização matemática. 3. Algoritmos de computador. I. Título</p> <p style="text-align: right;">CDU 65.012.4(043)</p>
-------	---



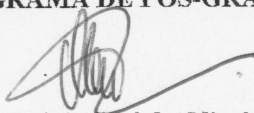
UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE GUARATINGUETÁ

ELEN YANINA AGUIRRE RODRIGUEZ

ESTA DISSERTAÇÃO FOI JULGADA ADEQUADA PARA A OBTENÇÃO DO TÍTULO DE
“MESTRE EM ENGENHARIA DE PRODUÇÃO”

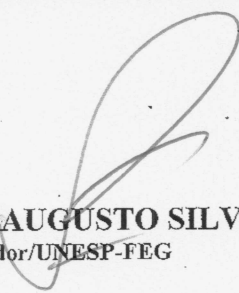
PROGRAMA: ENGENHARIA DE PRODUÇÃO
CURSO: MESTRADO ACADÊMICO

APROVADA EM SUA FORMA FINAL PELO PROGRAMA DE PÓS-GRADUAÇÃO



Prof. Dr. Otávio José de Oliveira
Coordenador

BANCA EXAMINADORA:



Prof. Dr. FERNANDO AUGUSTO SILVA MARINS
Orientador/UNESP-FEG



Prof. Dr. RAFAEL DE CARVALHO MIRANDA
UNIFEI



Prof.ª Dr.ª PALOMA MARIA SILVA ROCHA RIZOL
UNESP/FEG

Março de 2020

DADOS CURRICULARES

ELEN YANINA AGUIRRE RODRÍGUEZ

NASCIMENTO 17.08.1994 - CHEPÉN / PERÚ

FILIAÇÃO Elias Aguirre Tejada
Edita Rodríguez Ninaquispe

2011 / 2015 Curso de Graduação em Engenharia de
Sistemas
Universidade Nacional de Trujillo

Dedicado ao meu amor Alexander por ser um grande companheiro de vida, aos meus pais Elias e Edita pelo amor incondicional, aos meus irmãos Richard, Cristina, Elias e Christian pela cumplicidade e a parceria e, ao meu sobrinho Mattias que eu amo como se fosse filho.

AGRADECIMENTOS

Em primeiro lugar agradeço a Deus por tudo aquilo da minha vida.

Aos meus pais, por todo o amor e apoio incondicionais, por cultivar valores na minha formação e por ensinar-me a lutar por meus sonhos.

Aos meus irmãos, pelos momentos felizes, pela compreensão e, por estarem ao meu lado, mesmo estando distantes.

Ao meu amor Alexander, por me motivar a continuar sonhando e me acompanhar nos meus sonhos, pelo apoio incondicional, as palavras de carinho, por sempre acreditar em mim e por sempre ser compreensivo mesmo nas horas difíceis.

Ao meu orientador, o professor Dr. Fernando Silva Marins, por sua gentileza e preocupação no primeiro contato, pela oportunidade, pela motivação, pelo carinho, pela paciência e apoio constante para culminar o desenvolvimento desta pesquisa.

Ao meu co-orientador, Prof. Dr. Aneirson Francisco da Silva, pela paciência e orientação durante o desenvolvimento deste trabalho.

À Profa. Dra. Paloma Maria Silva Rocha Rizol e ao Prof. Dr. Rafael de Carvalho Miranda, pela contribuição e pela participação como membros da banca examinadora.

Aos professores da FEG-UNESP por sua contribuição em minha formação acadêmica.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES¹), pelo apoio financeiro para o desenvolvimento deste trabalho.

¹ <<http://www.capes.gov.br>>

Este trabalho contou com o apoio da seguinte entidade:

CAPES - Coordenação de Aperfeiçoamento de Pessoa de Nível Superior - Código de Financiamento
001

*“Todo mundo têm dentro de si um fragmento de boas notícias.
A boa notícia é que você não sabe quão extraordinário
você pode ser! O quanto você pode amar! O que você
pode executar! E qual é o seu potencial!”*

Ana Frank

RESUMO

Em diferentes aspectos da vida cotidiana, o ser humano é forçado a escolher entre várias opções, esse processo é conhecido como tomada de decisão. No nível do negócio, a tomada de decisões desempenha um papel muito importante, porque dessas decisões depende o sucesso ou o fracasso das organizações. No entanto, em muitos casos, tomar decisões erradas pode gerar grandes custos. Desta forma, alguns dos problemas de tomada de decisão que um gerente enfrenta comumente são, por exemplo, a decisão para determinar um preço, a decisão de comprar ou fabricar, em problemas de logística, problemas de armazenamento, etc. Por outro lado, a coleta de dados tornou-se uma vantagem competitiva, pois pode ser utilizada para análise e extração de resultados significativos por meio da aplicação de diversas técnicas, como estatística, simulação, matemática, econometria e técnicas atuais, como aprendizagem de máquina para a criação de modelos preditivos. Além disso, há evidências na literatura de que a criação de modelos com técnicas de aprendizagem de máquina têm um impacto positivo na indústria e em diferentes áreas de pesquisa. Nesse contexto, o presente trabalho propõe o desenvolvimento de um modelo preditivo para tomada de decisão, usando as técnicas supervisionadas de aprendizado de máquina, e combinando o modelo gerado com as restrições pertencentes ao processo de otimização. O objetivo da proposta é treinar um modelo matemático com dados históricos de um processo decisório e obter os preditores compostos por funções empíricas que serão posteriormente utilizadas e modeladas de acordo com as restrições do problema. Assim, este trabalho pode ser classificado como uma pesquisa aplicada, com objetivos empíricos descritivos e experiência prática que explicarão o modelo e suas vantagens. A maneira de abordar o problema deste trabalho será quantitativa, sendo os procedimentos técnicos de modelagem e simulação. A sistemática proposta é validada aplicando-se a um problema real em uma empresa multinacional brasileira do segmento de autopeças, situado no Vale do Paraíba-SP. O conjunto de dados vem do processo de transporte, visando aplicar a sistemática e criar um modelo matemático para prever o custo de transporte de uma quantidade de itens para um destino específico, de modo que este modelo finalmente ajuda a reduzir os tempos de estimativa e tomar melhores decisões. Finalmente, para a criação do modelo preditivo, foram testadas as técnicas de aprendizagem supervisionada de regressão linear, árvore de decisão CART, árvore de modelo M5P, máquina de vetor de suporte e redes neurais. Todos os modelos gerados foram avaliados por as métricas estatísticas de desempenho do erro médio absoluto, raiz do erro quadrático médio, erro absoluto relativo, erro quadrático relativo, coeficiente de correlação e coeficiente de determinação ou também conhecido como R^2 . Sendo que, o modelo M5P obteve os menores valores nas métricas de erro e os maiores valores na correlação e R^2 , mostrando ser um modelo eficiente, além de fornecer as equações necessárias para o processo de otimização.

PALAVRAS-CHAVE: Aprendizagem de Máquina. Otimização. Tomada de decisão. Árvore de decisão. Algoritmo M5P.

ABSTRACT

In different aspects of everyday life, the human being is forced to choose between several options, this process is known as decision making. At the business level, decision making plays a very important role, because the success or failure of organizations depends on these decisions. However, in many cases, making the wrong decisions can be costly. In this way, some of the decision-making problems that a manager commonly faces are, for example, the decision to determine a price, the decision to buy or manufacture, in logistics problems, storage problems, etc. On the other hand, data collection has become a competitive advantage, as it can be used for the analysis and extraction of significant results through the application of various techniques, such as statistics, simulation, mathematics, econometrics and current techniques, , such as machine learning to create predictive models. In addition, there is evidence in the literature that the creation of models with machine learning techniques has a positive impact on the industry and in different areas of research. In this context, the present work proposes the development of a predictive model for decision making, using supervised machine learning techniques, and combining the model generated with the constraints pertaining to the optimization process. The purpose of the proposal is to train a mathematical model with historical data from a decision-making process and obtain predictors composed of empirical functions that will later be used and modeled according to the constraints of the problem. Thus, this work can be classified as applied research, with descriptive empirical objectives and practical experience that will explain the model and its advantages. The way to approach the problem of this work will be quantitative, being the technical procedures of modeling and simulation. The proposed system is validated by applying it to a real problem in a Brazilian multinational company in the auto parts segment, located in Vale do Paraíba-SP. The data set comes from the transport process, aiming to apply the systematic and create a mathematical model to predict the cost of transporting a quantity of items to a specific destination, so that this model finally helps to reduce the estimation times and take better decisions. Finally, for the creation of the predictive model, the supervised learning techniques of linear regression, CART decision tree, M5P model tree, support vector machine and neural networks were tested. All the models generated were evaluated by the statistical metrics of performance of the mean absolute error, root of the mean square error, relative absolute error, relative quadratic error, correlation coefficient and determination coefficient or also known as R^2 . Since the M5P model obtained the lowest values in the error metrics and the highest values in the correlation and R^2 , showing to be an efficient model, in addition to providing the necessary equations for the optimization process.

KEYWORDS: Machine Learning. Optimization. Decision making. Decision tree. M5P algorithm.

LISTA DE ILUSTRAÇÕES

Figura 1	Classificação desta Pesquisa Científica	18
Figura 2	Fluxograma da pesquisa para execução da sistemática proposta	20
Figura 3	Diagrama do fluxo de dados no sistema de tomada de decisão	21
Figura 4	Número de publicações utilizando as palavras-chaves " <i>Machine Learning</i> "AND " <i>Decision making</i> " e " <i>Machine Learning</i> "AND " <i>Optimization</i> "	24
Figura 5	Número de publicações utilizando as palavras-chaves " <i>Machine Learning</i> "AND " <i>Decision making</i> "AND " <i>Optimization</i> "	24
Figura 6	Área das publicações com as palavras-chave " <i>Machine Learning</i> "AND " <i>Decision making</i> "AND " <i>Optimization</i> "	25
Figura 7	Países com publicações com as palavras-chave " <i>Machine Learning</i> "AND " <i>Decision making</i> "AND " <i>Optimization</i> "	25
Figura 8	Rede de palavras-chave " <i>Machine Learning</i> "AND " <i>Decision making</i> "AND " <i>Optimization</i> "	26
Figura 9	Mapeamento de clusters com " <i>Decision making</i> " e " <i>Optimization</i> "	27
Figura 10	Número de publicações utilizando as palavras-chaves " <i>Machine Learning</i> "AND " <i>Decision making</i> "AND " <i>Optimization</i> "AND (" <i>Supply chain</i> "OR " <i>Industry</i> ")	28
Figura 11	Área das publicações com as palavras-chave " <i>Machine Learning</i> "AND " <i>Decision making</i> "AND " <i>Optimization</i> "AND (" <i>Supply chain</i> "OR " <i>Industry</i> ")	28
Figura 12	Número de publicações utilizando as palavras-chaves (" <i>M5 model tree</i> "OR " <i>M5p</i> ") AND (" <i>Supply chain</i> "OR " <i>Industry</i> ") e (" <i>M5 model tree</i> "OR " <i>M5p</i> ") AND " <i>Decision making</i> "AND " <i>Optimization</i> "	29
Figura 13	Fluxo lógico do processo de resolução de problemas	31
Figura 14	Processo de tomada de decisão e pesquisa operacional	32
Figura 15	Técnicas de ML	35
Figura 16	Estrutura geral de ML	36
Figura 17	Processo de Aprendizagem Supervisionada	37
Figura 18	Processo de Aprendizagem não Supervisionada	38
Figura 19	Processo de Aprendizagem Semi-Supervisionada	38
Figura 20	Processo de Aprendizagem por Reforço	39
Figura 21	Processo de Aprendizagem Profunda	40
Figura 22	Conjunto de vectores de dados de entrada	41
Figura 23	Modelo de Caixa Branca e Caixa Preta	43
Figura 24	Regressão linear em duas e três dimensões	44
Figura 25	Subconjuntos de dados em Árvore de Decisão	46
Figura 26	Estrutura de Árvore de Decisão	46
Figura 27	Estrutura de um Modelo baseado em Regras	47
Figura 28	Estrutura de Árvore de Modelo M5	50

Figura 29	Ajuste suave perda de margem para uma SVM linear	53
Figura 30	Modelo de uma rede reural artificial e o neurônio biológico	54
Figura 31	Diagrama do processo estudado	58
Figura 32	Valores ausentes no conjunto de dados	60
Figura 33	Valores ausentes nas variáveis x_1 e x_3	61
Figura 34	Gráfico de correlações entre variáveis	64
Figura 35	Diagrama de dispersão dos dados reais e os dados previstos do modelo M5P . .	67
Figura 36	Diagrama de dispersão dos dados reais e os dados previstos do modelo LR . . .	68
Figura 37	Diagrama de dispersão dos dados reais e os dados previstos do modelo CART .	68
Figura 38	Diagrama de dispersão dos dados reais e os dados previstos do modelo SVM . .	69
Figura 39	Diagrama de dispersão dos dados reais e os dados previstos do modelo NN . . .	70
Figura 40	Valores das métricas de avaliação dos modelos testados	72
Figura 41	Gráfico de linhas de dados reais e previstos com M5P	72
Figura 42	Árvore gerada com o algoritmo M5P	74
Figura 43	Diagrama do fluxo de dados no sistema de tomada de decisão com M5P	80
Figura 44	Resultado na interface de usuário	81
Figura 45	Interface de aplicação feito em Shiny para tomada de decisão - Tela 1	117
Figura 46	Interface de aplicação feito em Shiny para tomada de decisão - Tela 2	118
Figura 47	Interface de aplicação feito em Shiny para tomada de decisão - Tela 3	119

LISTA DE TABELAS

Tabela 1 – Publicações em Scopus	23
Tabela 2 – Métricas de avaliação para modelos de regressão	56
Tabela 3 – Variáveis <i>Inputs</i> e <i>Output</i> do conjunto de dados	59
Tabela 4 – Alguns dados do conjunto de dados do estudo	61
Tabela 5 – Análise dos embarques por linha de produção	62
Tabela 6 – Análise estatística das variáveis do conjunto de dados	65
Tabela 7 – Resumo estatístico do modelo M5P	66
Tabela 8 – Resumo estatístico do modelo LR	67
Tabela 9 – Resumo estatístico do modelo NN	69
Tabela 10 – Desempenho das diferentes divisões do conjunto de dados	70
Tabela 11 – Estatísticas básicas das previsões dos modelos testados	71
Tabela 12 – Métricas de avaliação dos modelos testados	71
Tabela 13 – Dados para restrições	76
Tabela 14 – Detalhe de alguns itens enviados	77
Tabela 15 – Custos obtidos com o aplicativo	81
Tabela 16 – Comparação de custos	82
Tabela 17 – Valores da análise de correlação das variáveis	95

LISTA DE ABREVIATURAS E SIGLAS

AI	Inteligencia Artificial
BD	Big Data
CNN	Rede Neural Convolutacional
DL	Aprendizagem Profunda
DQN	Deep Q Network
DT	Árvores de decisão
FfNN	Rede Neural Feedforward
GANs	Redes adversarias geradoras
ICT	Tecnologias de Informação e Comunicação
K-m	k-means
LapSVMs	Máquinas de vetores de suporte laplaciano
LOF	Fator outlier local
LR	Regressão linear
ML	Aprendizagem de máquina
NN	Redes neurais
PPO	Otimização de Política Proximal
RBS	Sistema Baseado em Regras
RL	Aprendizagem por Reforço
RNN	Rede Neural Recorrente
SDR	Standard Deviation Reduction
SL	Aprendizagem Supervisionado
SSL	Aprendizagem Semi-Supervisionada
SVM	Máquinas de vetores de suporte
TSVM	Máquina de vetores de suporte transdutor
UNESP	Universidade Estadual Paulista
USL	Aprendizagem não Supervisionada

SUMÁRIO

1	INTRODUÇÃO	16
1.1	CONTEXTUALIZAÇÃO DA PESQUISA	16
1.2	QUESTÕES DA PESQUISA E OBJETIVOS	17
1.2.1	Questões da pesquisa:	17
1.2.2	Objetivo geral:	18
1.2.3	Objetivos específicos:	18
1.3	PROCEDIMENTO METODOLÓGICO E MATERIAIS	18
1.4	JUSTIFICATIVA E CONTRIBUIÇÃO CIENTÍFICA	22
1.5	DELIMITAÇÃO DA PESQUISA	30
1.6	ESTRUTURA DO TRABALHO	30
2	FUNDAMENTAÇÃO TEÓRICA	31
2.1	TOMADA DE DECISÃO	31
2.2	OTIMIZAÇÃO E TOMADA DE DECISÃO	32
2.3	APRENDIZAGEM DE MÁQUINA	34
2.3.1	Processo de uma estrutura geral de ML	35
2.3.2	Classificação das técnicas de ML	36
2.3.3	Descrição de terminologias em ML	40
2.4	APRENDIZAGEM SUPERVISIONADA VOLTADA PARA OBTENÇÃO DE FUNÇÕES DE REGRESSÃO COM MODELOS DE CAIXA BRANCA E CAIXA PRETA	42
2.4.1	Modelos de Caixa Branca	43
2.4.2	Modelos de Caixa Preta	52
2.4.3	Métricas de avaliação para modelos de regressão	55
3	DESCRIÇÃO E MODELAGEM DO PROBLEMA	57
3.1	DESCRIÇÃO E IDENTIFICAÇÃO DO PROBLEMA	57
3.2	BANCO DE DADOS	60
3.3	MODELAGEM DO PROBLEMA E ANÁLISE DOS RESULTADOS OBTIDOS	66
4	CONCLUSÕES E RECOMENDAÇÕES PARA FUTURAS PESQUISAS	83
4.1	Proposta para Futuras Pesquisas	85
	REFERÊNCIAS	86
	APÊNDICE A – VALORES DA ANÁLISE DE CORRELAÇÃO DAS VARIÁVEIS	95
	APÊNDICE B – ÁRVORE DE MODELO M5P	98

APÊNDICE C – EQUAÇÕES DE REGRESSÃO DA ÁRVORE DE MODELO M5P	101
APÊNDICE D – APLICATIVO FEITO EM SHINY PARA TOMADA DE DECISÃO - TELA 1	117
APÊNDICE E – APLICATIVO FEITO EM SHINY PARA TOMADA DE DECISÃO - TELA 2	118
APÊNDICE F – APLICATIVO FEITO EM SHINY PARA TOMADA DE DECISÃO - TELA 3	119

1 INTRODUÇÃO

Neste capítulo são apresentadas a contextualização da pesquisa, as questões da pesquisa, os objetivos do estudo, os materiais e métodos utilizados, a justificativa e contribuição científica, a delimitação da pesquisa e, por fim, a estrutura desta dissertação.

1.1 CONTEXTUALIZAÇÃO DA PESQUISA

A indústria está atualmente no meio de uma revolução baseada em dados. Com a integração das Tecnologias de Informação e Comunicação (ICT, *Information and Communication Technologies*) e seu rápido avanço tecnológico nos diferentes processos da cadeia de suprimento, aumentou rapidamente a quantidade de dados produzidos nas últimas décadas (GAMARRA; GUERRERO; MONTERO, 2016; MOKTADIR et al., 2019). Devido a esses avanços, há muitas oportunidades para desenvolver ferramentas e analisar esses grandes volumes de dados ou *Big Data* (BD). O termo de BD, tradicionalmente, têm sido definido por sua velocidade, volume, variedade, veracidade e complexidade em capturar, armazenar, analisar, compartilhar e vincular uma grande quantidade de dados criados por meio de diferentes tecnologias para uma melhor compreensão e tomada de decisão (CHEN; ZHANG, 2014; GANDOMI; HAIDER, 2015; ZHONG et al., 2016; FAVARETTO; CLERCQ; BERNICE, 2019).

A capacidade de extrair resultados significativos de os dados por meio da aplicação de estatística, simulação, otimização, econometria, matemática ou outras técnicas para tomada de decisão (MANDAL, 2019), têm um impacto positivo em toda uma organização (ACHARYA et al., 2018). Neste sentido, o foco das grandes tecnologias de dados em ambientes industriais é uma área de pesquisa interdisciplinar relativamente nova (O'DONOVAN et al., 2015), que tornou-se uma propriedade urgente e essencial para melhorar a eficiência dos processos, resultados de negócios e obter uma vantagem competitiva (AKTER et al., 2019).

Dada a quantidade de dados armazenados das organizações e as ferramentas tecnológicas emergentes nos últimos anos, a empresa de consultoria e pesquisa de tecnologia Gartner, em sua publicação "As 10 principais tendências tecnológicas para 2017" detalha que a Inteligência Artificial (AI, *Artificial Intelligence*) e Aprendizagem de Máquina (ML, *Machine Learning*) usam dados massivos para modelar e criar sistemas inteligentes, além disto detalha que ML atingiu um ponto crítico decisivo e terá um aumento nos próximos anos (GARTNER INC., 2016). Além disso, em sua publicação de tendências para o 2018, descreve que as aplicações com análise inteligente têm um alto impacto na indústria, além de que estas aplicações são uma maneira de aumentar a atividade humana na análise avançada da informação (aplicando comumente IA com ML) e não como uma forma de substituir pessoas (GARTNER INC., 2017).

Assim, o termo de ML refere-se à detecção automatizada de padrões significativos em dados, com o objetivo de construir modelos preditivos (ZHOU et al., 2017), concentrando-se na construção e estudo de algoritmos que podem aprender com os dados (GAMARRA; GUERRERO; MONTERO, 2016). Em um sentido amplo, o propósito do reconhecimento de padrões e da teoria da aprendizagem é compreender os mecanismos que regulam o desenvolvimento da inteligência, entendido como a

capacidade de extrair o conhecimento da experiência passada para aplicá-lo no futuro (VERCELLIS, 2009).

De acordo com Al-Jarrah et al. (2015), na última década, os sistemas de aprendizagem automática têm sido amplamente adotados em uma série de campos complexos de uso intensivo de dados, como climatologia, finanças, indústria, astronomia, medicina, biologia, economia, etc. Destarte, ML têm desempenhado um papel muito importante na construção de modelos com dados históricos de processos industriais. Conforme com Ge et al. (2017), as diferentes aplicações de ML na cadeia de suprimentos alterara fundamentalmente a gestão dos processos industriais, dado que, podem ser obtidas informações úteis, podem ser feitas previsões e as decisões podem ser tomadas de forma mais rápida e eficaz.

Neste contexto, a implementação de modelos preditivos focados na tomada de decisão têm um grande potencial para ajudar as empresas em seus diferentes processos. Os modelos podem ser desenvolvidos por meio de diferentes técnicas de ML e, essas técnicas podem ser classificadas como aprendizagem supervisionada (SL, *Supervised learning*), não supervisionada, semi supervisionada, por reforço e como aprendizagem profunda. As técnicas de SL, ao contrário de outras classificações, requer um conjunto de dados rotulados, ou seja, dados de entrada e dados de saída (ou resultados desejados). Outrossim, as técnicas de SL visam deduzir uma função com base nos dados rotulados de treinamento, e, podem ser usados para prever novas saídas, tanto saídas categóricas ou contínuas saídas (FLATH; STEIN, 2018).

Por outro lado, em conformidade com Kolbin (2003) o processo de tomada de decisão desempenha um papel muito importante na gestão e controle das organizações, e com o tempo tornou-se mais crucial. Além disso, pode-se afirmar que a tomada de decisão está intimamente relacionada às abordagens de pesquisa operacional, pois envolve a escolha entre várias alternativas possíveis, portanto está relacionada com a otimização. Assim, as técnicas de otimização fornecem ferramentas para modelar e tomar decisões ótimas, sendo que os modelos podem ser bastante simples com um pequeno número de variáveis controláveis ou podem ser de grande escala e bastante complexos, com milhares de variáveis e muitas condições (MURTY, 2003).

Concluindo esta seção, o objetivo geral deste estudo é desenvolver uma sistemática para tomada de decisões por meio de técnicas de SL; a fim de gerar funções empíricas, com o intuito de incorporá-las no processo de otimização. Assim, a sistemática proposta é validada estatisticamente por meio de uma aplicação real no processo de transporte da área de logística de uma empresa automotiva multinacional brasileira do segmento de autopeças; visando obter o melhor custo, na seleção da empresa de transporte terceirizada.

1.2 QUESTÕES DA PESQUISA E OBJETIVOS

Com base no exposto na seção anterior, as seguintes questões e objetivos são definidos.

1.2.1 Questões da pesquisa:

- Como implementar um modelo de previsão para a tomada de decisão utilizando técnicas de SL?
- Que técnica de SL utilizar para desenvolver funções empíricas de previsão?

- Como modelar o processo de otimização considerando o modelo de previsão?
- Como validar os resultados?

1.2.2 Objetivo geral:

Desenvolver uma sistemática para tomada de decisões por meio da implementação de um modelo de previsão utilizando técnicas de SL.

1.2.3 Objetivos específicos:

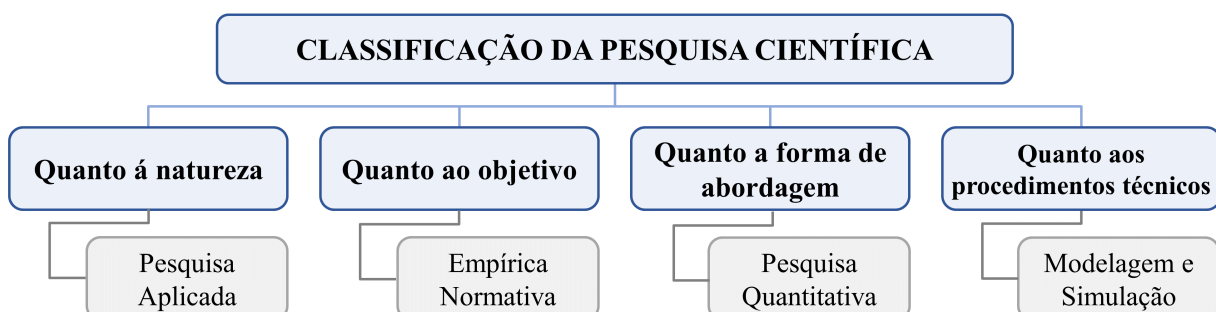
- Usar o modelo preditivo de SL.
- Selecionar a melhor técnica de SL para o desenvolvimento do modelo de previsão.
- Modelar o processo de otimização considerando as equações do modelo de previsão.
- Propor uma ferramenta computacional para a tomada de decisão.
- Analisar o impacto do uso do modelo de tomada de decisão no processo logístico para obter o melhor custo de transporte na empresa.
- Validar a abordagem proposta.

1.3 PROCEDIMENTO METODOLÓGICO E MATERIAIS

Existem diferentes abordagens metodológicas utilizadas no desenvolvimento dos trabalhos científicos. É importante identificar a melhor abordagem para estabelecer as consultas de pesquisa, bem como seus respectivos métodos e técnicas de planejamento e execução, resultando em trabalhos melhor estruturados (MIGUEL, 2007).

De acordo com Filippini (1997), Bertrand e Fransoo (2002), Cantamessa (2003), Bell e Bryman (2007) e Choi, Cheng e Zhao (2016) este trabalho pode ser classificado como uma pesquisa aplicada, com objetivos empíricos normativos, pois a sistemática proposta visa aprimorar a situação atual na tomada de decisões. A forma de abordar o problema é quantitativa, sendo o método de pesquisa a Modelagem e Simulação, conforme exibido na Figura 1.

Figura 1 – Classificação desta Pesquisa Científica



Fonte: Adaptação de (BERTRAND; FRANSOO, 2002) e (CANTAMESSA, 2003).

Entre os materiais utilizados para o desenvolvimento do projeto de pesquisa temos o seguinte:

- Banco de Dados: Os dados utilizados para o desenvolvimento desta pesquisa, são de uma empresa brasileira multinacional no ramo automotivo, com 31 unidades em 14 países e cerca de 15 mil funcionários; fabricante de componentes automotivos (cargas que serão transportadas) de veículos pequenos e pesados como chassis, longarinas, rodas de aço, etc. Os dados estudados vêm da área de logística, especificamente do processo de transporte.
- *Software R*: *software* livre R, é uma linguagem e ambiente para computação estatística e gráficos. Fornece uma ampla variedade de técnicas estatísticas, e é altamente extensível (R-PROJECT, 2018). Utilizou-se para treinar o modelo de previsão e modelar as restrições do processo, além de usar o pacote *shiny* para a criação de uma interface de usuário.

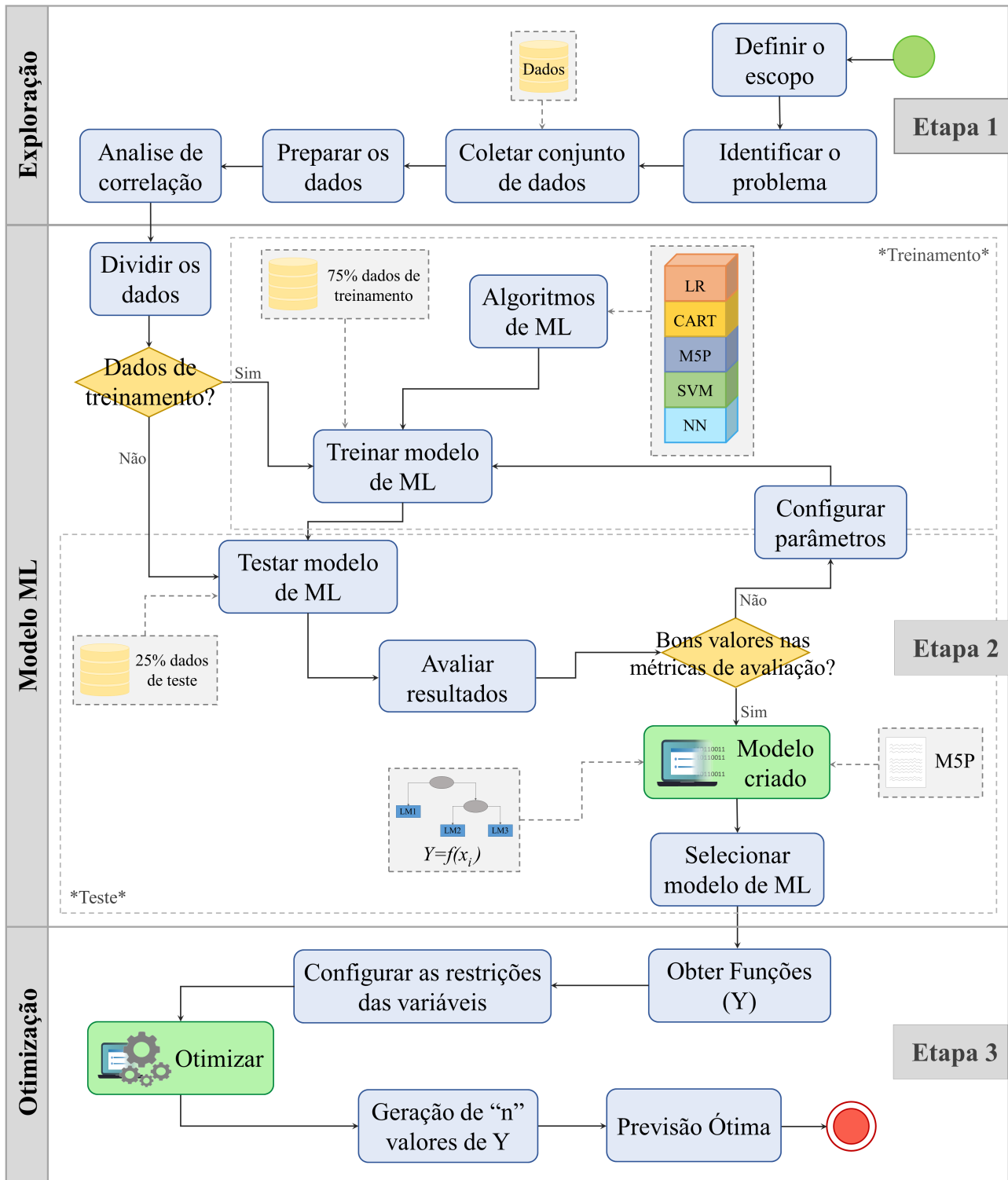
A Figura 2 apresenta a sistemática metodológica proposta. O fluxograma ou diagrama de processos apresentado é dividido em três etapas e, cada etapa é dividida em um conjunto de tarefas estipuladas:

- Etapa 1 - Exploração: Concerne a parte inicial do processo, com a definição do escopo e a identificação do problema, posto que, é importante compreender o contexto do processo que será estudado, identificando as variáveis independentes que influenciam no processo (variáveis com restrições e sem restrições), além de compreender a variável resposta. Em seguida, o conjunto de dados é coletado, com base na análise do problema, para subsequentemente preparar o conjunto de dados por meio de técnicas estatísticas, como a análise de correlação das variáveis.
- Etapa 2 - Modelo de aprendizagem de máquina: Trata-se da criação do modelo de previsão. Depois de preparar os dados coletados, são divididos em dados para treinamento e dados para teste ou validação. Regularmente a divisão do conjunto de dados é de 75% para treinamento e 25% restante para testes. O conjunto de dados de teste é usado para avaliar o modelo treinado e deve ser o suficientemente grande e representativo para gerar resultados significativos.

Dado que o SL têm o objetivo de inferir uma função com base nos dados de treinamento, são testadas as técnicas de aprendizagem supervisionada de regressão linear, árvore de decisão CART, árvore de modelo M5P, máquina de vetor de suporte e redes neurais. O modelo mais eficiente é selecionado, após a geração de um ciclo no qual se pretende obter um modelo com os melhores valores nas diferentes métricas de avaliação.

- Etapa 3 - Otimização: As funções de previsão de o modelo de ML gerado na etapa 2 são utilizadas, a fim de gerar sub-rotinas internas e atender às restrições modeladas do problema estudado. Assim, todo esse processo visa gerar um ciclo iterativo com os possíveis valores das variáveis e obter os melhores valores para a tomada de decisão.

Figura 2 – Fluxograma da pesquisa para execução da sistemática proposta



Fonte: Próprio autor (2020)

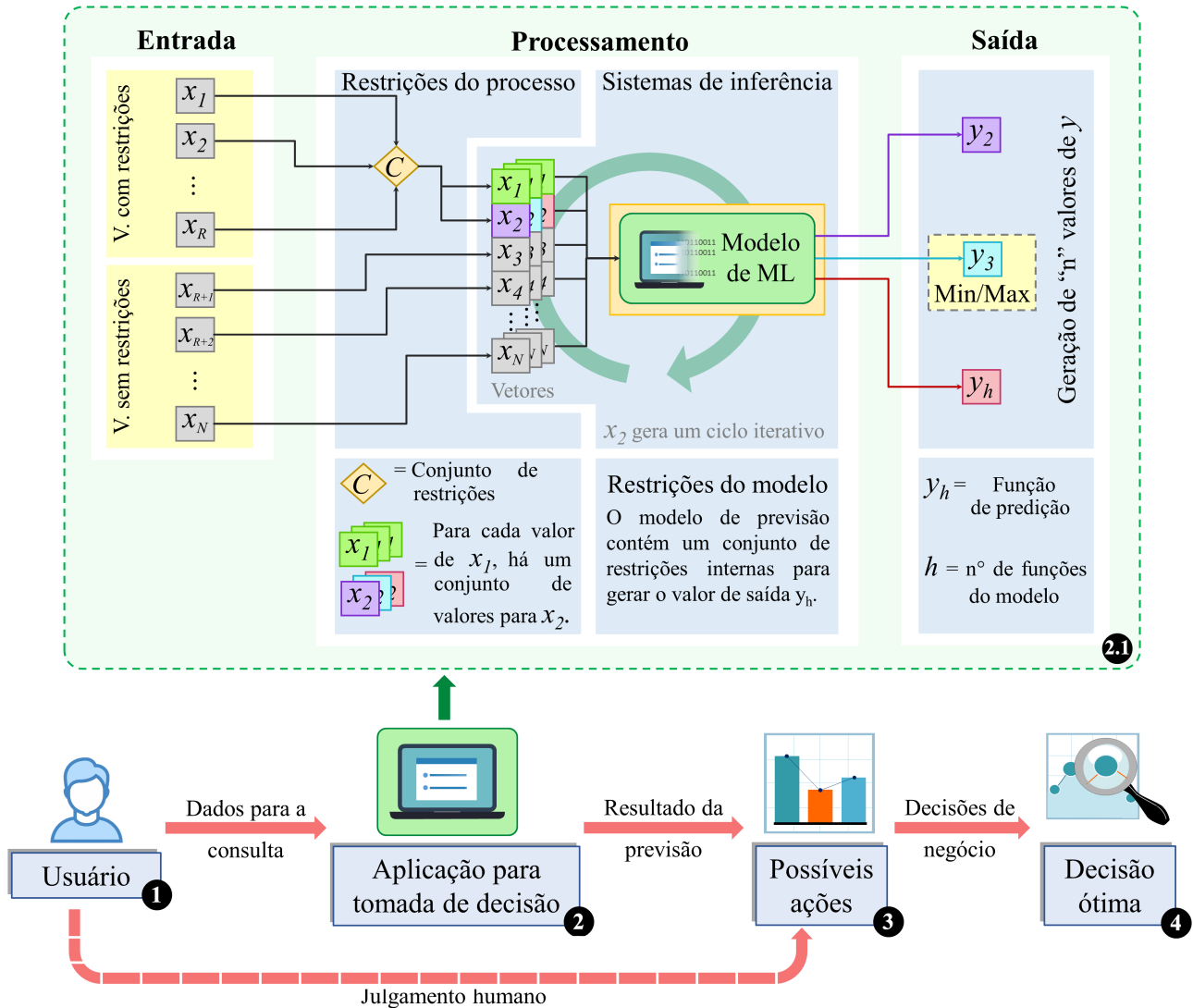
Por outro lado, a Figura 3 ilustra o fluxo de dados da ferramenta computacional proposta, o desenvolvimento da interface do usuário visa encontrar os valores ótimos para as variáveis de decisão, a fim de encontrar a melhor solução que maximize lucros ou benefícios ou minimize perdas ou custos. Esse processamento de dados é dividido em quatro fases:

- Fase 1: Consiste na parte inicial, o usuário insere no sistema os dados necessários para realizar a

consulta para tomada de decisão.

- Fase 2: Refere-se à fase de processamento dos dados inseridos para a consulta. Deste jeito, a maneira pela qual o sistema de tomada de decisão interage com o usuário e processa os dados inseridos pode ser agrupada nas atividades básicas de entrada, processamento e saída.

Figura 3 – Diagrama do fluxo de dados no sistema de tomada de decisão



Fonte: Próprio autor (2020)

A **entrada** é composta pelos dados inseridos pelo usuário e, no caso um valor não seja inserido, será considerado zero nessa variável. Assim, as variáveis são classificadas como conjunto de dados com restrições ou variáveis de decisão, de x_1 a x_R e, como conjunto de dados sem restrições, de x_{R+1} a x_N , sendo N o número total de variáveis consideradas para a criação do modelo preditivo.

No **processamento**, com base nas restrições do processo de uma ou mais variáveis, é gerado um valor ou conjunto de valores que serão armazenados em sua variável correspondente. No caso de uma variável assumir mais de um valor (o valor ótimo das variáveis de decisão está

entre um conjunto de valores), é gerado um conjunto de vetores com os valores das variáveis de entrada e a predição é calculada por um ciclo iterativo a partir do modelo de ML, conforme apresentado na Figura 3. Deste modo, por exemplo, na figura é possível observar um conjunto de três vetores, onde a variável x_1 é sombreada em verde, referindo-se ao mesmo valor sendo repetido nos três vetores. Por outro lado, a variável x_2 têm uma cor diferente em cada vetor, o que significa que o valor ótimo para x_2 está entre um grupo de valores, o que gerará um ciclo iterativo no cálculo da previsão da variável de saída y (cada valor de previsão y é gerado por uma função diferente do modelo ML).

Na **saída**, como um ciclo iterativo é gerado internamente, em cada iteração, o sistema de inferência ML gera um conjunto de valores y_h , que é calculado pela função de previsão do modelo de ML e, h é o número de funções de previsão do modelo. Posteriormente, é selecionado o melhor valor que atenda às restrições e o objetivo inicial, maximização ou minimização.

- Fase 3: Os resultados da previsão são apresentados, mostrando as possíveis ações que foram calculadas no processo interno do sistema, além de exibir o valor ótimo mínimo ou máximo calculado pelo sistema.
- Fase 4: O uso de sistemas inteligentes não substitui as pessoas, pelo contrário, é uma ferramenta válida para análise avançada de informações. Os valores apresentados na previsão são analisados e, com o julgamento do usuário, é tomada a melhor decisão.

1.4 JUSTIFICATIVA E CONTRIBUIÇÃO CIENTÍFICA

Nos últimos anos, uma enorme quantidade de dados foram coletados e armazenados pelos desenvolvimentos tecnológicos recentes, entre essas tecnologias está a Internet das coisas, Sistema cibernético, tecnologia da informação e comunicação, computação em nuvem, big data, comunicações sem fio e sistemas de aquisição de dados, que estão intimamente relacionados à Indústria 4.0 (LU, 2017; YAO; GE, 2018). O termo «Industrie 4.0» ou indústria inteligente, foi estabelecido para a próxima revolução industrial (HENNING; WOLF-DIETER; WOLFGANG, 2011), esta revolução foi precedida por outras três revoluções industriais na história da humanidade; a primeira revolução industrial foi no campo da mecanização, o uso intensivo de energia elétrica foi a chamada segunda revolução e a terceira revolução industrial foi a revolução digital (HERMANN; PENTEK; OTTO, 2016). Indústria 4.0 foi origem da iniciativa estratégica Alemã e está agora na agenda de vários países europeus, nos EUA e na Ásia (STRANDHAGEN et al., 2017).

Dado que, a Indústria 4.0 gera uma quantidade inumerável de dados não estruturados provenientes de diferentes domínios (YUE et al., 2015), os dados podem ser consumidos por diferentes técnicas, como a análise de dados ou modelagem para permitir que as indústrias desenvolvam uma melhor compreensão de seus processos e atividades para obter informações que possam melhorar as operações existentes (O'DONOVAN et al., 2015). Neste sentido, a análise de dados pode ser definida como a gestão baseada em fatos para direcionar decisões e ações por meio de análises estatísticas, modelos explicativos e preditivos (HAZEN et al., 2018).

Na indústria, os chefes das diferentes áreas da cadeia de suprimentos dependem cada vez mais de dados gerados a partir dos diferentes processos para obter informações para reduzir custos, melhorar o desempenho, visualizar despesas, monitorar estoques, etc. Desse jeito, com o uso dos dados, os gerentes pretendem melhorar a compreensão, otimizar e melhorar processos para obter uma vantagem competitiva (HAZEN et al., 2014). No entanto, aproveitar os dados por meio das diferentes ferramentas de análise (estatísticas, simulação, otimização, AI entre ML e data mining), oferecem grandes oportunidades para melhorar a tomada de decisão e aumentar a eficiência e a eficácia dos processos industriais da cadeia de suprimento.

Na literatura, várias técnicas de ML têm sido aplicadas a uma série de problemas na cadeia de suprimentos para tomada de decisão, podendo-se observar quão eficiente pode ser esta ferramenta. Sendo assim, ML têm desempenhado um papel muito importante na construção de modelos preditivos (GE et al., 2017), dado que, o processamento de dados é transparente e seus dados previstos são suficientemente precisos para suportar um problema de tomada de decisão (RUDIN et al., 2012).

Por outro lado, acrescentando que para destacar a contribuição desta pesquisa, foi realizado um estudo bibliométrico descritivo-retrospectivo dos artigos publicados em um período desde a primeira publicação em 1962 até 08 de março de 2020, com as palavras-chaves de "*Decision making*", "*Optimization*", *Machine Learning*, *Supervised learning*, "*Supply chain*", "*Industry*", "*M5 model tree*" e "*M5P*" na base de dados de Scopus. A Tabela 1 apresenta os resultados do estudo realizado em relação ao número de publicações e ao total de citações.

Tabela 1 – Publicações em Scopus

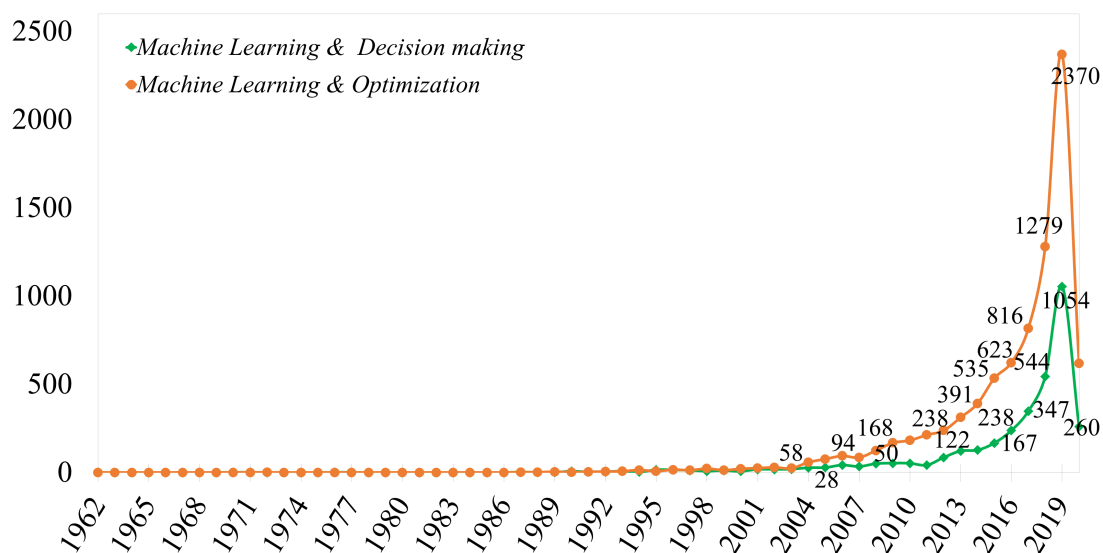
Palavras-chaves	Total	Citações
<i>"Decision making"AND "Optimization"</i>	22 323	348 493
<i>"Machine Learning"AND "Decision making"</i>	3 429	54 986
<i>"Machine Learning"AND "Optimization"</i>	8 394	184 970
<i>"Decision making"AND "Optimization"AND ("Supply chain"OR "Industry")</i>	2 787	51 380
<i>"Machine Learning"AND "Decision making"AND "Optimization"</i>	363	4 329
<i>"Machine Learning"AND "Decision making"AND "Optimization"AND ("Supply chain"OR "Industry")</i>	25	246
<i>"Supervised learning"AND "Decision making"AND "Optimization"</i>	26	1 316
<i>"M5 model tree"OR "M5P"</i>	307	4 886
<i>("M5 model tree"OR "M5P") AND "Optimization"</i>	38	440
<i>("M5 model tree"OR "M5P") AND "Decision making"</i>	13	380
<i>("M5 model tree"OR "M5P") AND ("Supply chain"OR "Industry")</i>	8	108
<i>("M5 model tree"OR "M5P") AND "Decision making"AND "Optimization"</i>	1	0
<i>("M5 model tree"OR "M5P") AND "Decision making"AND "Optimization"AND ("Supply chain"OR "Industry")</i>	0	0
<i>("M5 model tree"OR "M5P") AND "Decision making"AND "Optimization"AND ("Logistic process"OR "Transport")</i>	0	0

Fonte: Base de dados de Scopus (2020).

Seguindo a análise bibliométrica, alguns resultados das combinações de palavras-chave apresenta-

das na Tabela 1 são detalhados. Sendo assim, utilizando combinações das palavras-chave "*Machine Learning*" AND "*Decision making*" e "*Machine Learning*" AND "*Optimization*", foram identificados 3 429 artigos publicados no período de 1962-2020 e 8 394 artigos entre 1972-2020 respectivamente. Além disso, a Figura 4 apresenta a tendência das publicações, na qual é possível identificar que nos últimos 14 anos houve um interesse constante na realização de estudos e publicação de artigos com tópicos que envolvam as palavras-chave apresentadas.

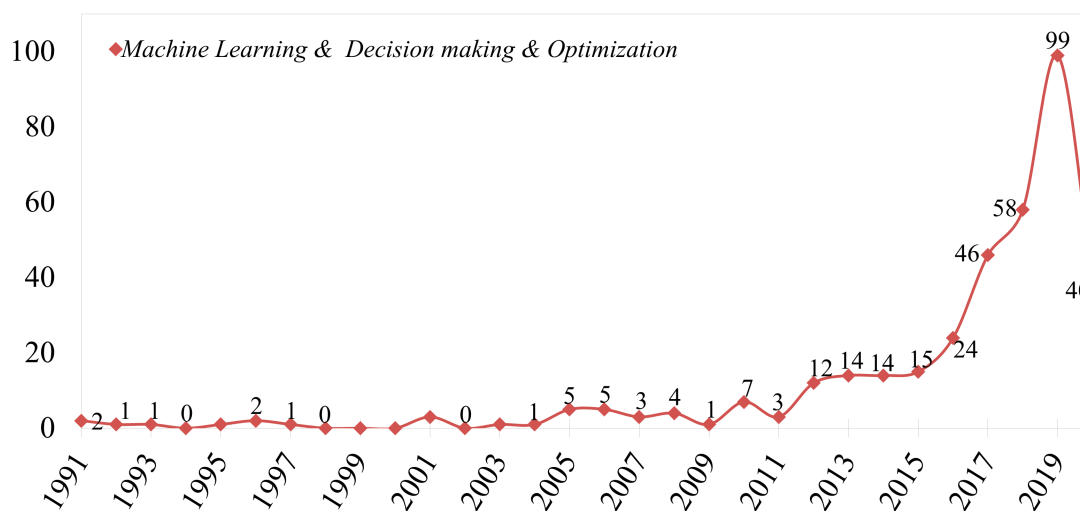
Figura 4 – Número de publicações utilizando as palavras-chaves "*Machine Learning*" AND "*Decision making*" e "*Machine Learning*" AND "*Optimization*"



Fonte: Base de Dados de SCOPUS (2020)

Por outro lado, a Figura 5 ilustra a tendência dos artigos que contêm a combinação das palavras-chave "*Machine Learning*" AND "*Decision making*" AND "*Optimization*" no período de 1991-2020, sendo identificados 363 publicações; também é possível identificar que nos últimos 5 anos houve um aumento no número de artigos, evidenciando a contemporaneidade e relevância do assunto.

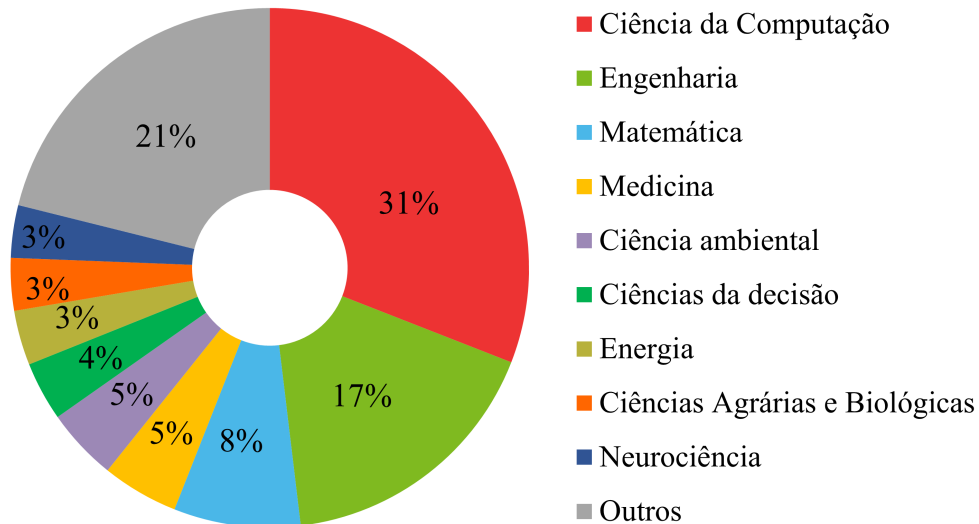
Figura 5 – Número de publicações utilizando as palavras-chaves "*Machine Learning*" AND "*Decision making*" AND "*Optimization*"



Fonte: Base de Dados de SCOPUS (2020)

Conjuntamente, a Figura 6 indica a porcentagem do total de publicações pertencentes a cada área; sendo que dos 363 artigos publicados, 31% pertencem ou têm tópicos fortemente relacionados com a área de ciência da computação e 17% à engenharia.

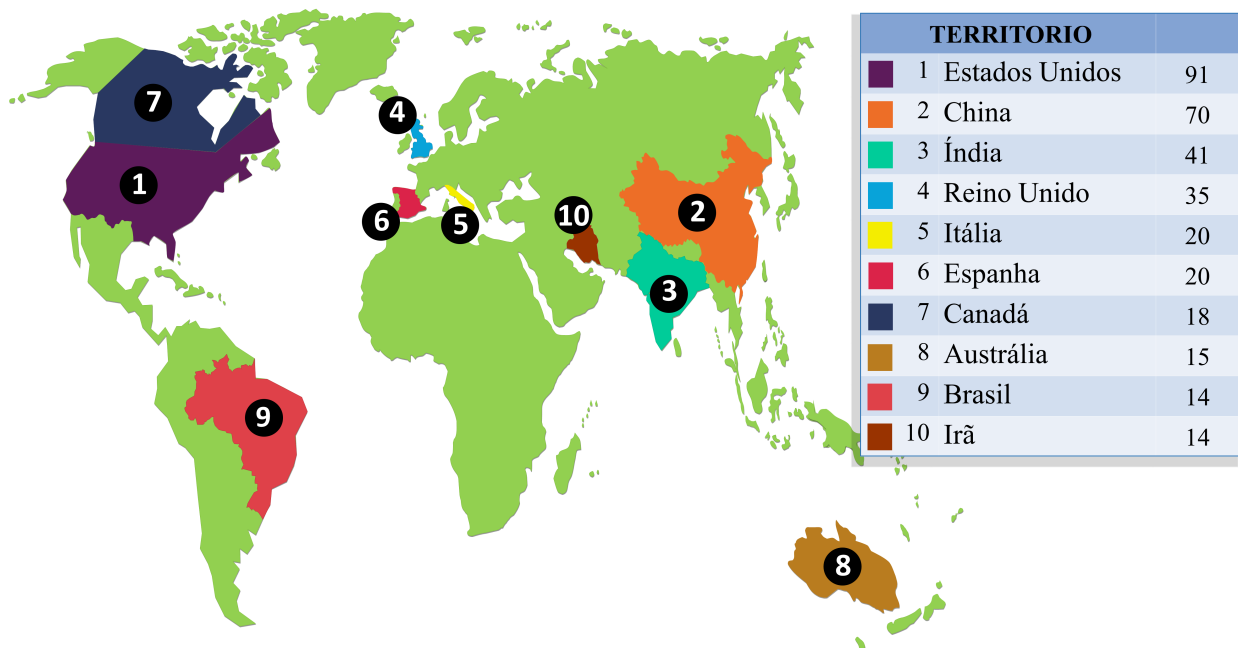
Figura 6 – Área das publicações com as palavras-chave "*Machine Learning*"AND "*Decision making*"AND "*Optimization*"



Fonte: Base de Dados de SCOPUS (2020)

Além do mais, a Figura 7 apresenta o ranking dos dez primeiros países com mais publicações com as palavras-chave "*Machine Learning*"AND "*Decision making*"AND "*Optimization*". Dado que, os dois países que lideram a legenda são Estados Unidos e China, com 91 e 70 artigos publicados respectivamente; no entanto, o Brasil está na 9ª posição com apenas 10 publicações.

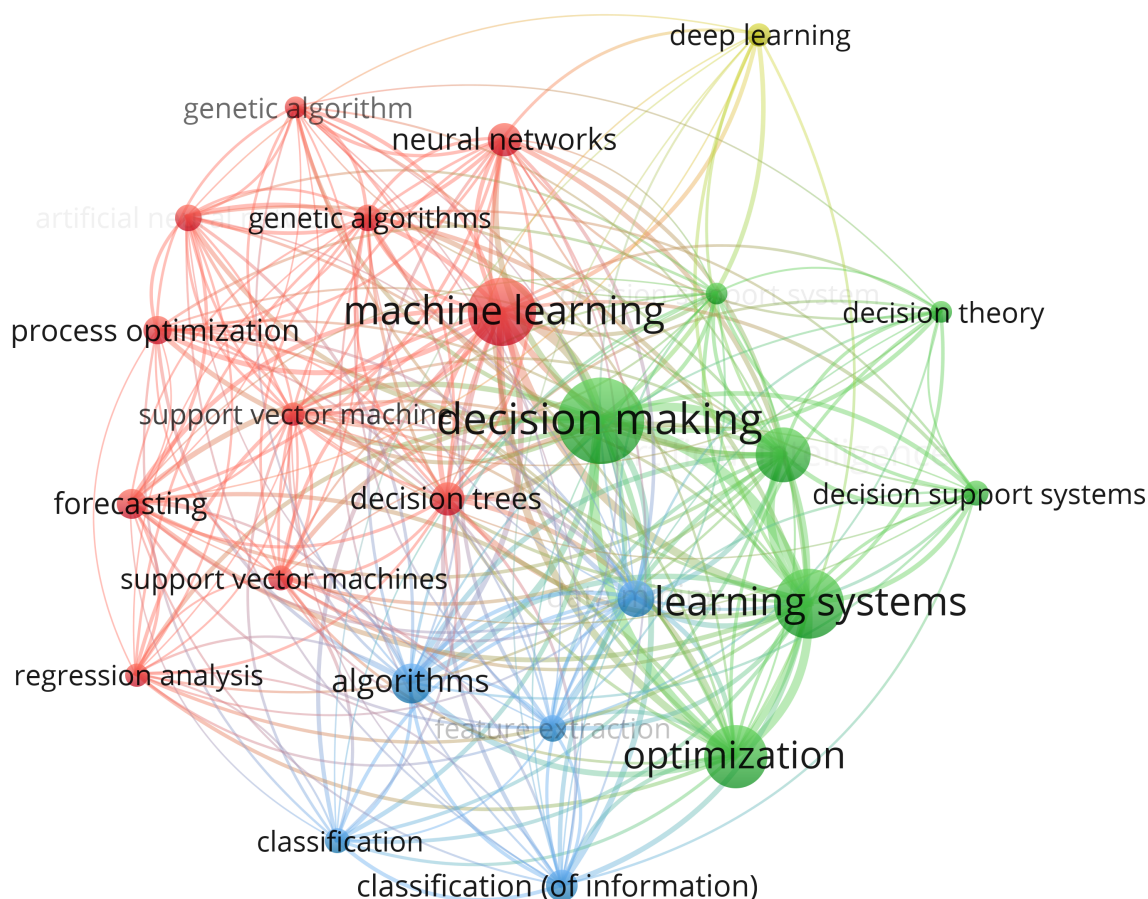
Figura 7 – Países com publicações com as palavras-chave "*Machine Learning*"AND "*Decision making*"AND "*Optimization*"



Fonte: Base de Dados de SCOPUS (2020)

A ferramenta bibliométrica VOSviewer, comumente usado para a construção de redes de relacionamento (ECK; WALTMAN, 2011), é utilizada para a criação do gráfico de co-ocorrência de palavras-chaves combinando os termos de "*Machine Learning*" AND "*Decision making*" AND "*Optimization*"; foram analisados um total de 3 189 palavras com o requisito mínimo de 7 ocorrências, essa filtragem resultou em 4 clusters, composto de 26 termos para o maior cluster e 9 termos para o menor, conforme a Figura 8.

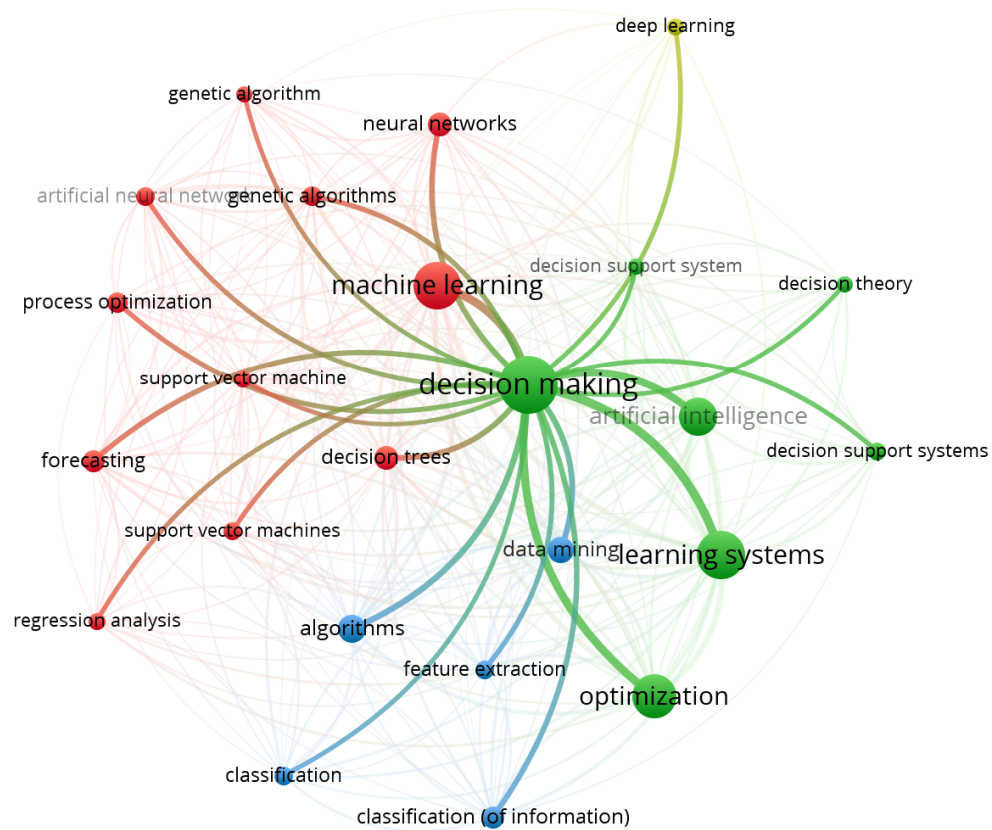
Figura 8 – Rede de palavras-chave "*Machine Learning*" AND "*Decision making*" AND "*Optimization*"



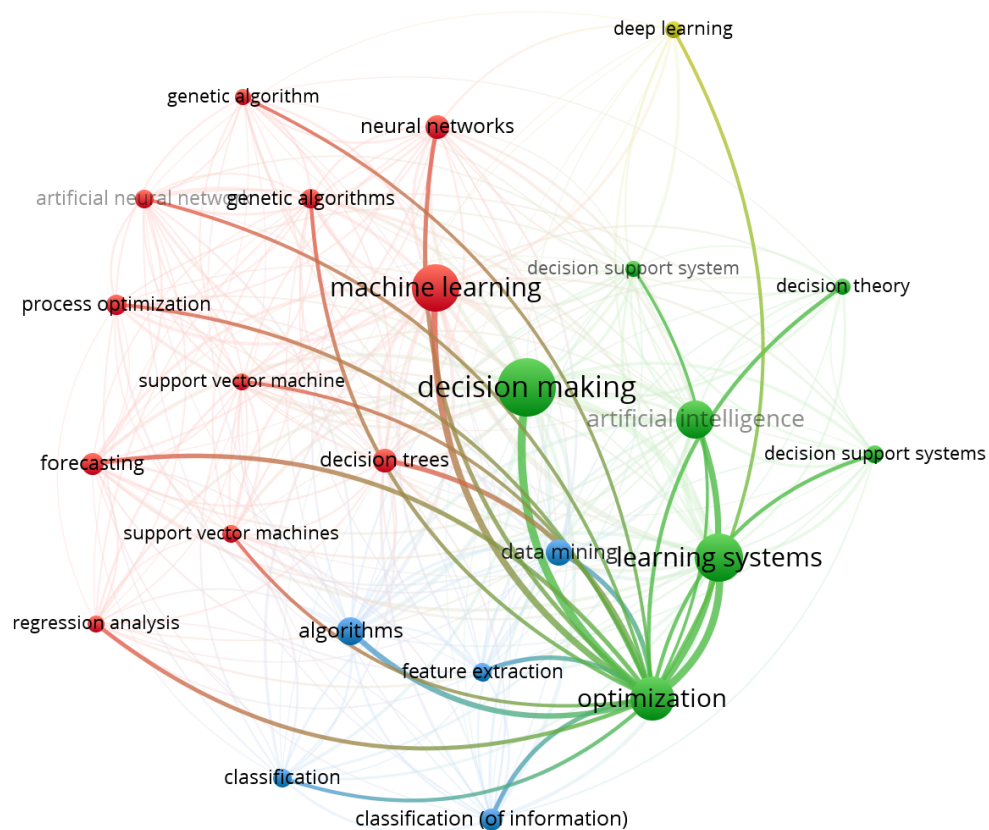
Fonte: VOSviewer e Base de Dados de SCOPUS (2020)

De acordo com as palavras-chaves analisadas que formam a rede, e destacando as conexões das palavras-chave "*Decision making*" e "*Optimization*" (Figura 9), pode-se inferir que existem artigos que utilizam técnicas de ML com otimização focado na tomada de decisão; mostrando assim o interesse científico em combinar essas técnicas. Além disso, as figuras mostram algumas das técnicas de ML utilizadas nas diferentes publicações analisadas; no entanto, pode-se observar que existe uma ligação com as técnicas de máquinas de vetores, redes neurais e árvores de decisão, sendo do tipo de SL. Assim, pode-se destacar de acordo com a figura que as técnicas de árvores de decisão também foram de interesse para o desenvolvimento do modelo de previsão.

Figura 9 – Mapeamento de clusters com "Decision making" e "Optimization"



(a) Clusters com "Decision making"

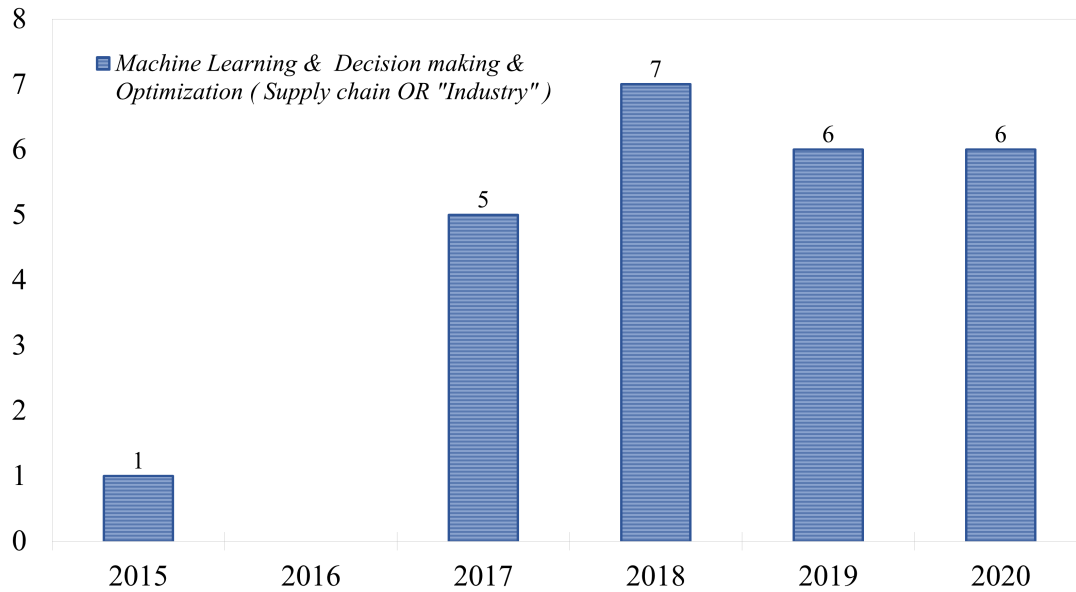


(b) Clusters com "Optimization"

Fonte: VOSviewer e Base de Dados de SCOPUS (2020)

Seguindo com a análise de artigos publicados, a Figura 10 apresenta o resultado da combinação das palavras-chave "*Machine Learning*" AND "*Decision making*" AND "*Optimization*" AND ("*Supply chain*" OR "*Industry*"), sendo identificadas 25 publicações no período de 2015-2020. Adicionalmente, pode destacar-se que existem muito poucas publicações que contenham os tópicos mencionados na área de gestão de processos industriais e tomada de decisão; isso pode ser considerado como uma oportunidade de estudo, tendo em conta que entre 2017-2020 existem apenas 24 artigos publicados.

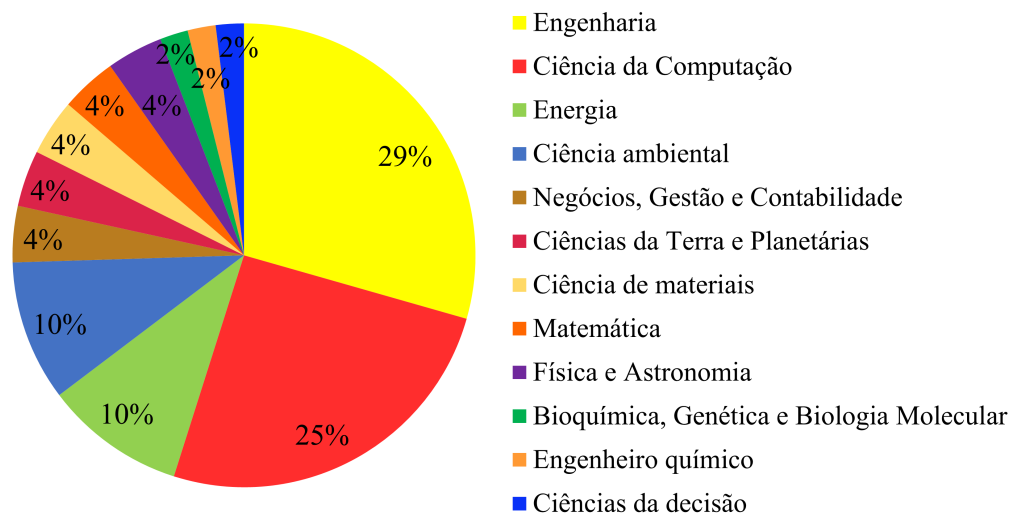
Figura 10 – Número de publicações utilizando as palavras-chaves "*Machine Learning*" AND "*Decision making*" AND "*Optimization*" AND ("*Supply chain*" OR "*Industry*")



Fonte: Base de Dados de SCOPUS (2020)

Assim, os artigos analisados podem ser categorizados de acordo com as áreas apresentadas na Figura 11, sendo que o 29% dos artigos têm tópicos fortemente relacionados com a área de engenharia e o 25% com ciência da computação.

Figura 11 – Área das publicações com as palavras-chave "*Machine Learning*" AND "*Decision making*" AND "*Optimization*" AND ("*Supply chain*" OR "*Industry*")



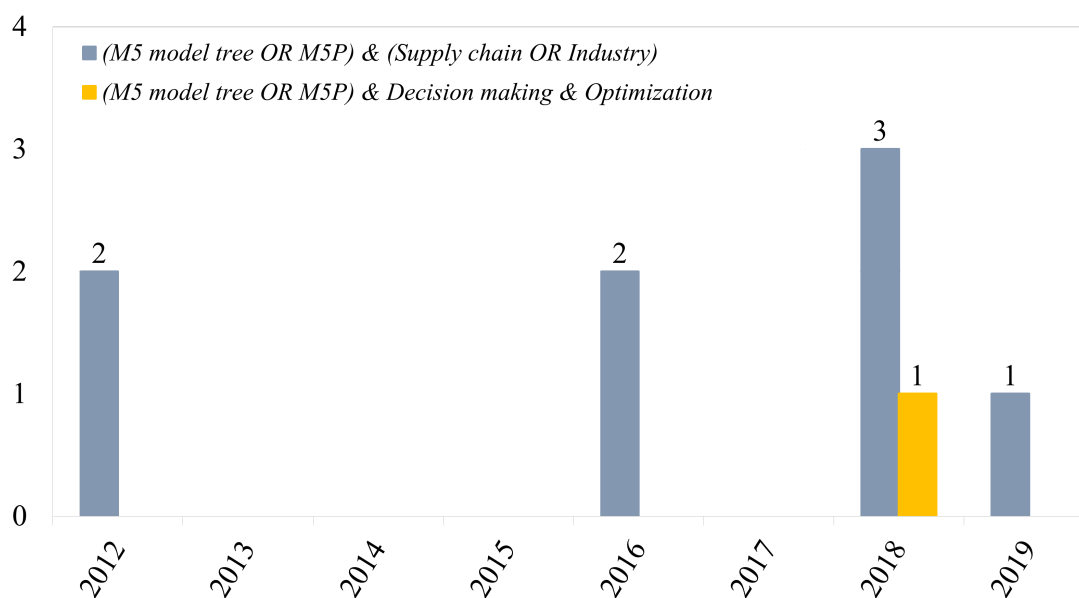
Fonte: Base de Dados de SCOPUS (2020)

Por outro lado, analisando os artigos obtidos na análise bibliométrica, pode-se destacar que a técnica de árvore de decisão é uma técnica de ML muito eficiente e, o algoritmo de árvore de Modelo M5P gera funções empíricas que podem ser usadas no processo de otimização. Tudo isso resultando em uma grande vantagem para ser usado para o desenvolvimento desta pesquisa. Nesse sentido, como mostra a Tabela 1, foi realizado uma análise de artigos publicados com a palavra-chave (*"M5 model tree"OR "M5P"*), resultando em 258 artigos publicados entre 1991-2019, e tendo um crescimento a partir de 2019.

Nessa sequência, a Figura 12 ilustra o número de artigos publicados por ano, com as combinações de palavras-chave de (*"M5 model tree"OR "M5p"*) AND (*"Supply chain"OR "Industry"*) e (*"M5 model tree"OR "M5p"*) AND *"Decision making"AND "Optimization"*, sendo identificados 8 publicações no período 2012-2020 e 1 publicação no 2018 respectivamente. Esses resultados sugere uma oportunidade de pesquisa pelo baixo número de publicações.

Porém, na única publicação de 2018, Bazargan-Lari (2018) propõe determinar um conjunto de alternativas para a descarga de poluentes por meio do NSGA-II, a fim de minimizar a quantidade média de perdas de água durante a descarga, a duração da descarga e o número de tentativas mal sucedidas de descarga; usando o algoritmo M5P, a fim de obter as regras internas para superar o gargalo de otimização que normalmente consome muito tempo. A abordagem proposta é aplicada a uma rede de distribuição de água potável no mundo real.

Figura 12 – Número de publicações utilizando as palavras-chaves (*"M5 model tree"OR "M5p"*) AND (*"Supply chain"OR "Industry"*) e (*"M5 model tree"OR "M5p"*) AND *"Decision making"AND "Optimization"*



Fonte: Base de Dados de SCOPUS (2020)

Finalmente, deve-se ressaltar que ao incluir a palavra-chave *"Supply chain"OR "Industry"* ou *"Logistic process"OR "Transport"*, com as palavras-chaves anteriores, não foram reportados trabalhos publicados. Assim, pode-se destacar a originalidade desta pesquisa, além do fato de que é possível perceber que é um campo pouco explorado, resultando em uma oportunidade de pesquisa para estudos

de tomada de decisão com técnicas de ML, a fim de encontrar valores ótimos em diferentes problemas industriais.

1.5 DELIMITAÇÃO DA PESQUISA

Como delimitações desta pesquisa, serão utilizadas as técnicas de Aprendizagem Supervisionado para o desenvolvimento do modelo de previsão, desta forma, não serão estudadas outras técnicas de ML (WITTEN; FRANK, 2005).

Os estudos aqui relatados serão aplicados à realidade de uma empresa multinacional do segmento de autopeças, e, portanto, os seus resultados não serão generalizados para outras empresas de outros segmentos. Além disso, para as limitações técnicas, os algoritmos de ML utilizados correspondem aos pacotes do *software* livre R.

1.6 ESTRUTURA DO TRABALHO

Este trabalho está estruturado em mais três capítulos além deste capítulo. No Capítulo 2 é apresentado o referencial teórico, contendo temas pertinentes ao trabalho. Já no Capítulo 3 são abordadas a descrição e a modelagem do problema, bem como a apresentação dos resultados e as suas análises. Por fim, no Capítulo 4 são apresentadas as conclusões, verificação dos objetivos e recomendações para futuras pesquisas, seguidas das referências bibliográficas.

2 FUNDAMENTAÇÃO TEÓRICA

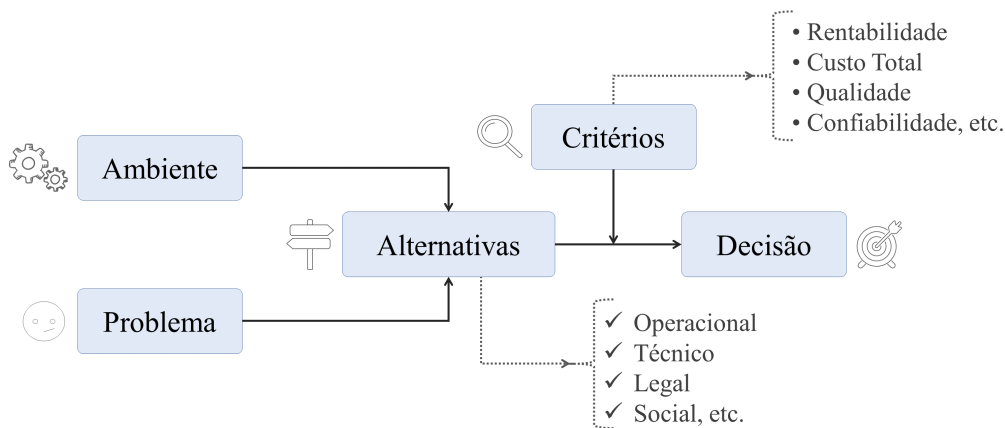
Neste capítulo serão abordados conceitos de tomada de decisão, bem como conceitos de otimização e aprendizagem de máquina com foco principal em algumas das técnicas de caixa branca e caixa preta de aprendizagem de máquina supervisionado; sendo os mesmos conceitos que fundamentam os objetivos e as metodologias utilizadas neste trabalho.

2.1 TOMADA DE DECISÃO

A tomada de decisão em geral pode ser definida como o processo de resolver um problema selecionando a melhor alternativa. Esse processo de escolha pode ser aplicado no ambiente familiar, pessoal, sentimental, político, religioso ou empresarial. Por outro lado, conforme Murty (2003), no ambiente organizacional, independente da posição que uma pessoa possa ocupar, existem problemas de tomada de decisão que surgem na operação de um determinado sistema. Tais problemas podem ser simples, exigindo a determinação de condições simples que devem ser atendidas com um pequeno número de variáveis controláveis, ou, podem ser problemas de grande escala e bastante complexo, composta de muitas variáveis e muitas condições.

Neste contexto, a Figura 13 descreve a estrutura lógica da solução de problemas na tomada de decisão. Assim, dado um problema em um determinado ambiente, as diferentes alternativas representam as possíveis ações para atingir os objetivos, e as melhores alternativas que atendem os critérios de desempenho são selecionadas para uma decisão final (VERCELLIS, 2009).

Figura 13 – Fluxo lógico do processo de resolução de problemas



Fonte: Vercellis (2009)

Segundo Murty (2003), os problemas de decisão podem ser classificados em duas categorias:

- Categoria 1: Esta categoria inclui um pequeno conjunto de alternativas, sendo cada alternativa totalmente conhecida e qualquer um pode ser selecionado como uma decisão.

- Categoria 2: Nesta categoria estão os problemas que possuem alternativas que devem necessariamente satisfazer algumas restrições. Além disso, para identificar todas as alternativas, é necessário construir modelos matemáticos com restrições.

Deste modo, as organizações, no dia a dia, enfrentam diferentes tipos de decisões; por exemplo, decisões com tecnologia, produtos, fabricação, transporte, vendas, clientes, entre outros. Sendo assim, existem várias estratégias para a tomada de decisão, como a otimização, análise de decisão de múltiplos critérios, o processo de hierarquia analítica (AURUM; WOHLIN, 2003); ou também a estatística, os diagramas de influência, tabelas de decisão, processos analíticos de rede, árvores de decisão, entre outros (FRENCH; MAULE; PAPAMICHAIL, 2009).

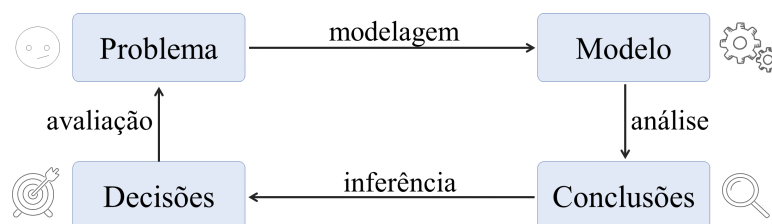
No entanto, devido ao fato de atualmente existir uma grande quantidade de dados armazenados nas organizações, surgiram diferentes tecnologias para modelar e analisar essa quantidade de dados, resultando em uma alternativa vantajosa para a tomada de decisão. Assim, em conformidade com Mortenson, Doherty e Robinson (2015), os dados disponíveis nas organizações e os avanços tecnológicos para análise de dados, permitem o aprimoramento de operações com efeito significativo na gestão, permitindo decisões mais efetivas.

Entre as tecnologias de análise de dados emergentes estão os modelos de previsão, que usam os dados para fornecer informações valiosas sobre possíveis estados futuros (BUTZ; SIGAUD; GERARD, 2003), desta forma gerando um estado de tomada de decisão sobre o que fazer ou o que decidir em relação às previsões. Por outro lado, há uma forte conexão entre a tomada de decisão e a otimização, de modo que um problema de otimização têm a tarefa de selecionar uma decisão a partir de um conjunto de alternativas que satisfaçam um objetivo e um conjunto de restrições (MEISEL, 2011). Não obstante, a otimização também é particularmente associada ao campo de pesquisa operacional.

2.2 OTIMIZAÇÃO E TOMADA DE DECISÃO

Empresas ou organizações enfrentam diferentes problemas de tomada de decisão, que têm a estrutura comum de um conjunto de decisões factíveis e critérios de avaliação estabelecidos para identificar a decisão ótima que gere uma minimização no custo ou maximização no lucro (VERCELLIS, 2009). Todo esse processo pode ser modelado e têm uma forte ligação com a otimização e a pesquisa operacional (RARDIN, 2015). Assim, pode-se afirmar que os modelos de otimização surgem naturalmente dos processos de tomada de decisão. Neste seguimento, a Figura 14 ilustra as etapas de tomada de decisão e pesquisa operacional, com a estrutura de um ciclo que começa com a avaliação e modelagem do problema, seguida da análise dos resultados que leva a uma conclusão e decisão final.

Figura 14 – Processo de tomada de decisão e pesquisa operacional



Fonte: Rardin (2015)

No entanto, segundo Chong (1996), pode-se dizer que um problema de otimização é a maximização ou minimização de uma função objetivo, que envolve a seleção sistemática de um valor que pertence a um conjunto de valores possíveis. Além disso, conforme (MEISEL, 2011) um problema de otimização envolve um vetor de variáveis independentes, também conhecidas como variáveis de decisão $x = [x_1, x_2, x_3 \dots x_n] \in \mathbb{R}^n$, um conjunto de soluções viáveis \mathcal{W} e uma função objetivo $f(x) : \mathcal{W} \rightarrow \mathbb{R}$; sendo o principal objetivo da otimização a determinação de uma solução ótima $w^* \in \mathcal{W}$, conforme a expressão (2.2.1).

$$\forall w \in \mathcal{W} : f(w^*[x_n]) \geq f(w[x_n]) \quad (2.2.1)$$

para maximização, sendo w^* a solução ótima das variáveis de decisão x_n que maximiza a função $f(x)$; e a (2.2.2) para minimização

$$\forall w \in \mathcal{W} : f(w^*[x_n]) \leq f(w[x_n]) \quad (2.2.2)$$

Por outro lado, de acordo com Murty (2003) e Rardin (2015) a forma padrão de um modelo de otimização é composta por uma função objetiva e restrições.

$$\begin{aligned} &\text{min ou max} \quad (\text{função objetivo}) \\ &\text{sujeito a} \\ &\quad (\text{principais restrições}) \\ &\quad (\text{restrições de tipo variável}) \end{aligned}$$

As expressões (2.2.3) - (2.2.8) são outro jeito para o caso de minimização, pode ser representado como:

$$\text{Minimizar} \quad f(x) \quad (2.2.3)$$

$$\text{sujeito a} \quad \sum x_j = b_i \quad , i = 1, 2, \dots, n \quad (2.2.4)$$

$$x_j + x_{j+p} \leq b_i \quad (2.2.5)$$

$$l_j \leq x_j \leq u_j \quad , j = 1, 2, \dots, m \quad (2.2.6)$$

$$x_j \in \Delta_j \quad , j \in J \quad (2.2.7)$$

$$x_j \geq 0 \quad (2.2.8)$$

Sendo x_j a variável de decisão, b_i uma constante que limita uma variável ou conjunto de variáveis, l_j e u_j são os limites superior e inferior da variável de decisão x_j e Δ_j é um conjunto de valores discretos, dos quais um será selecionado para x_j . Além disso, qualquer restrição de desigualdade " \geq " pode ser transformada em uma restrição " \leq " multiplicando ambos os lados por -1.

2.3 APRENDIZAGEM DE MÁQUINA

A aprendizagem é uma das funções mais importantes para os seres humanos, animais e sistemas inteligentes. Os humanos ou animais aprendem com a experiência, nesse sentido, a aprendizagem é o processo pelo qual o conhecimento é adquirido por meio de experiências passadas.

Sistemas inteligentes como a ML usam computadores para simular a aprendizagem humana, permitindo que os computadores adquiram conhecimento por meio de algoritmos e dados do mundo real (PORTUGAL; ALENCAR; COWAN, 2018). A ML é um subcampo da AI, e nas últimas décadas ele melhorou exponencialmente. Certamente, isto reflectiu-se em numerosas aplicações como na indústria, medicina, economia, ciências naturais e técnicas, ecologia, finanças, entre outras (KONONENKO; KUKAR, 2007). Inicialmente, o termo ML foi cunhado por Samuel (1959), em sua publicação utilizando o jogo de damas; não obstante, ML foi estudado como um campo separado a partir de 1990 em diante (MICHALSKI; CARBONELL; MITCHELL, 2013).

De acordo com Langley e Simon (1995), Mjolsness e DeCoste (2001) e Jordan e Mitchell (2015) ML pode ser definida como o estudo de métodos computacionais com a capacidade de identificar padrões complexos em milhões de dados para aprender por meio da aquisição de conhecimento com base na experiência passada (dados históricos), para mais tarde ser aplicado no futuro para prever comportamentos aprendidos. ML usa técnicas estatísticas para a construção de modelos matemáticos, posto que sua principal tarefa é fazer uma inferência a partir de uma quantidade de dados de amostra (ALPAYDIM, 2004).

Além disso, algumas das possíveis aplicações utilizando ML são:

- Aplicações em negócios financeiros, com a implementação de sistemas automáticos para a avaliação de crédito e da avaliação de risco (POZZOLO et al., 2014; LI et al., 2019), ou aplicado em operações bancárias (CARRASCO et al., 2019).
- Aplicações em marketing, por exemplo para a construção de modelos de venda cruzada que ajuda a maximizar os lucros de uma organização (QIU et al., 2008), para a previsão de rotação de clientes (ZHANG et al., 2012), ou para analisar a comunicação entre uma determinada empresa e os consumidores por meio de redes sociais como Facebook, Twitter, entre outras (KWOK; YU, 2013).
- Aplicações em logística, com a criação de modelos inteligentes para o monitoramento preditivo de processos da cadeia de suprimentos, por exemplo, na área de transporte (METZGER et al., 2015), além de aplicativos para tomada de decisão.
- Aplicações em fabricação, por exemplo, para estimar o custo de fabricação dos componentes de um motor (LOYER et al., 2016), ou também para a criação de um modelo inteligente para a detecção de quebras de ferramentas, monitorando as forças de corte e o consumo de energia em uma fresagem (CHO et al., 2005).
- Aplicações em saúde, como a implementação de sistemas inteligentes de uma rede de conhecimento para diagnóstico médico (JIANG et al., 2017), para a criação de um quadro de avaliação

do câncer de mama por meio da classificação de lesões mamográficas (CHOUGRAD; ZOUAKI; ALHEYANE, 2018), ou também para a detecção de retinopatia diabética por meio de fotografias do fundo do olho (GULSHAN et al., 2016). Enfim, ML pode ser aplicado em diferentes áreas.

Por outro lado, como mostra a Figura 15, as técnicas de ML podem ser classificadas como supervisionada, não supervisionada, semi supervisionada, por reforço ou como aprendizagem profunda; sendo em cada um classificado uma variedade de algoritmos para sua aplicação.

Figura 15 – Técnicas de ML



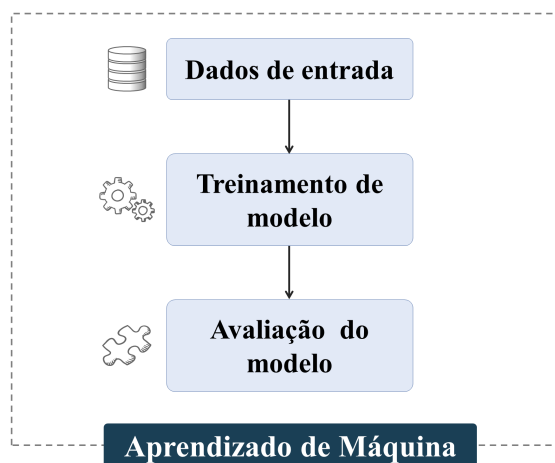
Fonte: Próprio autor(2020)

Na Subseção 2.3.2, é feita uma descrição básica da classificação das técnicas da Figura 15. Tendo em conta que uma descrição mais profunda sobre SL com técnicas de regressão é feita na Seção 2.4.

2.3.1 Processo de uma estrutura geral de ML

Em conformidade com Marsland (2015), o processo geral das técnicas de ML é dividida em 3 etapas conforme a Figura 16, levando em conta que o processo possui algumas pequenas variações no desenvolvimento das diferentes tarefas que compõem cada etapa de acordo com a classificação dos algoritmos (Subseção 2.3.2).

Figura 16 – Estrutura geral de ML



Fonte: Próprio autor(2020)

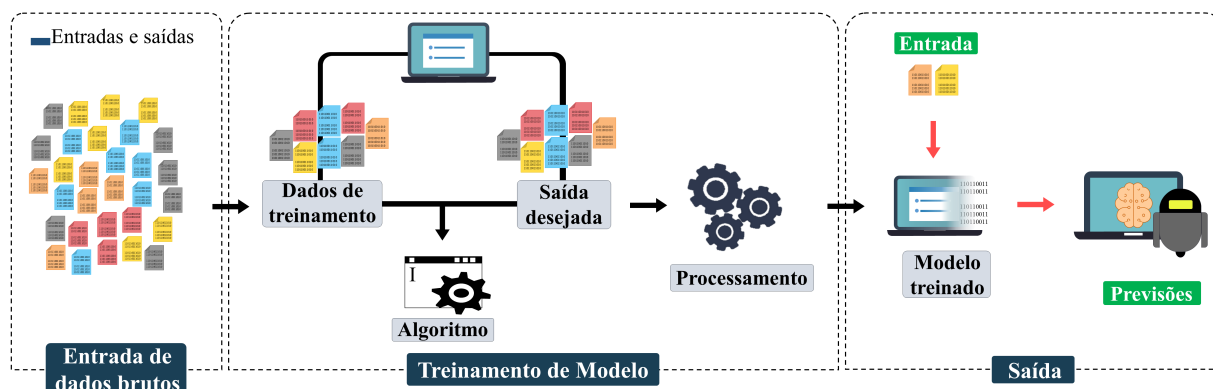
- **Dados de entrada:** Inicialmente os dados devem ser coletados e, posteriormente serão tratados com o objetivo que o conjunto de dados não possui erros significativos, falta de dados, além de que sejam selecionadas as características necessárias para o processo de aprendizagem.
- **Treinamento de modelo:** Um algoritmo deve ser selecionado de acordo com os objetivos, seguidamente, o processo de treinamento para a construção do modelo matemático é computado por meio de recursos computacionais. Em muitos casos no processo de treinamento é necessário selecionar ou modificar alguns parâmetros para obter um modelo ótimo.
- **Avaliação do modelo:** Este processo é fundamental para a implementação de um sistema, a precisão deve ser testada e avaliada com um grupo de dados que não foram utilizados para o treinamento.

2.3.2 Classificação das técnicas de ML

- **Aprendizagem Supervisionada:**

Aprendizagem Supervisionada (SL, *Supervised Learning*) ou aprendizagem com rótulos, têm a tarefa de aprender por meio de um conjunto de dados históricos a relação entre as características e suas saídas correspondentes, com o objetivo de representar a estrutura aprendida por meio de um modelo (KRAWCZYK et al., 2017). Por outro lado, a SL pode ser descrita como a técnica para inferir uma função a partir de dados de treinamento rotulados, que pode ser usado para prever novas saídas. (FLATH; STEIN, 2018). Como ilustrado na Figura 17, este modelo requer dados rotulados, que serão processados para a criação de um modelo de aprendizagem para prever novas saídas contínuas ou discretas.

Figura 17 – Processo de Aprendizagem Supervisionada



Fonte: Próprio autor(2020)

Além disso, pode-se distinguir dois modelos principais de SL:

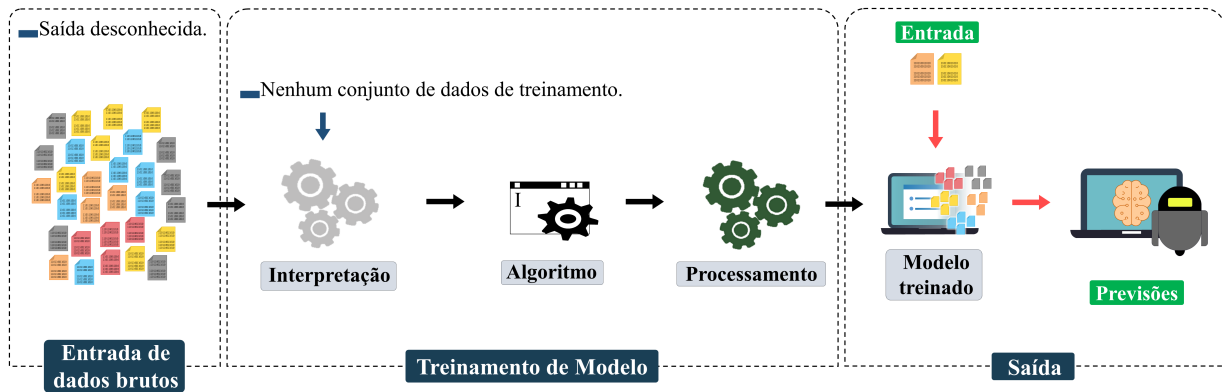
- **Modelo de Classificação:** O objetivo desta aprendizagem é classificar os padrões de entrada de um conjunto de dados rotulados em um conjunto de categorias ou classes, a ideia principal desta metodologia é calibrar um conjunto de classificadores individuais e combiná-los para obter um classificador superior a cada um (ROKACH, 2010). Assim, pode ser descrito que a variável de saída deste modelo têm uma natureza categórica (FIGUEIREDO, 2003), com uma função de saída chamada classificador.
- **Modelo de Regressão:** De acordo com (GOLDSTEIN; NAVAR; CARTER, 2016), os métodos ML manipulam diretamente o modelo tradicional de regressão para melhorá-lo, possibilitando a obtenção de um modelo atrativo de aprendizagem. Além disso, os modelos de regressão foram estudados a partir do campo da estatística, fornecendo diferentes abordagens para problemas de ML (FERNÁNDEZ-DELGADO et al., 2019). Por outro lado, os modelos de regressão, ao contrário dos modelos de classificação, são usados para aprender e prever funções contínuas (OUYANG et al., 2011).
- **Aprendizagem não Supervisionada:**

Aprendizagem não Supervisionada (USL, *Unsupervised learning*), muitas vezes é chamado de "aprender sem um instrutor", isso implica que um sistema de aprendizagem humano, animal ou artificial observa seu ambiente e, com base nas observações, gera um critério de aprendizagem associando as observações para obter respostas. Portanto, USL parece ser o mecanismo básico para a adaptação, tornando-se uma abordagem muito poderosa e promissora para alguns problemas práticos de processamento, mineração de dados e descoberta de conhecimento (OJA, 2002).

Nesse sentido, segundo Lee et al. (2018), pode-se descrever que o USL é capaz de processar dados não-rotulados em grande escala e detectar padrões interessantes, com o objetivo de classificar os dados e combiná-los em grupos por características semelhantes. Todo esse processo pode ser visto como a aprendizagem de um modelo probabilístico com dados não-rotulados (GHAHRAMANI, 2004).

Por outro lado, como apresentado na Figura 18, para o processo de treinamento de um modelo de aprendizagem não supervisionado, a máquina recebe apenas o conjunto de dados de entrada sem as saídas desejadas, passando por um processo de interpretação e processamento dos dados para construir representações da entrada, que pode ser utilizado para tomar decisões, fazer previsões, etc.

Figura 18 – Processo de Aprendizagem não Supervisionada



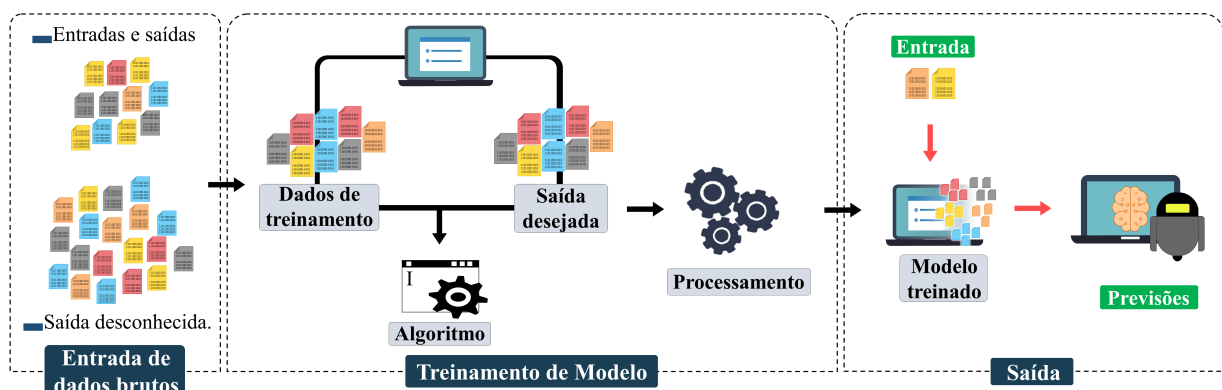
Fonte: Próprio autor(2020)

De acordo com Ghahramani (2004), os modelos clássicos representativos da aprendizagem não supervisionada são agrupamento ou clusterização (*Clustering* em inglês) e redução de dimensionalidade, sendo a redução da dimensionalidade o processo de redução do número de variáveis aleatórias que está sendo tratado, uma técnica bem conhecida é a Análise de Componentes Principais. Além disso, a clusterização visa agrupar os dados para descobrir o comportamento natural de um conjunto de padrões. A técnica mais conhecida é a Análise de Cluster ou Análise de Conglomerado, que é uma técnica de classificação estatística (JAIN, 2010).

- Aprendizagem Semi-Supervisionada:

Como o mesmo nome indica, Aprendizagem Semi-Supervisionada (SSL, *Semi-Supervised Learning*) mantém uma relação estreita entre aprendizagem supervisionada com dados rotulados e o aprendizagem não supervisionado com dados não-rotulados, como é apresentado na Figura 19.

Figura 19 – Processo de Aprendizagem Semi-Supervisionada



Fonte: Próprio autor(2020)

Os algoritmos utilizados na SSL assumem que os padrões detectados nos dados rotulados de entrada seguem uma estrutura de grupo e são extraídos da mesma distribuição marginal. Por outro lado, os dados não-rotulados fornecem informações muito úteis para a exploração dos dados no espaço de entrada no processo de aprendizagem. Conseqüentemente, o SSL requer menos esforço para processar os dados rotulados e, assim, oferecer maior precisão aproveitando a combinação dos dois tipos de dados (Huang et al., 2014).

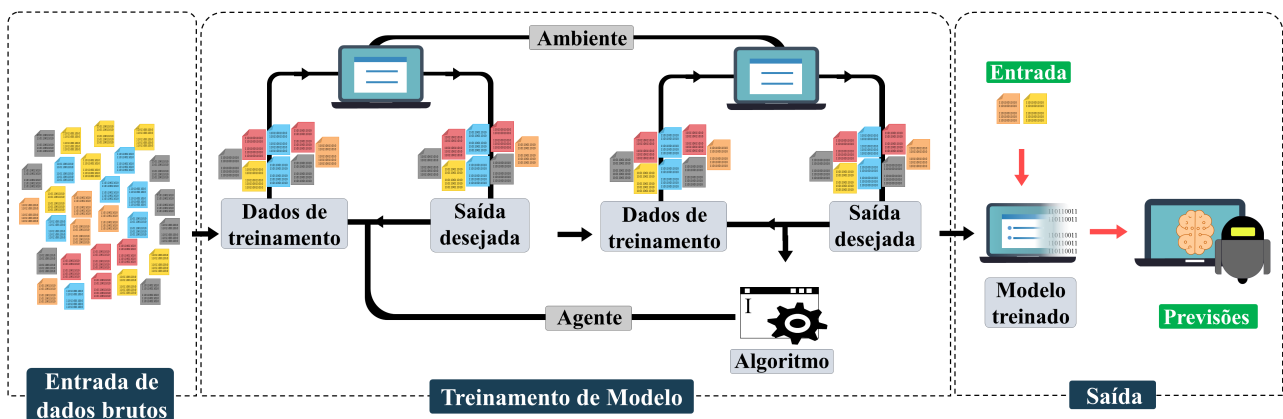
Além do mais, segundo Zhu e Goldberg (2009), a SSL pode abranger diferentes entornos de desenvolvimento, incluindo:

- Classificação, conhecido como classificação com dados rotulados e não-rotulados (ou dados parcialmente rotulados), é uma extensão do problema de classificação supervisionada. Seu objetivo principal é treinar um modelo de classificação SSL, a fim de ser melhor do que o classificador supervisionado.
 - Agrupamento ou clusterização, é uma extensão da clusterização sem supervisão. Seu objetivo é obter um melhor agrupamento com SSL do que o agrupamento usando apenas dados não-rotulados.
- Aprendizagem por Reforço:

O Aprendizagem por Reforço (RL, *Reinforcement Learning*) reproduz um princípio muito simples da natureza. Este aprendizagem consiste em aprender a decidir, criando um modelo baseado em tentativa e erro em um ambiente dinâmico, analisando qual ação leva a um melhor resultado (QIU et al., 2016); por outro lado, os dados de treinamento para esse modelo são não-rotulados.

Como é mostrado na Figura 20, na RL podem ser identificados os componentes de: agente e ambiente (ou entorno). Além disso, no modelo padrão de RL o agente está ligado ao seu ambiente por meio da percepção e ação e, o agente percebe o estado completo do ambiente e executa uma ação que gera imediatamente que o ambiente modifique seu estado, para mais tarde, o agente receber uma sinal de recompensa escalar. Por conseguinte, todo esse processo resulta em um estado de *feedback* (BUSONI; BABUSKA; SCHUTTER, 2008).

Figura 20 – Processo de Aprendizagem por Reforço



Fonte: Próprio autor(2020)

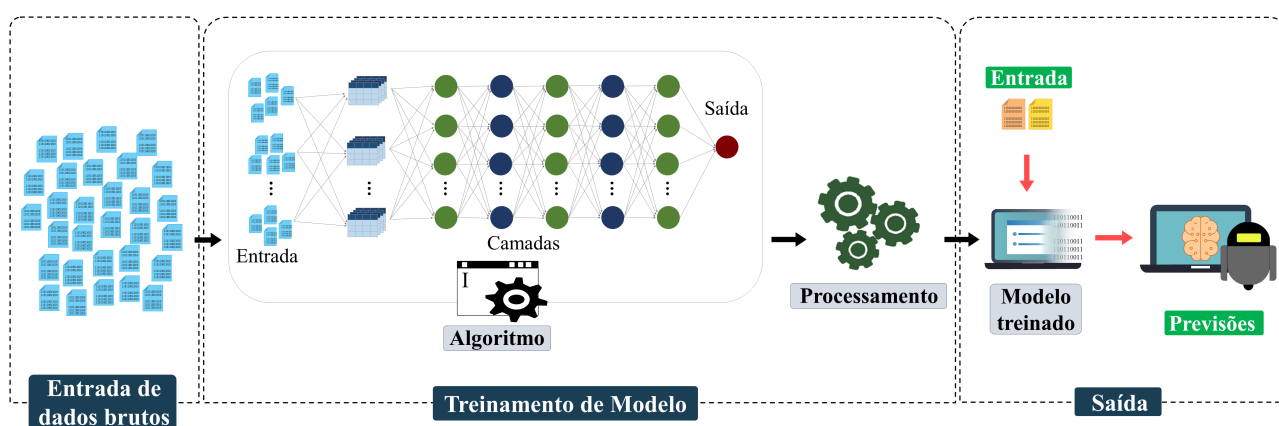
Em conformidade com Kaelbling, Littman e Moore (1996), existem duas estratégias para resolver problemas de RL:

- A primeira estratégia é buscar no espaço dos comportamentos gerados um que gere resultados ótimos no ambiente, esta abordagem têm sido aplicada por algoritmos genéticos e programação genética.
- A segunda estratégia é usar técnicas estatísticas e métodos de programação dinâmica para estimar a ação que afetará o estado do ambiente.

- **Aprendizagem Profunda:**

Conforme Deng e Yu (2013) e LeCun, Bengio e Hinton (2015), Aprendizagem Profunda (DL, *Deep Learning*) é uma classe de técnicas que processam dados não-lineares, com muitos níveis ou camadas de processamento que são explorados. A função das camadas de processamento é extrair e transformar as informações das variáveis de entrada, usando como entrada a saída da camada anterior. A Figura 21 descreve de maneira geral o processo de DL, começando com o conjunto de dados que é tratado por um tipo de algoritmo NN para treinar um modelo, as variáveis de entrada passam pelas camadas, gerando os pesos e limites. Todos os dados são processados, criando assim um modelo ML para ser testado e usado posteriormente.

Figura 21 – Processo de Aprendizagem Profunda



Fonte: Próprio autor(2020)

Além disso, os algoritmos que o DL usam principalmente são das classes supervisionadas e não supervisionadas (comumente redes neurais), utilizados para que o modelo aprenda automaticamente as representações hierárquicas. Comumente DL é aplicado em dar sentido aos dados de imagens, som ou texto, tendo com abordagens principais nas pesquisas propostas durante a última década os algoritmos de Rede de crenças profundas e Redes neurais convolucionais (QIU et al., 2016).

2.3.3 Descrição de terminologias em ML

De acordo com Bishop (2006), Shai e Shai (2014) e Marsland (2015), dentro do campo de ML são usados vários termos como dados de entrada, rótulo e dados de treinamento, sendo, nas linhas

seguintes brevemente descritos.

- Dados de entrada: Também conhecido como conjunto de domínios, é um conjunto de objetos ou instâncias. Uma instância é frequentemente representada por um vetor de características, na qual o vetor \mathcal{X} têm um conjunto de elementos ou dados, conforme a expressão (2.3.1).

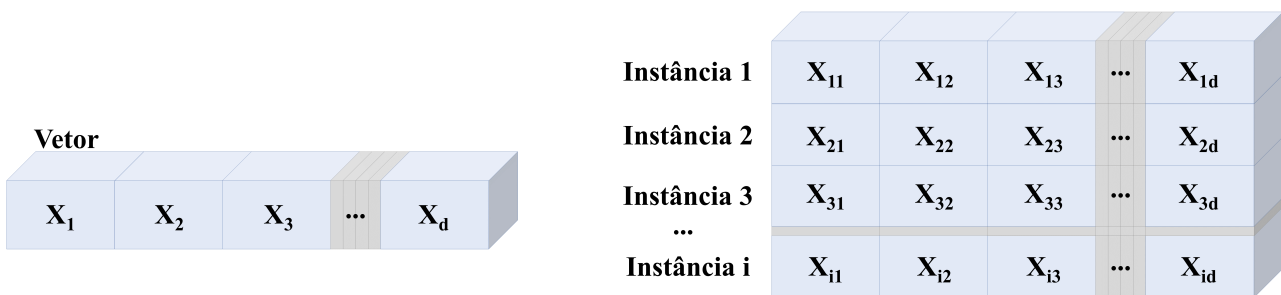
$$\mathcal{X} = \{x_1, x_2, \dots, x_d\} \in \mathbb{R}^d \quad d \geq 1 \quad (2.3.1)$$

sendo d a dimensão do vetor ou longitude do vetor e cada x_d é uma característica. Quando existem várias instâncias fala-se de um conjunto empilhado de vetores, como apresentado em (2.3.2).

$$\mathcal{X}_i = \{x_{i1}, x_{i2}, \dots, x_{id}\} \in \mathbb{R}^d \quad (2.3.2)$$

E cada característica do conjunto pode ser denotado como x_{id} , no qual d é a dimensão e i é a instância, como ilustrado na Figura 22.

Figura 22 – Conjunto de vetores de dados de entrada



Fonte: Próprio autor(2020)

- Rotulo: São as saídas pertencentes ao conjunto de dados, geralmente este conjunto de dados é chamado de dados rotulados e é representado por $\mathcal{Y}_i \in \mathbb{R}$, sendo i a instância e cada \mathcal{Y}_i corresponde a uma instância.
- Dados de treinamento: Eles são uma coleção de instâncias aleatórias que atuam como entradas para o processo de aprendizagem. Assim, o conjunto de dados não-rotulados, pode ser representado como a expressão (2.3.3)

$$S = \{\mathcal{X}_i\}_n \quad (2.3.3)$$

sendo n o número de instâncias de treinamento e X_i uma variável aleatória que se extrai de uma distribuição de probabilidade $P(X)$. Por outro lado, a expressão (2.3.4) é a representação dos dados rotulados ou dados com saída.

$$S = \{(\mathcal{X}_1, \mathcal{Y}_1)_1, (\mathcal{X}_5, \mathcal{Y}_5)_2, \dots, (\mathcal{X}_i, \mathcal{Y}_i)_n\} \quad (2.3.4)$$

Sendo a expressão (2.3.4) uma sequência finita de pares de $\mathcal{X} \times \mathcal{Y}$ e cada etiqueta \mathcal{Y}_i corresponde a \mathcal{X}_i em uma instância n de treinamento, sendo as \mathcal{X}_i e \mathcal{Y}_i variáveis aleatórias que são extraídos de uma distribuição de probabilidade $P(X, Y)$.

2.4 APRENDIZAGEM SUPERVISIONADA VOLTADA PARA OBTER FUNÇÕES DE REGRESSÃO COM MODELOS DE CAIXA BRANCA E CAIXA PRETA

O principal objetivo da SL é aprender a prever uma variável aleatória \mathcal{Y} , também conhecida como variável de saída, baseada em um conjunto de variáveis aleatórias denotadas por \mathcal{X} , também conhecidas como variáveis de entrada ou características. De acordo com Crisci, Ghattas e Perera (2012) e Marsland (2015), a tarefa de SL é aprender uma função, conforme a expressão (2.4.1)

$$f : \mathcal{X} \longrightarrow \mathcal{Y} \quad \text{ou} \quad \mathcal{X} \longrightarrow f(\mathcal{X}) \quad (2.4.1)$$

para uma família de funções (2.4.2)

$$f = \{f : \mathcal{X} \longrightarrow \mathcal{Y}\} \quad (2.4.2)$$

baseado nas variáveis $\mathcal{X} \times \mathcal{Y}$ pertencentes ao conjunto de dados de treinamento S da (2.3.4), com o objetivo de que a função apreendida $f(\mathcal{X})$ preveja um novo rótulo \mathcal{Y} com base no conjunto de dados de entrada \mathcal{X} .

Além disso, em conformidade com Figueiredo (2003), a função obtida é avaliada de acordo com sua precisão, assumindo que os novos dados de entrada $\{x_1, x_2, \dots, x_d\}$ seguem a mesma distribuição do conjunto de dados de treinamento S , além que a estrutura da função apreendida depende de um conjunto de parâmetros β (chamado coeficientes em regressão ou pesos em Redes Neurais-NN). Portanto, a expressão (2.4.3) apresenta a estrutura da função \mathcal{Y}

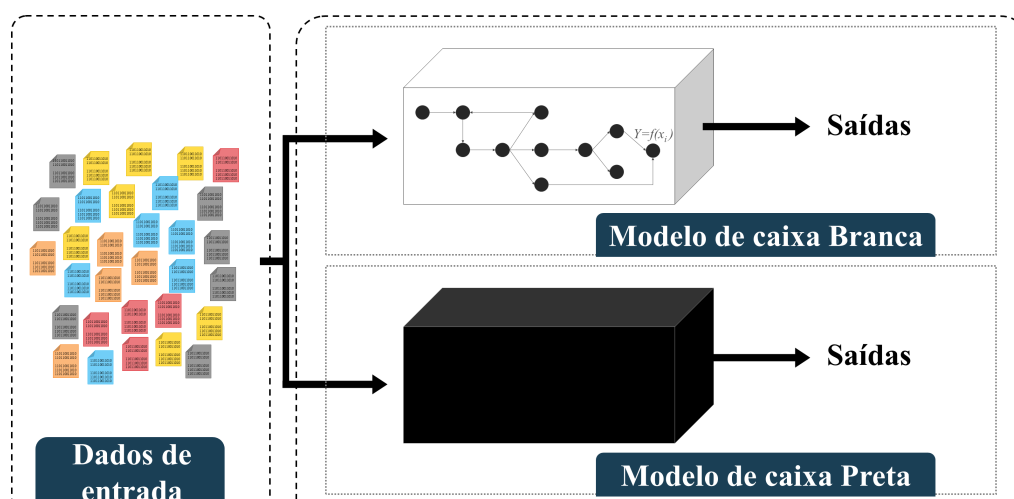
$$\mathcal{Y} = f(\mathcal{X}, \beta) \quad (2.4.3)$$

quem têm a tarefa de aprender uma função $f(\mathcal{X})$ e estimar os parâmetros β a partir do conjunto de dados de treinamento S .

Por outro lado, na SL, os problemas de classificação e regressão podem ser categorizados em dois grupos principais, tais como os modelos de caixa branca e os modelos de caixa preta. Dependendo da necessidade de interpretar a caixa de provas são classificados os algoritmos, seja por sua estrutura interna para entender a causa de uma decisão ou em contraste por sua funcionalidade, conforme exibido na Figura 23.

Em conformidade com Zhang e Tsai (2007) e Flath e Stein (2018), estes modelos podem ser descritos como:

Figura 23 – Modelo de Caixa Branca e Caixa Preta



Fonte: Próprio autor(2020)

2.4.1 Modelos de Caixa Branca

Também chamado de modelos transparentes, que possuem uma abordagem clássica permitindo ao usuário visualizar e analisar a estrutura do modelo criado. Entre os modelos mais antigos, estão os modelos de regressão linear para problemas de regressão e regressão logística para problemas de classificação, que são baseados em suposições de uma relação linear entre as variáveis de entrada e as variáveis de saída. As suposições entre a regressão linear e as variáveis podem ser sua maior fortaleza, porque são modelos simples de entender e eficiente de aprender. Por outro lado, é também uma fraqueza, porque seu poder preditivo é restrito. Nesse sentido, os modelos de árvores de decisão ou modelos baseados em regras são considerados para superar a deficiência dos relacionamentos não-lineares nos dados, levando em consideração que esses modelos são mais adequados para dados complexos e possuem um poder preditivo melhor que a regressão linear ou regressão logística.

a. Regressão Linear:

A Regressão Linear (LR, *Linear Regression*) é um algoritmo estudado e usado tanto no ML quanto na estatística, sendo neste último, um tópico estudado muito importante. A regressão linear investiga a influência de um conjunto de variáveis regressoras \mathcal{X} (variáveis independentes) em uma variável de resposta \mathcal{Y} (variáveis dependentes), com o objetivo de modelar a relação entre estas variáveis (GROMPING, 2015).

De acordo com Montgomery e Runger (2003), matematicamente, a regressão linear simples pode ser representada como a expressão (2.4.4)

$$\mathcal{Y} = \beta_0 + \beta_1 x + \varepsilon \quad (2.4.4)$$

em que os parâmetros β_0 e β_1 são os coeficientes de regressão. Por outro lado, a expressão (2.4.5) apresenta a representação da regressão linear múltipla

$$\mathcal{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_d x_d + \varepsilon \quad (2.4.5)$$

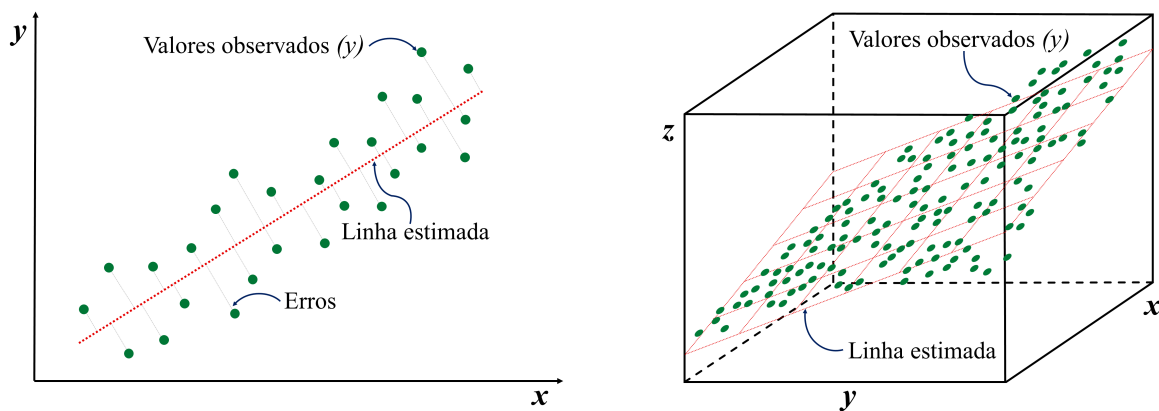
na qual, os parâmetros β_d representa a variação esperada na resposta \mathcal{Y} por unidade de variação unitária em x_d , sendo d a dimensão.

Em síntese, podem ser representados como a expressão (2.4.6), sendo as saídas de \mathcal{Y} a resultante da soma dos valores x_i , com $i = 1, 2, \dots, d$, multiplicado por os parâmetros constantes

$$\mathcal{Y} = \beta_0 + \sum_{i=1}^d \beta_i x_i + \varepsilon \quad (2.4.6)$$

na qual os parâmetros ou valores β_i definem uma linha reta que cruza ou se aproxima aos pontos de $\hat{\mathcal{Y}}$ no plano (MARSLAND, 2015), como apresentado na Figura 24.

Figura 24 – Regressão linear em duas e três dimensões



Fonte: Adaptação de Marsland (2015)

O processo de LR é calcular a distância entre um ponto e uma linha arbitrária definida, essa distância são os resíduos ou também chamados de erros de previsão (Figura 24), Assim, a linha é recalculada procurando uma que melhor se adéque aos dados, gerando o mínimo erro. Matematicamente, o valor residual é representado na expressão (2.4.7)

$$residual_i = y_i - \hat{y}_i \quad (2.4.7)$$

em que \hat{y}_i é o valor desviado de y_i . Além disso, o método mais utilizado é o método dos mínimos quadrados, que é a busca pelo valor mínimo da soma dos quadrados dos resíduos (SSR ou $SS(residual_i)$, *Sum of Squares of residual*) (BALI et al., 2016). Apresentado nas expressões (2.4.8) e (2.4.9).

$$SSR = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{ou} \quad (2.4.8)$$

$$SSR = \sum_{i=1}^n (y_i - (\beta_0 + \sum_{i=1}^d \beta_i x_i))^2 \quad (2.4.9)$$

Deste jeito, a medição do desempenho da função ou função de perda é denominada como erro quadrático médio (MSE, *Mean Squared Error*), que avalia a qualidade da função de previsão,

considerando que a função está constituída por as variáveis de entrada ou características e os parâmetros. Matematicamente a representação do MSE é apresentado na expressão (2.4.10)

$$MSE = \frac{1}{n} SSR = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.4.10)$$

na qual, n é o tamanho da amostra (SHAI; SHAI, 2014). Por outro lado, de acordo com Montgomery e Runger (2003) e Marsland (2015), as expressões (2.4.5) e (2.4.6) podem ser escritas em forma de matriz com a expressão (2.4.11)

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \quad (2.4.11)$$

sendo \mathbf{y} um vetor ($n \times 1$) dos rótulos, \mathbf{X} uma matriz ($n \times p$) dos dados \mathcal{X} , β um vetor ($p \times 1$) dos coeficientes de regressão e ε um vetor ($n \times 1$) dos erros aleatórios; na qual n é o tamanho da amostra e em $p = k + 1$ o k é a dimensão das características.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \text{ e } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Além disso, os parâmetros β da equação de regressão podem ser estimados a partir da expressão (2.4.12)

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (2.4.12)$$

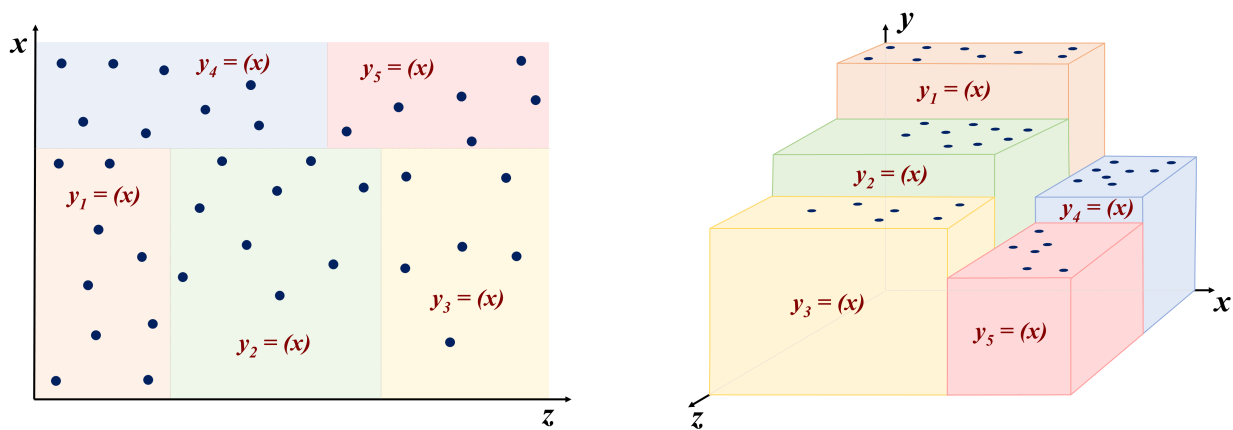
o símbolo "'"denota a transposta do vetor \mathbf{X} , a matriz $\mathbf{X}'\mathbf{X}$ é frequentemente não singular e também uma matriz simétrica ($p \times p$) e a matriz $\mathbf{X}'\mathbf{y}$ é um vetor coluna ($p \times 1$).

b. Árvores de Decisão:

As Árvores de Decisão (DT, *Decision Trees*) são uma das técnicas mais populares em mineração de dados (KARABADJI et al., 2017), ML e pesquisa operacional por sua simplicidade e eficiência em diversos problemas nas diferentes áreas de pesquisa, além de ter uma estrutura simples e de fácil interpretação (LI; ZHAO; ZHU, 2015).

Em síntese, segundo Fournier e Crémilleux (2002) e Wang et al. (2012), DT é uma estrutura em forma de árvore que é construída a partir de um conjunto de dados de treinamento particionado recursivamente, dividindo-se assim em subconjuntos de dados cada vez mais homogêneos, conforme Figura 25.

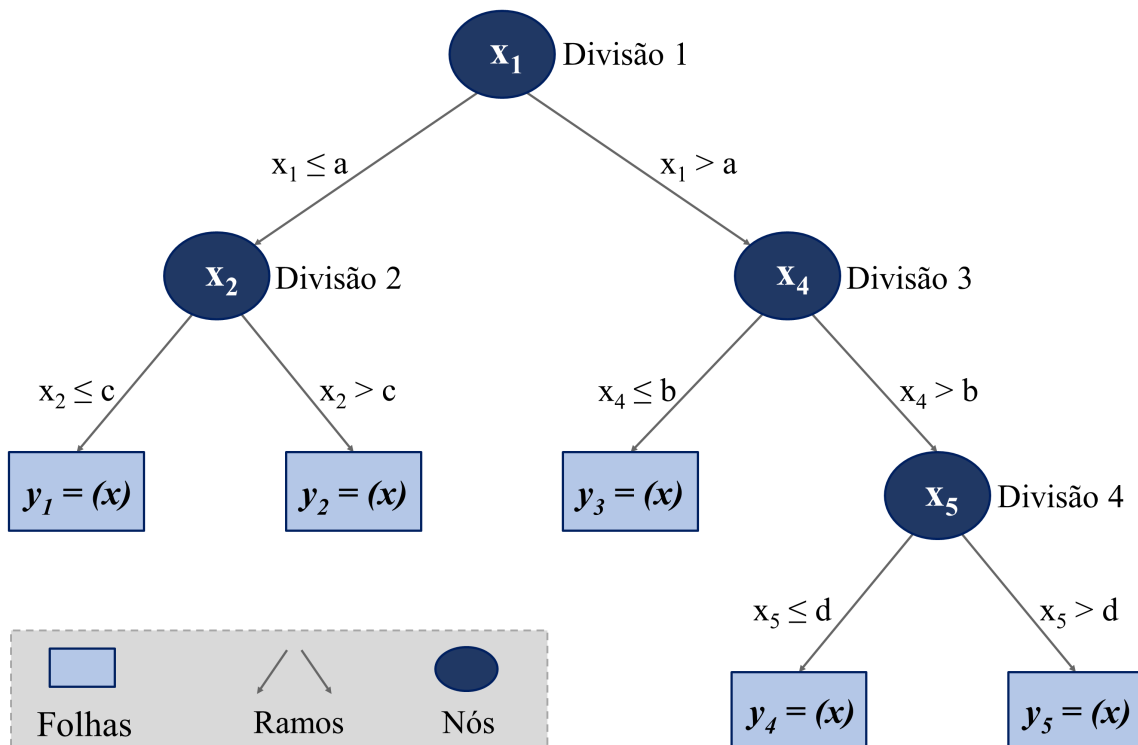
Figura 25 – Subconjuntos de dados em Árvore de Decisão



Fonte: Adaptação de Breiman et al. (1998) e Bishop (2006)

Com base na Figura 25, da partição em subconjuntos de dados, é construído uma DT com a estrutura mostrada na Figura 26, que é uma estrutura clássica de DT.

Figura 26 – Estrutura de Árvore de Decisão



Fonte: Adaptação de Breiman et al. (1998) e Bishop (2006)

A construção de DT apresenta um processo iterativo, em que, para construir o primeiro vértice ou nó raiz (Divisão 1 na Figura 26), é selecionada uma característica, a mais acentuada. Com base nos valores da característica é executado o processo de bifurcação, resultando em uma divisão binária do nó raiz, para que posteriormente os nós internos também sejam divididos em forma binária para formar os ramos que definirão as folhas da árvore. Finalmente, os campos

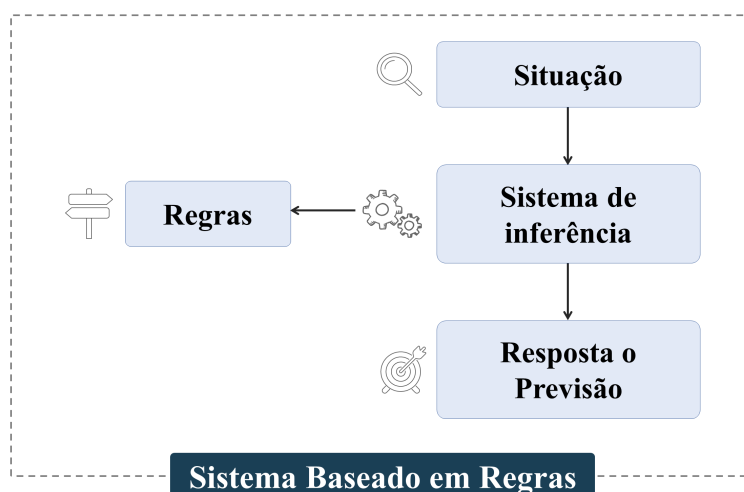
de partição potenciais ou nós são chamados de 'entradas' e as folhas são constituídas por os subconjuntos, que são os conjuntos de exemplos ou elementos da mesma classe (VILLE, 2013; GENRIKHOV; DJUKOVA; ZHURAVLEV, 2017a).

Por outro lado, a árvore da Figura 26 pode ser interpretada como um conjunto de regras, como:

$$\text{Regras de decisão} \left\{ \begin{array}{l} \text{SE } (x_1 \leq a) \text{ e } (x_2 \leq c) \text{ ENTÃO } y_1 = (x) \\ \text{SE } (x_1 \leq a) \text{ e } (x_2 > c) \text{ ENTÃO } y_2 = (x) \\ \text{SE } (x_1 > a) \text{ e } (x_4 \leq b) \text{ ENTÃO } y_3 = (x) \\ \text{SE } (x_1 > a) \text{ e } (x_4 > b) \text{ e } (x_5 \leq d) \text{ ENTÃO } y_4 = (x) \\ \text{SE } (x_1 > a) \text{ e } (x_4 > b) \text{ e } (x_5 > d) \text{ ENTÃO } y_5 = (x) \end{array} \right.$$

Nesta perspectiva, de acordo com Zhang e Tsai (2007), DT pode ser considerada como um sistema baseado em regras (RBS, Rule-based System), posto que DT consiste em uma série de regras "IF-THEN"(SE-ENTÃO, em português) que processam as informações ou dados de entrada em um processo de inferência usando o conjunto de regras para produzir uma saída ou previsão específica (representada na Figura 27). Portanto, "IF" é chamado como premissa ou antecedente e é composto por um conjunto de ações e "THEN" é chamado de consequente ou conclusão. Além disso, nos sistemas de previsão criados pelas DT de ML as regras são aprendidas automaticamente e podem ser extraídas, dado que DT é um método de caixa branca, assim, em sua operação do RBS de DT para previsão é utilizado o método de raciocínio de encadeamento progressivo, que consiste em um vetor de características de entrada que é aplicado até chegar a uma conclusão ou previsão.

Figura 27 – Estrutura de um Modelo baseado em Regras



Fonte: Zhang e Tsai (2007)

Por conseguinte, a DT em ML pode ser aplicado a problemas de classificação, também chamado de árvore de classificação, quando os rótulos são do tipo qualitativo, ou para problemas de regressão, chamados de árvore de regressão, quando os rótulos são valores contínuos.

Da mesma forma que na LR, o objetivo na árvore de regressão é definir uma regra para prever um valor numérico contínuo de saída y , ao inserir um vetor de variáveis de entrada \mathcal{X} , com base em um conjunto de observações do conjunto dados de treinamento S , conforme a expressão (2.3.4). Desta forma, na literatura podem ser encontradas diferentes ferramentas em DT. Assim, para problemas de regressão, um dos algoritmos mais conhecidos por sua flexibilidade é o algoritmo de Árvores de Classificação e Regressão (CART, *Classification and regression trees*), que baseia-se principalmente na previsão de um valor contínuo nos nós preditores ou folhas. Por outro lado, outro algoritmo muito eficiente é o algoritmo M5, também conhecido como Árvore de Modelo M5 (M5 model tree, em inglês), este algoritmo têm uma estrutura semelhante ao CART (GENRIKHOV; DJUKOVA; ZHURAVLEV, 2017b), com a diferença de que gera funções de LR nos nós do preditor (CASTILLO et al., 2017).

Algoritmo CART:

Em conformidade com Chou (1991), Breiman et al. (1998), Siciliano e Mola (2000), Alpaydim (2004) e Bishop (2006), nas linhas seguintes é descrito, matematicamente, o processo de construção da árvore de regressão com CART.

A árvore de regressão CART é formada pela divisão binária do conjunto de n observações de \mathcal{X} e as repetidas divisões descendentes deste conjunto em subconjuntos de \mathcal{X} mais homogêneo, até encontrar uma redução substancial do erro em cada folha do modelo criado, além de que em cada folha da árvore é atribuído um valor de resposta. Portanto, o conjunto completo de dados de treinamento S é composto pelos dados de entrada $\mathcal{X} = \{x_1, x_2, \dots, x_d\}$ e pelo conjunto de rótulos \mathcal{Y} . Na qual, é dividido em \mathcal{X}_t e \mathcal{Y}_t que correspondem ao conjunto de instâncias de treinamento associadas ao nó t . Além disso, o conjunto \mathcal{X}_t pode ser particionado nos subconjuntos \mathcal{X}_t^a e \mathcal{X}_t^b :

$$\mathcal{X}_t = \begin{cases} 1 & \text{if } x \in \mathcal{X}_t^a \\ 0 & \text{then } x \in \mathcal{X}_t^b \end{cases} \quad (2.4.13)$$

Essa divisão binária de cada nó da árvore deve atender $(\mathcal{X}_t^a \cup \mathcal{X}_t^b = \mathcal{X}_t) \wedge (\mathcal{X}_t^a \cap \mathcal{X}_t^b = \emptyset)$, e a união dos elementos de \mathcal{X}_t^a e dos elementos de \mathcal{X}_t^b é igual ao total de elementos de \mathcal{X}_t , também esses conjuntos de elementos não têm instâncias em comum. Por ourto lado, no caso em que não é necessário dividir um nó, ele será chamado de nó de folha ou folha e conterà a regra de predição ou também chamado de preditor $d(x)$ para atribuir um valor $y(t)$ em cada folha, que pode ser um valor constante ou uma função.

Além disso, para medir o desempenho ou precisão do preditor de uma amostra com n instâncias, é usado o método dos mínimos quadrados para procurar o melhor ajuste para os valores do preditor a fim de minimizar a soma dos resíduos ou erros, com a expressão (2.4.14)

$$R^*(d) = \sum_{i=1}^n (y_i - d(x_i))^2 = \sum_{i=1}^n (y_i - d(x_{i,d}))^2 \quad (2.4.14)$$

sendo y_i o valor real de saída da instância i , $i = 1, 2, \dots, n$, e $d(x_{i,d})$ é o valor previsto com base

no número de características d da instância i de x . Assim, a precisão do preditor $d(t)$ é medido por meio do MSE para avaliar a qualidade de previsão na amostra, com a expressão (2.4.15)

$$R(d) = \frac{1}{n} R^*(d)$$

ou

$$R(d) = \frac{1}{n} \sum_{i=1}^n (y_i - d(x_{i,d}))^2 \quad (2.4.15)$$

na qual $R(d)$ é a taxa de erro do preditor, com o objetivo de obter o valor mínimo de $R(d)$ para maximizar a precisão.

Por outro lado, para medir a precisão dos preditores de um determinado nó t da árvore, é calculada a taxa de erro ou eficácia do preditor do nó por meio da função $R(t)$, apresentado na expressão (2.4.16)

$$R(t) = \frac{1}{n} \sum_{i=1}^n (y_{i,t} - \bar{y}_i(t))^2 \quad (2.4.16)$$

na qual, a expressão (2.4.17) apresenta a função para o calculado do $\bar{y}_i(t)$

$$\bar{y}_i(t) = \frac{1}{n} \sum_{i=1}^n y_{i,t} \quad (2.4.17)$$

sendo, $\bar{y}_i(t)$ a média dos valores reais de saída $y_{i,t}$. Posteriormente, a expressão (2.4.18) de $R(T)$ é definida para medir a eficácia da predição da árvore T a partir dos preditores das folhas t .

$$R(T) = \sum_{i=1}^n R(t)$$

ou

$$R(T) = \sum_{i=1}^n \frac{1}{n} \sum_{i=1}^n (y_{i,t} - \bar{y}_i(t))^2 \quad t \in T \quad (2.4.18)$$

Nessa continuidade, a divisão do conjunto de dados em nós ou subgrupos destina-se a melhorar a eficiência da árvore, tendo em conta que as equações mostradas nas linhas anteriores são usadas para medir a precisão, minimizando os erros ou maximizando a eficiência, e obter o melhor ajuste do modelo. Assim, a divisão binária do nó t em subconjuntos de t_a e t_b (equivalente à Equação 2.4.13) é realizada dividindo uma variável x_d , baseado em um valor de corte C_t . A fim de encontrar a melhor divisão s^* da divisão s dos dados rotulados em grupos de valores altos e baixos, de um conjunto de todas as divisões D_s feitas, além de satisfazer a propriedade de $R(t) \geq R(t_a) + R(t_b)$ com o objetivo de maximizar a eficiência ou minimizar os erros.

Deste modo, a expressão (2.4.19) é usada para qualquer divisão s de um nó t , baseado em um valor de corte C_t , para dividir o nó em t_a e t_b :

$$\Delta R(s, t) = R(t) - R(t_a) - R(t_b) \quad C_t \in x_d \quad (2.4.19)$$

na qual, de todas as possíveis divisões s da variável x_d do conjunto D_s , é selecionada a melhor divisão s^* que maximiza a eficiência da árvore:

$$\Delta R(s^*, t) = \max_{s \in D_s} \Delta R(s, t)$$

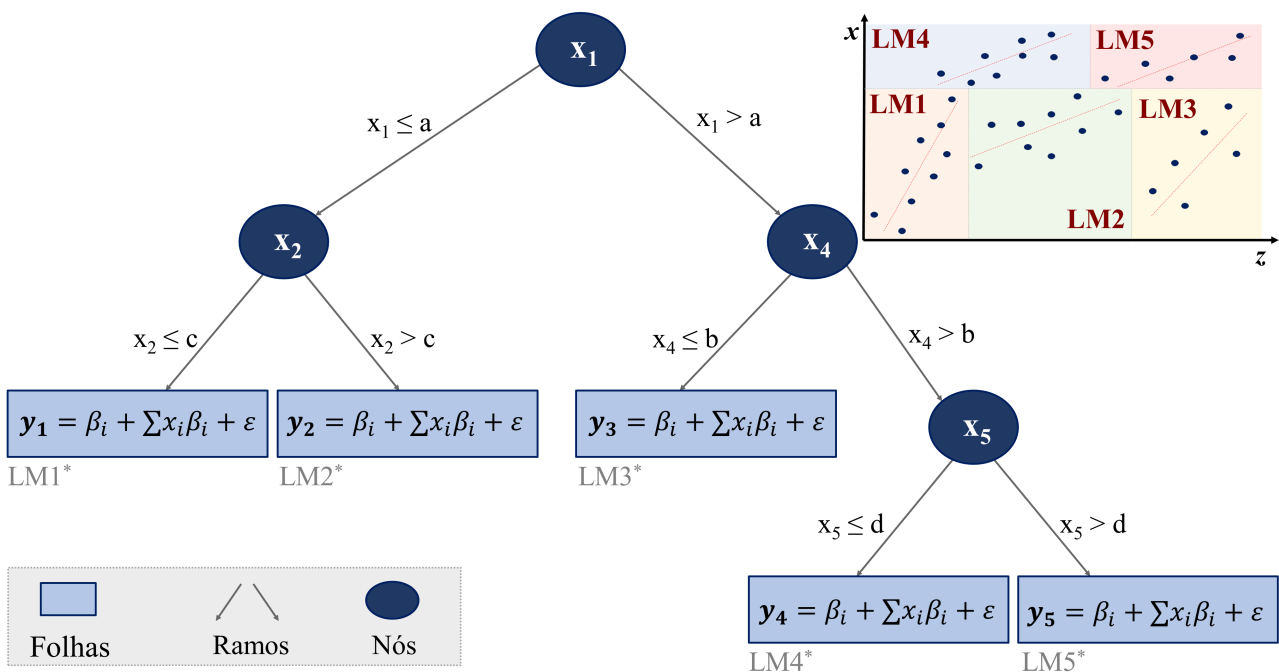
sendo todo este processo um ciclo iterativo para minimizar os desvios ou erros de um nó com o objetivo de diminuir $R(T)$.

Algoritmo de Árvore de modelo:

O nome de Árvore de modelo M5 ou *M5 Model Tree* foi utilizado, porque o conjunto de dados inseridos para o treinamento da DT e a criação de um modelo de ML são divididos em subconjuntos e representados por modelos lineares, ou seja, todas as folhas da árvore são representadas por equações lineares. Sendo assim, tais equações geradas são uma das principais diferenças com o algoritmo CART, pois o CART apresenta valores fixos em suas folhas e M5 apresenta as funções de LR.

Em conformidade com Etemad-Shahidi e Mahjoobi (2009) e Castillo et al. (2017), com M5 é possível criar modelos muito eficientes, além de utilizar dados de alta dimensionalidade sem problemas, mesmo com centenas de variáveis. A Figura 28 apresenta a estrutura da árvore do modelo M5, e pode ser claramente visualizado que sua estrutura é muito semelhante à estrutura clássica; com a diferença que o M5 têm funções em suas folhas.

Figura 28 – Estrutura de Árvore de Modelo M5



Fonte: Adaptação de Breiman et al. (1998), Bishop (2006), Keshtegar e Kisi (2018) e Seghier et al. (2019)

De acordo com a literatura, as árvores M5 foram inicialmente introduzidas por Quinlan (1992). Mais tarde o modelo inicial foi melhorado por Wang e Witten (1997) e passou a ser chamado algoritmo M5' ou M5P (*M5-Prime*), esse algoritmo é o estado atual das árvores de modelo.

Assim, o processo básico na construção da árvore de modelo é basicamente simples. Além disso, em conformidade com Wang e Witten (1997), Jung (2009) e Zhan, Gan e Hadi (2011), nas seguintes linhas são descritas matematicamente algumas equações importantes para a construção da M5P.

Nesta sequência, o processo de construção de um modelo M5P começa com a divisão do conjunto de dados de entrada \mathcal{X} em vários subespaços ou subconjuntos, da mesma forma que com o algoritmo CART. Além disso, o critério de divisão na M5P baseia-se principalmente na seleção do atributo ou variável x_i que maximiza a redução do erro, por meio do desvio padrão dos valores \mathcal{Y} que chegam a um determinado nó. Neste sentido, a redução do erro é calculada pela expressão (2.4.20), que representa a redução do desvio padrão (SDR, Standard Deviation Reduction)

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i) \quad (2.4.20)$$

na qual T é o conjunto de instâncias que atinge um nó, e T_i é um subconjunto $\{T_1, T_2, \dots, T_i\}$ da i -ésima divisão das instâncias de um nó com base em uma variável escolhida. Ainda, $sd(T)$ é o desvio padrão do conjunto T e $sd(T_i)$ é o desvio padrão dos valores-alvo de T_i como uma medida de erro.

Conjuntamente, cada folha do modelo criado contém um modelo de LR que foi calculado por meio da LR padrão (método dos mínimos quadrados), usando apenas as variáveis que são aplicadas nos nós inferiores, tendo o formato:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_j x_j \quad (2.4.21)$$

sendo as $\{x_1, x_2, x_3, \dots, x_j\}$ as variáveis de entrada ou atributos, e as $\{\beta_1, \beta_2, \beta_3, \dots, \beta_j\}$ são os parâmetros da equação de regressão. Além disso, geralmente no processo de construção da árvore de modelo pode ocorrer um ajuste excessivo, primeiro gerando que a precisão do modelo criado aumente, e depois diminua por causa do ajuste. Não obstante, para reduzir este problema é utilizado o método de poda, no qual os atributos são eliminados um por um, a fim de minimizar o erro estimado.

Esse procedimento de eliminação de atributos usa uma estimativa do erro esperado, que é aplicado em cada nó para o conjunto de dados de teste. Nesse procedimento, primeiro calcula-se a diferença absoluta entre o valor real e o valor previsto de cada uma das instâncias do conjunto de dados de teste, para depois ser multiplicado por um fator, como mostrado na expressão (2.4.22)

$$\frac{n+v}{n-v} \times \frac{\sum |y_i - \hat{y}_i|}{n} \quad (2.4.22)$$

sendo, n o número de instâncias de treinamento que chegam a um determinado nó, e v o número de parâmetros da equação de LR desse nó. Então, se o erro obtido do nó pai for menor, consequentemente, a folha pode ser eliminada.

Finalmente, após a poda, é usado o processo de suavização para aumentar a precisão da

previsão ao compensar as descontinuidades agudas que inevitavelmente ocorrerão entre os modelos lineares da árvore podada. Particularmente em modelos construídos a partir de um pequeno subconjunto de dados ou instância de treinamento. Esse processo de suavização usa um modelo de folha para prever um valor y , para subsequentemente filtrar esse valor previsto ao longo da rota percorrida até o nó raiz, com o objetivo de suavizar o valor de cada nó ao combiná-lo com o valor previsto do modelo linear. Matematicamente, esse cálculo é representado na expressão (2.4.23)

$$p' = \frac{np + kq}{n + k} \quad (2.4.23)$$

em que p' é o valor previsto que passará para o próximo nó superior, p é o valor previsto que passa de um nó inferior para o nó atual, q é o valor previsto para o nó atual, k é um valor predeterminado e n é o número de instâncias do subconjunto de dados que atinge o nó inferior.

Por outro lado, ao levar em conta um conjunto de dados de treinamento com valores perdidos nos atributos, o SDR é modificado na expressão (2.4.24), a fim de considerar esse conjunto de dados.

$$SDR = \frac{m}{|T|} \times \beta(i) \times \left[sd(T) - \sum_{j \in \{L,R\}} \frac{|T_j|}{|T|} \times sd(T_j) \right] \quad (2.4.24)$$

sendo m o número de instâncias sem valores ausentes para um atributo x_d , T é o conjunto de instâncias que chegam ao nó atual, $\beta(i)$ é o fator de correção mencionado acima, e finalmente, " TL, TR " são os conjuntos de dados resultantes da divisão binária do atributo x_d .

2.4.2 Modelos de Caixa Preta

Esses modelos são mais completos e menos interpretáveis do que os da caixa branca, fornecendo ao usuário valores de saída sem indicar a estrutura ou a forma na qual foi calculado. Entre os modelos considerados estão as máquinas de vetores de suporte-SVM que tratam cada observação como um vetor e usam superfícies multidimensionais para definir a relação entre características e resultados (LANTZ, 2015), e, as NN que consistem em várias camadas de neurônios artificiais que estão conectados uns aos outros para processar a informação.

a. Máquinas de vetores de suporte:

As Máquinas de vetores de suporte (*SVM, Support Vector Machine*) foi introduzido na década de 1960 por Vapnik, baseia-se na teoria da aprendizagem estatística, e é usado para a estimativa de valores de saída de classificação e regressão. Segundo Witten e Frank (2005) e Nagi et al. (2011), o SVM é uma poderosa técnica de ML usada também para problemas de regressão, sendo a ideia básica encontrar uma função em que seus resultados tenham uma aproximação aos valores de treinamento, minimizando o erro de previsão.

O conceito básico de SVM é mapear os dados de entrada \mathcal{X} , em um espaço não linear de maior dimensão (NAGI et al., 2011). De acordo com Smola e Scholkopf (2004), Alpaydim (2004) e Chang e Lin (2011) o principal objetivo do SVM é encontrar uma função $f(x)$, e pode

ser definida como o modelo linear (2.4.25).

$$f(x) = \langle w, x \rangle + b \quad , w \in \mathcal{X}, b \in \mathbb{R} \quad (2.4.25)$$

sendo os coeficientes w e b , o peso e o viés do vetor de suporte. O $\langle \cdot, \cdot \rangle$ denota o produto escalar em \mathcal{X} . O principal propósito da expressão (2.4.25), é obter um pequeno w . Isso pode ser escrito como um problema de otimização convexa, conforme (2.4.26).

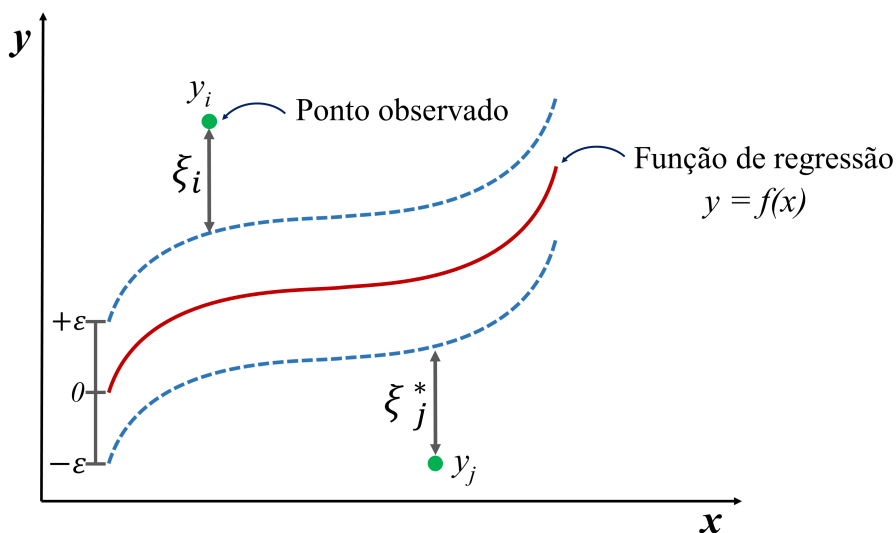
$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{sujeito a} \quad & \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \quad i = 1, \dots, l \end{cases} \end{aligned} \quad (2.4.26)$$

na qual, C e ε são parâmetros prescritos, $C > 0$ determina a compensação entre o nivelamento da função e a quantidade na qual os desvios maiores que ε são tolerados. O $(\xi_i + \xi_i^*)$ é a função de perda insensível a ε , e é descrito pela equação (2.4.27).

$$|\xi|_\varepsilon = \begin{cases} 0 & \text{if } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & \text{otherwise} \end{cases} \quad (2.4.27)$$

A Figura 29 apresenta o ajuste suave perda de margem para uma SVM linear.

Figura 29 – Ajuste suave perda de margem para uma SVM linear



Fonte: Adaptação de Smola e Scholkopf (2004) e Nagi et al. (2011)

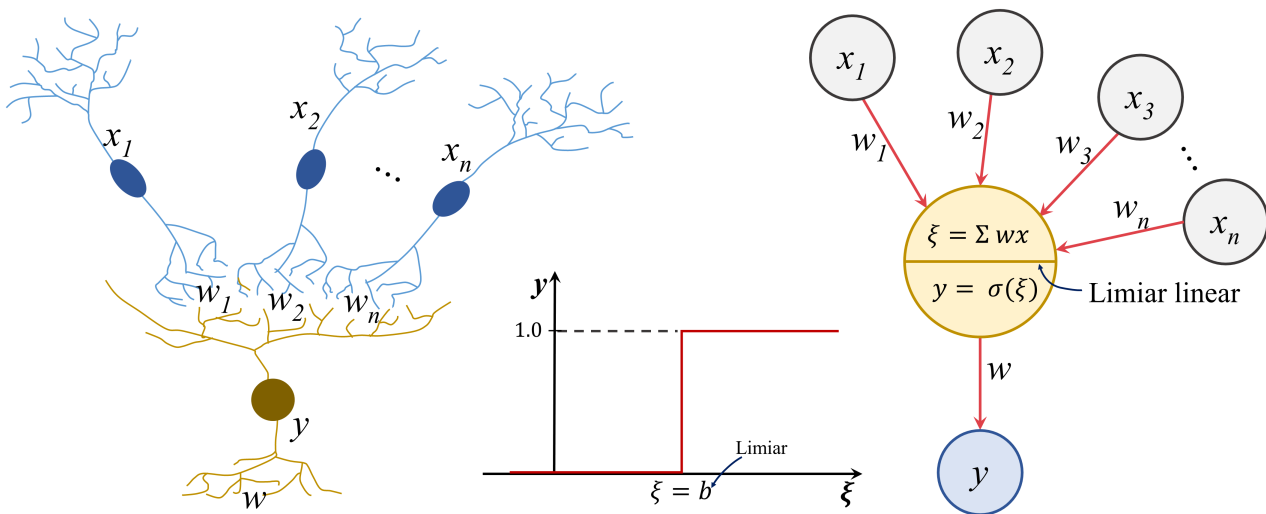
Conforme a Figura 29, somente os pontos fora do " ε -tubo" contribuem para o custo, sendo estas as variáveis de folga " ξ_i e ξ_i^* " que representam a distância entre os valores reais e os valores-limite correspondentes. Finalmente, vale esclarecer que o SVM não funciona bem com

conjuntos de dados extremamente grandes, pois os cálculos não são bem dimensionados e se tornam computacionalmente muito caros (MARSLAND, 2015).

b. Redes Neurais:

Uma rede neural artificial (NN) é inspirada na biologia do pensamento humano e modela a relação entre um conjunto de sinais de entrada e um sinal de saída (LANTZ, 2015). O modelo geral de um único neurônio artificial pode ser entendido de maneira muito semelhante ao modelo biológico, como é ilustrado na Figura 30.

Figura 30 – Modelo de uma rede neural artificial e o neurônio biológico



Fonte: Adaptação de Basheer e Hajmeer (2000) e Lantz (2015)

De acordo com a Figura 30, os sinais de entrada recebidos pelas dendritas x são ponderados com w de acordo com sua importância, o corpo da célula adiciona os sinais de entrada e o sinal é transmitido de acordo com uma função de ativação indicada por f com o limiar b , produzindo assim o sinal de saída y (BASHEER; HAJMEER, 2000) e (LANTZ, 2015). Em síntese, os NN são um grupo de nós interconectados a uma função matemática, além de possuir um meso adaptativo que é modificado de acordo com os valores de entrada de cada nó.

Segundo Bishop (2006) e Voyant et al. (2017), uma NN tem uma estrutura de um vetor de variáveis de entrada de dimensão d , uma camada oculta com neurônios m e com uma única variável de saída que pode ser representada como equação (2.4.28).

$$y(x, w) = \sigma \left(\sum_{j=1}^m w_j f \left(\sum_{i=1}^d w_{ji} x_i \right) \right) \quad (2.4.28)$$

na qual, cada um dos m neurônios ocultos é geralmente relacionado a uma função hiperbólica tangente:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.4.29)$$

e o vector $w = \{w_i, w_{ij}\}$ rege o mapeamento não-linear e é estimado durante a fase de treinamento ou aprendizagem do modelo por meio do conjunto de dados S .

2.4.3 Métricas de avaliação para modelos de regressão

Existem várias medidas alternativas que podem ser usadas para avaliar o sucesso de um modelo inteligente de previsão. As métricas de avaliação para modelos de ML são muito importante para medir o desempenho do modelo, avaliando as taxas de erro de previsão, a qualidade do ajuste e o desempenho do modelo em novas previsões (CASTILLO et al., 2017). Na literatura, pode-se constatar que, para avaliar a eficiência da previsão de modelos inteligentes podem ser usadas várias métricas estatísticas que medem o erro de previsão a fim de obter um modelo com qualidade em suas previsões. Assim, as métricas de avaliação medem a diferença entre os valores previstos e os valores reais (SHCHERBAKOV et al., 2013), e, esse processo é realizado usando dados de teste.

Embora existam várias métricas de avaliação, ainda não foi selecionada uma métrica padrão para medir a eficiência de um modelo de previsão (LI; SHI, 2010). Por esse motivo, e de acordo com Witten e Frank (2005), é aconselhável avaliar o desempenho de um ou vários modelos com base na comparação das diferentes métricas de desempenho. Deste jeito, Witten e Frank (2005), Shcherbakov et al. (2013) e Adetiloye e Awasthi (2017) descrevem algumas das métricas comumente usadas como o erro médio absoluto (MAE, *Mean absolute error*), raiz do erro quadrático médio (RMSE, *Root mean square error*), erro absoluto relativo (RAE, *Relative absolute error*), raiz do erro quadrático relativo (RRSE, *Root relative squared error*); que são usados para medir a aproximação dos valores previstos com os valores reais com base no erro calculado.

Por outro lado, o coeficiente de correlação (r , *correlation coefficient*) é uma medida que quantifica e mede a correlação estatística entre os valores obtidos na previsão e os valores reais, e varia entre ± 1 para resultados perfeitamente correlacionados e 0 quando não há correlação (MUN, 2014). O coeficiente de determinação (R^2 , *Coefficient of determination*) é comumente usado como um indicador que mede como o modelo criado se ajusta aos dados, e, indica a porcentagem da variação nas saídas do modelo; também têm uma variação entre $0 \leq R^2 \leq 1$ e, com um valor próximo a 1, indica que o modelo possui um bom ajuste e, quando está próximo de 0, indica que o modelo têm um ajuste baixo ou ruim. (ROSS, 2010).

Deste modo, as métricas usadas neste estudo para medir o desempenho do modelo de ML são apresentadas na Tabela 2. A aplicação dessas métricas neste trabalho ajudará a selecionar o modelo que tenha os valores mais baixos nas medições do MAE, RMSE, RAE e RRSE. Por outro lado, o modelo selecionado deve ter um valor muito próximo de 1 nos coeficientes r e R^2 .

Tabela 2 – Métricas de avaliação para modelos de regressão

Medida de desempenho	Fórmula
Erro médio absoluto	$MAE = \frac{1}{n} \sum_{i=1}^n \hat{y} - y_i $
Raiz do Erro Quadrático Médio	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2}$
Erro absoluto relativo	$RAE = \frac{\sum_{i=1}^n \hat{y} - y_i }{\sum_{i=1}^n \bar{y} - y_i }$
Raiz do erro Quadrático Relativo	$RRSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y} - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2}}$
Coefficiente de correlação	$r = \frac{n(\sum y_i \hat{y}) - (\sum y_i)(\sum \hat{y})}{\sqrt{[n \sum y_i^2 - (\sum y_i)^2][n \sum \hat{y}^2 - (\sum \hat{y})^2]}}$
Coefficiente de Determinação	$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$

Fonte: Witten e Frank (2005) e Castillo et al. (2017).

Sendo, em cada uma das equações, \hat{y} é o valor predito pelo modelo, y_i é o valor real da instância $i = 1, 2, \dots, n$, e \bar{y} é a média de y_i .

3 DESCRIÇÃO E MODELAGEM DO PROBLEMA

Neste capítulo é apresentada a aplicabilidade da sistemática proposta em um problema real na indústria, que objetiva implementar um modelo de previsão para a tomada de decisões utilizando técnicas de ML e otimização, a fim de estimar os valores ótimos para as variáveis de decisão. Além disso, de acordo com a Figura 2, neste capítulo também é detalhado o desenvolvimento das três principais etapas da abordagem metodológica proposta, começando com a Etapa 1 com a exploração e identificação do problema, além da coleta e preparação dos dados, a Etapa 2 com a seleção e criação do modelo preditivo com o algoritmo de aprendizagem supervisionado para obter equações de regressão, que são modeladas na Etapa 3 para o processo de otimização.

ETAPA 1: EXPLORAÇÃO

3.1 DESCRIÇÃO E IDENTIFICAÇÃO DO PROBLEMA

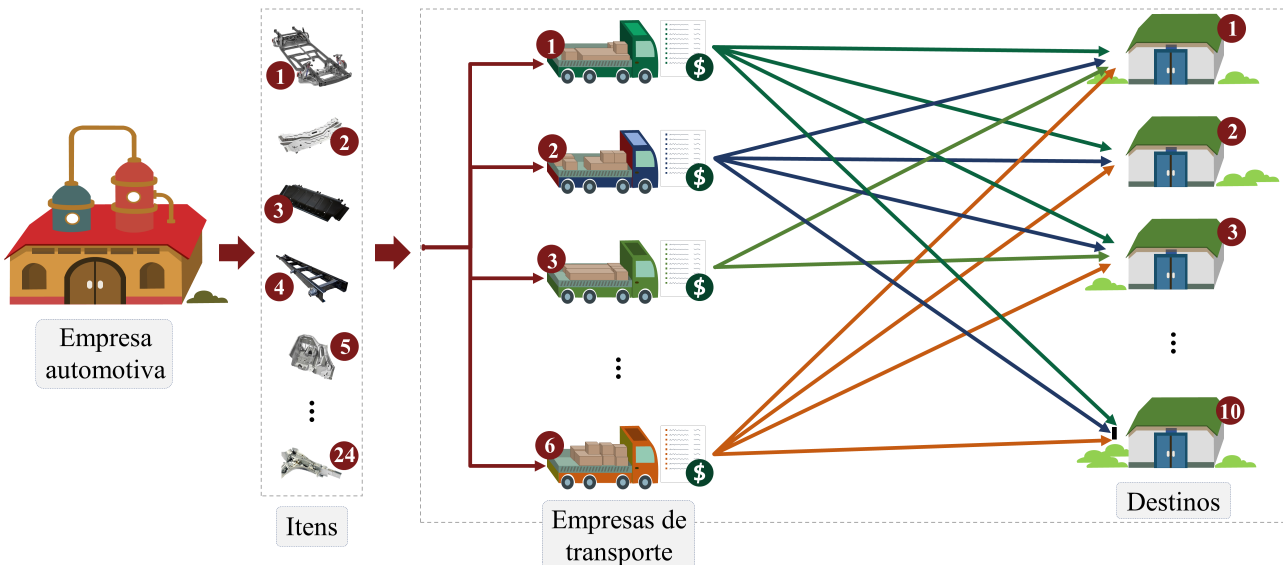
Atualmente a era digital é uma realidade. Com o tempo, as empresas viram a necessidade de integrar algumas das tecnologias emergentes em seus diferentes processos, trazendo como consequência a geração e armazenamento de uma grande quantidade de dados (CUI, 2019). Os dados armazenados, por si só não geram informações relevantes para uma organização, por isso, os dados armazenados são analisados para entender o mistério que os grandes conjuntos de dados ocultam, a fim de extrair seu verdadeiro valor e capturar novas oportunidades para melhorar a eficiência operacional (AGARWAL; DHAR, 2014; BAG et al., 2020).

Neste sentido, a exploração dos dados com técnicas de ML surge como uma alternativa para explorar os dados gerados e extrair seu valor (NIETO et al., 2019; KARTAL et al., 2016). Atualmente, a aplicação de técnicas de ML como uma ferramenta de análise para tomada de decisão tornou-se popular em várias áreas de estudo como medicina (BRADLEY; MEER; MCKAY, 2019; FUSE et al., 2019), economia (POURNARAS; PILGERSTORFER; ASIKIS, 2018), engenharia (FILIPIC; JUNKAR, 2000; KARTAL et al., 2016; NAJAFI, 2019), etc. Na área industrial, especificamente na área de logística, a tomada de decisão não é um processo fácil, pois o principal objetivo desse processo é sempre buscar a máxima eficiência das operações e gerar o menor prejuízo ou maior lucro para a empresa (KARTAL et al., 2016; BOHANEK; ROBNIK-SIKONJA; BORSTNAR, 2017; RODRIGUEZ-PADIAL; MARIN; DOMINGO, 2017). Por essa razão, a sistemática proposta é uma alternativa para a tomada de decisão, que visa implementar um modelo preditivo usando um algoritmo das técnicas supervisionadas de ML e integrando-o à abordagem de otimização

A aplicabilidade da sistemática proposta é comprovada em um problema real, em uma empresa automotiva multinacional brasileira do segmento de autopeças, situado no Vale de Paraíba. Assim, é estudado o problema da área de logística, especificamente, no contrato de empresas externas para o transporte dos diferentes itens fabricados (itens como chassi, longarinas, componentes para carrocerias o chassi, componentes leves, entre outros produtos metálicos para caminhões, ônibus ou pick-ups) enviados para um determinado destino (Figura 31); visando aplicar a sistemática para a implementação

de um modelo preditivo para estimar os melhores valores para as variáveis de decisão (empresas de transporte), a fim de tomar decisões mais rápidas, eficaz e que gere um custo menor.

Figura 31 – Diagrama do processo estudado



Fonte: Próprio autor (2020)

Por outro lado, vale ressaltar que, de acordo com o relatório "Logística, cadeia de suprimentos e transporte 2023: Mudança em alta velocidade" publicado pelo Forbes-Insights e Penske (2018), atualmente cerca de um terço das empresas terceirizam algumas operações de logística, cadeia de suprimentos e transporte. Essa terceirização aumentará com o tempo e, no futuro, estima-se que 61% das empresas dependerão significativamente mais de fontes externas para atender às suas necessidades. Além disso, o relatório descreve que a terceirização do transporte está tendo um crescimento significativo, pois essa medida traz benefícios para as empresas, como a geração de menos problemas de manutenção, menos problemas de recursos humanos, manutenção de uma frota atualizada, menos problemas de gerenciamento e também menores custos de capital, entre outros.

Nesse seguimento, a empresa em estudo terceiriza o processo de transporte dos itens produzidos. Conforme apresentado na Figura 31, a problemática estudada começa quando a empresa automotiva produz diferentes partes de carros que são transportadas para vários clientes (destinos), o transporte desses itens é realizado por diferentes empresas de transporte. Cada uma dessas empresas de transporte têm uma variação no preço de frete, além de nem sempre cobrir as mesmas rotas e nem sempre transportar os mesmos itens produzidos. No entanto, a aplicação da sistemática proposta nos dados disponíveis têm o objetivo de criar um modelo de ML de previsão para tomada de decisão, considerando também as restrições como as empresas de transporte e os destinos para modelar o processo de otimização, o que dará uma visão para selecionar a empresa externa que gerará um menor custo de transporte.

A empresa automotiva disponibilizou no conjunto de dados, para sua segurança, apenas algumas características que interferem no processo de transporte. A matriz de dados disponível é composta de 27 variáveis de entrada e 1 variável de saída, conforme a Tabela 3. Deste modo, na tabela pode-se observar que entre as variáveis de entrada estão as empresas de transporte x_1 e os destinos x_2 para os

quais um certo número de mercadorias é transportado, sendo que a empresa possui contrato de trabalho com 6 transportadoras e realiza o transporte para 10 destinos diferentes. Além disso, as variáveis x_4, x_5, \dots , até x_{27} têm o registro das quantidades enviadas de 24 linhas de produção diferentes, e a variável x_3 têm o registro das quantidades de outros itens enviados. Vale ressaltar que os valores registrados nas variáveis pertencem a peças produzidas por cada linha e podem ser para caminhões, ônibus ou pick-ups. Por último, a variável de saída é um vetor que contém o registro do custo da carga total transportada para um determinado destino, além disso, pode-se enfatizar que os itens das 24 linhas de produção nem sempre são transportados na mesma carga para o mesmo destino.

Tabela 3 – Variáveis *Inputs* e *Output* do conjunto de dados

Identificação	Nome da variável	Tipo	Descrição
<i>Inputs</i>			
Transporte	x_1	Empresas de transporte	Numérico\inteiro $x_1 = \{10, 20, 30, 40, 50, 60\}$
Destinos	x_2	Destinos	Numérico\inteiro $101 \leq x_2 \leq 110$
Outros itens	x_3	Quantidade outros	Numérico\inteiro $x_3 \geq 0$
Linha 1	x_4	Quantidade L1	Numérico\inteiro $x_4 \geq 0$
Linha 2	x_5	Quantidade L2	Numérico\inteiro $x_5 \geq 0$
Linha 3	x_6	Quantidade L3	Numérico\inteiro $x_6 \geq 0$
Linha 4	x_7	Quantidade L4	Numérico\inteiro $x_7 \geq 0$
Linha 5	x_8	Quantidade L5	Numérico\inteiro $x_8 \geq 0$
Linha 6	x_9	Quantidade L6	Numérico\inteiro $x_9 \geq 0$
Linha 7	x_{10}	Quantidade L7	Numérico\inteiro $x_{10} \geq 0$
Linha 8	x_{11}	Quantidade L8	Numérico\inteiro $x_{11} \geq 0$
Linha 9	x_{12}	Quantidade L9	Numérico\inteiro $x_{12} \geq 0$
Linha 10	x_{13}	Quantidade L10	Numérico\inteiro $x_{13} \geq 0$
Linha 11	x_{14}	Quantidade L11	Numérico\inteiro $x_{14} \geq 0$
Linha 12	x_{15}	Quantidade L12	Numérico\inteiro $x_{15} \geq 0$
Linha 13	x_{16}	Quantidade L13	Numérico\inteiro $x_{16} \geq 0$
Linha 14	x_{17}	Quantidade L14	Numérico\inteiro $x_{17} \geq 0$
Linha 15	x_{18}	Quantidade L15	Numérico\inteiro $x_{18} \geq 0$
Linha 16	x_{19}	Quantidade L16	Numérico\inteiro $x_{19} \geq 0$
Linha 17	x_{20}	Quantidade L17	Numérico\inteiro $x_{20} \geq 0$
Linha 18	x_{21}	Quantidade L18	Numérico\inteiro $x_{21} \geq 0$
Linha 19	x_{22}	Quantidade L19	Numérico\inteiro $x_{22} \geq 0$
Linha 20	x_{23}	Quantidade L20	Numérico\inteiro $x_{23} \geq 0$
Linha 21	x_{24}	Quantidade L21	Numérico\inteiro $x_{24} \geq 0$
Linha 22	x_{25}	Quantidade L22	Numérico\inteiro $x_{25} \geq 0$
Linha 23	x_{26}	Quantidade L23	Numérico\inteiro $x_{26} \geq 0$
Linha 24	x_{27}	Quantidade L24	Numérico\inteiro $x_{27} \geq 0$
<i>Output</i>			
	y	Custo de transporte	Numérico\decimal $x \geq 0$

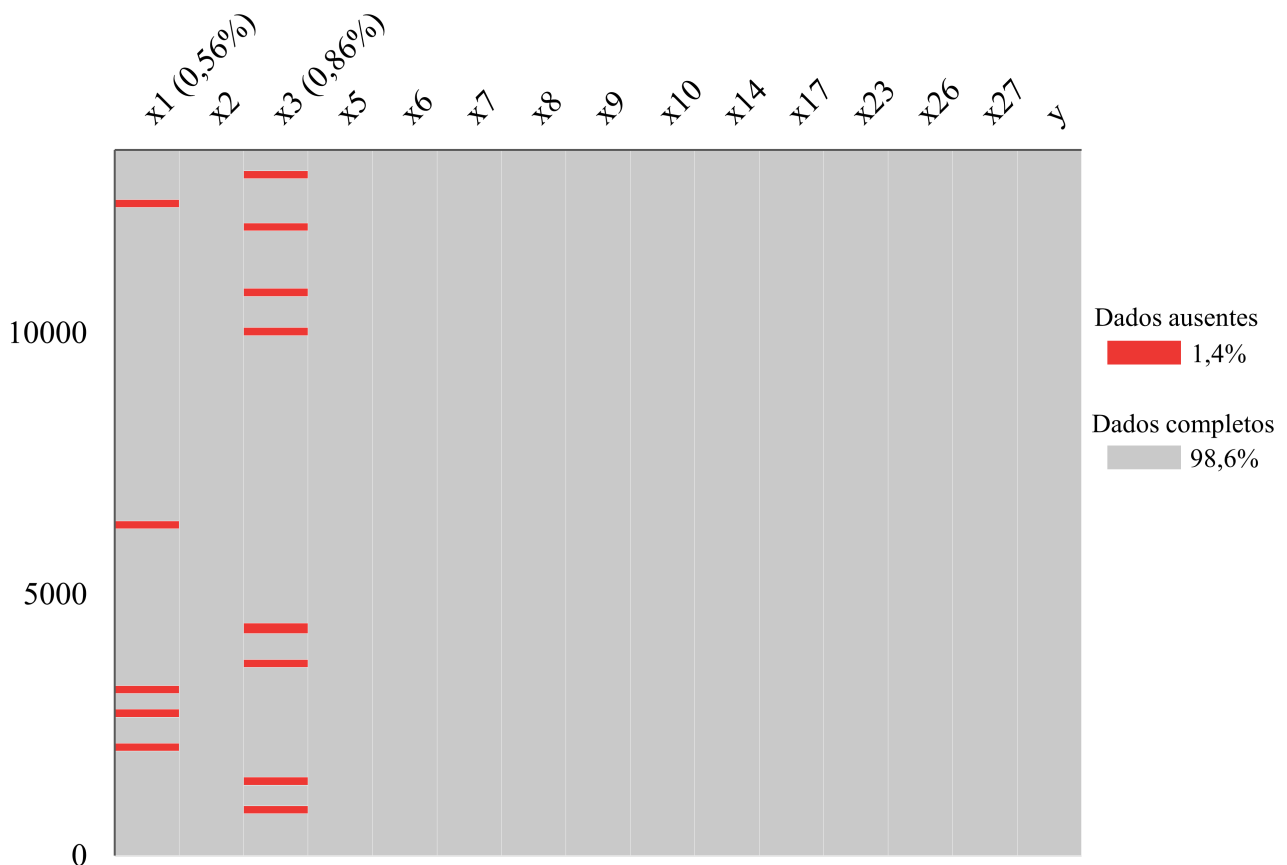
Fonte: Próprio autor (2020)

3.2 BANCO DE DADOS

A a matriz de dados fornecida pela empresa de estudo, além das 27 variáveis de entrada e uma variável de saída, é inicialmente formado por 15406 instâncias ou linhas de dados. O conjunto de dados foi inicialmente analisado, a fim de encontrar dados ausentes (*missing data*). Muitas vezes os valores ausentes são uma situação recorrente nas bases de dados e são considerados como um problema para a análise e criação de modelos preditivos, pois pode haver uma diferença considerável entre as previsões feitas com um modelo treinado com um conjunto de dados com valores ausentes e um modelo com um conjunto de dados sem problemas.

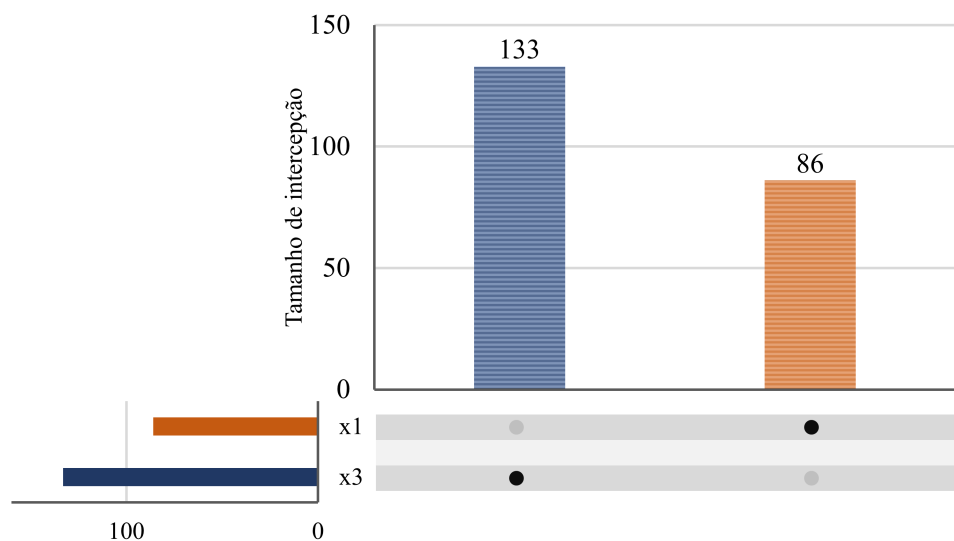
Nesta sequência, foram encontrados um total de 219 instâncias com dados incompletos em algumas variáveis, ou seja, 98,6% das instâncias estão com dados completos em cada uma das variáveis, como ilustra a Figura 32.

Figura 32 – Valores ausentes no conjunto de dados



Fonte: R Core Team (2020)

Note-se que a Figura 32 ilustra a porcentagem de valores omissos no conjunto de dados de 27 variáveis, sendo que as variáveis x_1 e x_3 estão incompletas. Além disso, a Figura 33 apresenta o número de total de valores ausentes em cada uma das variáveis, na qual pode ser visualizado que a variável x_1 contém 133 instâncias com dados incompletos e a variável x_3 contém 86 instâncias incompletas.

Figura 33 – Valores ausentes nas variáveis x_1 e x_3 

Fonte: R Core Team (2020)

Instâncias com valores ausentes foram excluídas, para não prejudicar o modelo. Assim, a Tabela 4 apresenta um resumo dos dados usados para o desenvolvimento da pesquisa, que contém um total de 15 187 instâncias.

Tabela 4 – Alguns dados do conjunto de dados do estudo

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	...	x_{26}	x_{27}	y
20	106	127	0	12	0	0	80	0	349	1097	...	59	0	1168.2
30	104	35	0	0	0	0	0	0	0	230	...	4799	0	1646.5
20	105	224	0	0	0	0	0	0	0	0	...	0	95	2050.0
10	102	127	0	0	1	0	0	0	189	0	...	200	0	2319.1
60	103	13	3	0	0	0	100	0	175	0	...	118	0	1644.6
20	105	0	0	0	0	17	0	0	0	0	...	0	0	966.8
50	109	6	0	0	0	0	0	0	0	0	...	0	0	1109.4
10	110	8	0	0	0	0	0	0	0	0	...	0	11417	2430.3
20	105	0	0	0	0	16	0	0	0	0	...	0	0	968.9
60	101	189	0	0	0	0	0	0	329	2470	...	481	0	4736.5
40	107	0	0	0	0	0	0	0	0	0	...	12	0	1212.4
10	102	20	0	0	6	0	0	0	0	0	...	0	0	1507.3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
60	103	4	8	0	0	0	5	0	9	0	...	18	0	1648.5
20	107	0	0	0	0	0	0	0	0	0	...	3019	83	3955.1

Fonte: Próprio autor (2020).

Analisando as variáveis das linhas de produção do conjunto de dados, a Tabela 5 apresenta o número de vezes que uma linha de produção foi enviada, o valor total enviado, a média dos embarques, etc. Os valores apresentados na tabela têm uma diferença no número de embarques de cada linha, porque nem sempre os mesmos itens são enviados para um destino.

Tabela 5 – Análise dos embarques por linha de produção

(continua)

Variáveis	Embarques	Quantidade	Média	Desvio padrão	Mínimo	Quartil 1	Quartil 2	Quartil 3	Máximo
Outros itens	x_3	9756	79	101,28	1	8	19	125	589
Linha 1	x_4	822	9	3,84	1	6	9	12	24
Linha 2	x_5	393	6	6,50	1	4	4	8	60
Linha 3	x_6	241	2	1,45	1	1	2	3	6
Linha 4	x_7	2858	15	2,40	1	13	15	16	26
Linha 5	x_8	360	203	130,81	1	85	161	357	550
Linha 6	x_9	2253	49	6,18	46	47	48	48	78
Linha 7	x_{10}	4018	218	267,48	1	32	103	286	1330
Linha 8	x_{11}	2379	2884	2428,91	1	859	2236	4501	9954
Linha 9	x_{12}	2410	777	709,46	1	236	620	1196	5871
Linha 10	x_{13}	702	164	124,36	1	99	122	185	760
Linha 11	x_{14}	3893	69	50,65	1	35	58	92	402
Linha 12	x_{15}	199	344	45,06	270	301	342	387	420
Linha 13	x_{16}	406	258	127,01	54	163	249	380	489
Linha 14	x_{17}	159	33	17,18	5	19	30	47	65
Linha 15	x_{18}	125	378	71,18	250	319	385	442	500
Linha 16	x_{19}	101	13	10,49	3	5	7	23	34
Linha 17	x_{20}	318	266	128,48	150	190	228	267	699
Linha 18	x_{21}	266	27	16,18	1	13	27	41	56
Linha 19	x_{22}	415	22	8,18	8	15	23	29	35

Fonte: Próprio autor (2020).

Tabela 5 - Análise dos embarques por linha de produção

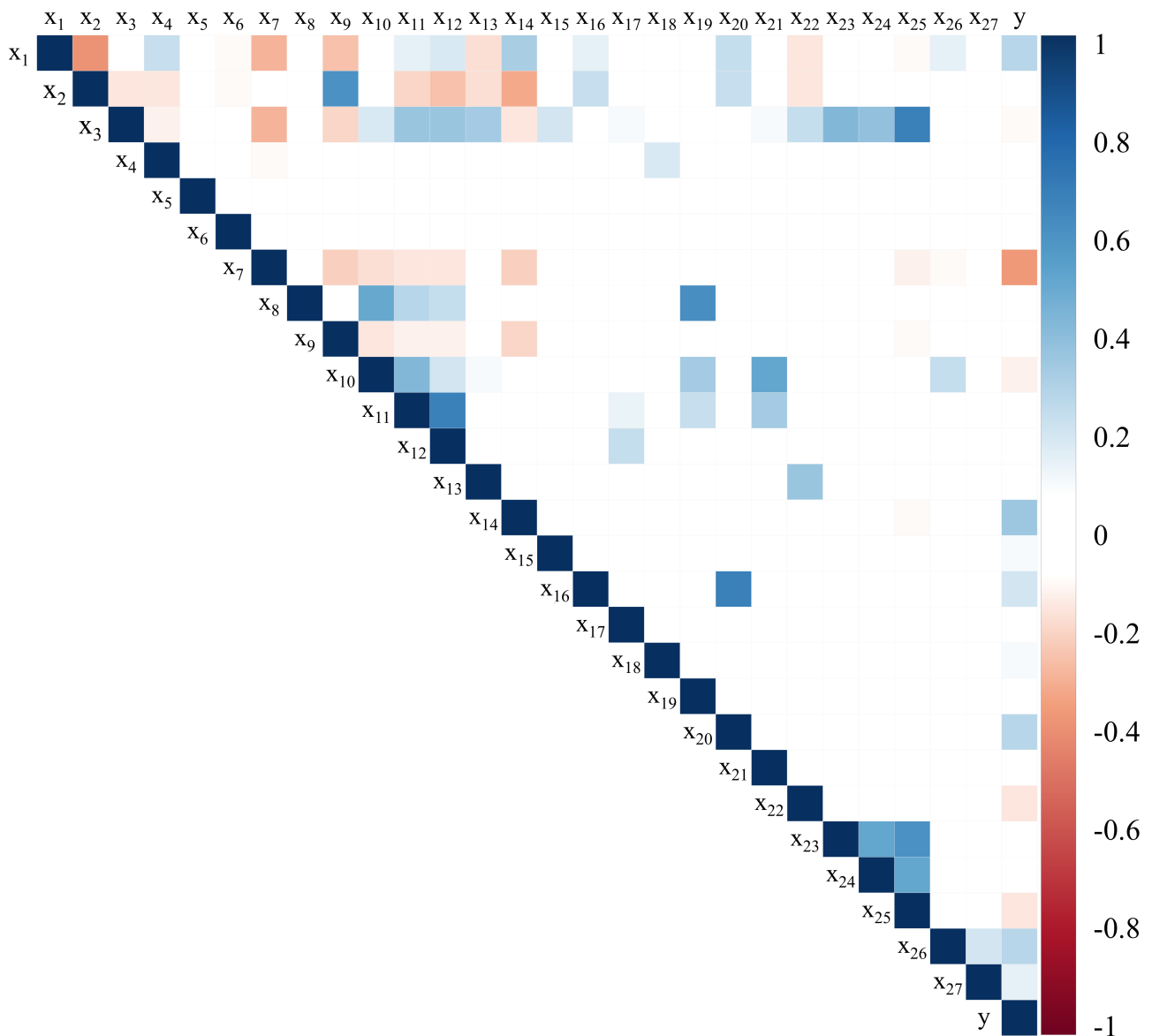
Variáveis	Embarques	Quantidade	Média	Desvio padrão	Mínimo	Quartil 1	Quartil 2	Quartil 3	(conclusão)
									Máximo
Linha 20 x_{23}	314	104697	333	33,35	276	302	337	362	388
Linha 21 x_{24}	342	11594	34	19,56	2	23	32	37	128
Linha 22 x_{25}	994	684131	688	227,93	1	705	771	827	986
Linha 23 x_{26}	2779	1176896	423	698,94	1	108	248	408	6753
Linha 24 x_{27}	264	264296	1001	2129,17	1	72	97	120	11417

Fonte: Próprio autor (2020).

Por outro lado, é analisada a correlação entre as variáveis de entrada. Essa análise é importante, porque o treinamento de modelos de ML com um conjunto de dados com variáveis altamente colineares resulta em um modelo que gera previsões seriamente afetadas em sua precisão.

Assim, um dos métodos estatísticos mais utilizados para o diagnóstico da correlação das variáveis é o coeficiente de correlação (r). Portanto, de acordo com Dormann et al. (2013), se as correlações entre duas variáveis excederem o limite de $\pm 0,7$, a correlação é considerada alta e pode distorcer seriamente a estimativa de um modelo e as previsões futuras. Por conseguinte, a Figura 34 apresenta o gráfico de correlação entre as variáveis do banco de dados.

Figura 34 – Gráfico de correlações entre variáveis



Fonte: R Core Team (2020)

No entanto, a Tabela 17 do Apêndice A apresenta os valores da análise de correlação das variáveis de estudo. Nesta sequência, na Figura 34 e na Tabela 17, pode-se ver que a correlação entre as variáveis de entrada do conjunto de dados é inferior a $\pm 0,7$, por conseguinte, não há colinearidade entre as variáveis e podem ser considerada para o treinamento do modelo de previsão.

Além disso, a Tabela 6 apresenta a análise estatística de todas as variáveis. Para essa análise, foram utilizadas estatísticas como média, desvio padrão, valores mínimos e máximos e quartis do conjunto de dados. Os valores zero nos quartis são devidos à quantidade de zeros registrados nessas variáveis, esses valores não se devem à falta de dados ou perda de dados, pelo contrário, significa que foram enviados zero produtos dessa linha de produção.

Tabela 6 – Análise estatística das variáveis do conjunto de dados

	Média	Desvio padrão	Mínimo	Quartil 1	Quartil 2	Quartil 3	Máximo
x_1	29,90	18,32	10	20	20	60	60
x_2	104,84	2,89	101	102	105	106	110
x_3	50,48	89,49	0	0	8	58	589
x_4	0,46	2,13	0	0	0	0	24
x_5	0,17	1,46	0	0	0	0	60
x_6	0,03	0,33	0	0	0	0	6
x_7	2,77	5,86	0	0	0	0	26
x_8	4,80	36,81	0	0	0	0	550
x_9	7,30	17,66	0	0	0	0	78
x_{10}	57,80	167,97	0	0	0	8	1330
x_{11}	451,73	1422,15	0	0	0	0	9954
x_{12}	123,38	400,69	0	0	0	0	5871
x_{13}	7,60	43,66	0	0	0	0	760
x_{14}	17,59	39,44	0	0	0	8	402
x_{15}	4,51	39,51	0	0	0	0	420
x_{16}	6,90	46,54	0	0	0	0	489
x_{17}	0,35	3,80	0	0	0	0	65
x_{18}	3,11	34,76	0	0	0	0	500
x_{19}	0,09	1,38	0	0	0	0	34
x_{20}	5,58	42,43	0	0	0	0	699
x_{21}	0,48	4,18	0	0	0	0	56
x_{22}	0,60	3,84	0	0	0	0	35
x_{23}	6,89	47,69	0	0	0	0	388
x_{24}	0,76	5,82	0	0	0	0	128
x_{25}	45,05	179,93	0	0	0	0	986
x_{26}	77,49	340,85	0	0	0	0	6753
x_{27}	17,40	309,24	0	0	0	0	11417
y	1503,63	616,65	644,06	1088,82	1420,42	1634,92	4952,01

Fonte: Próprio autor (2020).

ETAPA 2: MODELO DE APRENDIZAGEM DE MÁQUINA

3.3 MODELAGEM DO PROBLEMA E ANÁLISE DOS RESULTADOS OBTIDOS

Para a criação do modelo de previsão de ML são testados os algoritmos de aprendizagem supervisionado de máquina de vetores de suporte, redes neurais, regressão linear e árvore de decisão com os algoritmos CART e M5P que gera equações de regressão nas folhas finais para o cálculo das previsões. Nesta seção, são usados o *software* livre R e o conjunto de dados com 15187 instâncias.

Dado que, todo modelo de previsão precisa ser testado, o conjunto de dados é dividido em dados de treinamento e dados de teste. Deste modo, não há uma maneira padrão de dividir o conjunto de dados, pois essa divisão depende dos dados a serem utilizados (KOTU; DESHPANDE, 2019). Na literatura, existem estudos em que o conjunto de dados foi dividido em diferentes proporções (FENG; ZHOU; DONG, 2019; RANASINGHE et al., 2017), essa divisão depende da quantidade, variabilidade e incerteza dos dados. Por esse motivo, é importante que os dados sejam divididos em subconjuntos de dados representativos para treinamento e teste. Além disso, a eficácia do modelo é analisado em relação a esses subconjuntos de dados (MAY; MAIER; DANDY, 2010).

O conjunto de dados é dividido aleatoriamente em 75% (11390 instâncias) dos dados para o treinamento e 25% (3793 instâncias) para o teste. A seguir, os modelos testados de NN, SVM, LR, CART e M5P são descritos e mostradas as estatísticas dos modelos e o gráfico de dispersão dos valores reais com os previstos de cada modelo, com finalidade de selecionar o modelo de previsão com a maior eficiência.

- Árvore de modelo M5P:

O modelo M5P possui uma estrutura de 77 nós e 78 modelos lineares (também chamado de folhas de previsão). A Tabela 7 apresenta o resumo estatístico do modelo M5P, obtido com o *software* R.

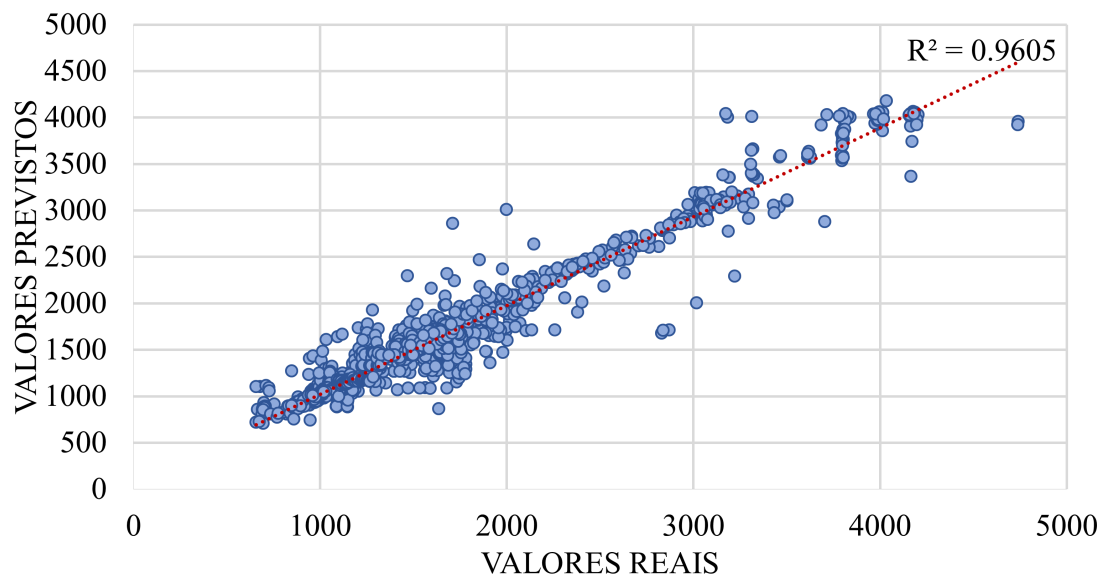
Tabela 7 – Resumo estatístico do modelo M5P

Resumo estatístico	
Coefficiente de correlação	0,9745
Erro absoluto médio	69,2642
Erro médio quadrático da raiz	139,9426
Erro absoluto relativo	0,1769
Erro quadrático relativo da raiz	0,2251
Número total de instâncias	11390

Fonte: Próprio autor (2020).

Por outro lado, o modelo é testado , e, a Figura 35 ilustra o gráfico de dispersão dos valores reais com os valores previstos obtidos com o modelo de previsão, demonstrando assim a relação entre esses valores.

Figura 35 – Diagrama de dispersão dos dados reais e os dados previstos do modelo M5P



Fonte: Próprio autor (2020)

- Regressão linear (LR):

A função básica obtida do modelo de regressão linear é:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_{25} x_{25} + \beta_{26} x_{26} + \beta_{27} x_{27}$$

e os valores β_i são:

β_0	β_1	β_2	β_3	β_4	β_5	β_6
-7,282e+03	1,438e+00	8,456e+01	-6,345e-01	1,555e+01	-5,042e+01	7,682e+01
β_7	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}
-3,732e+01	1,810e-01	-1,032e+01	-9,402e-01	1,985e-02	1,214e-01	1,126e+00
β_{14}	β_{15}	β_{16}	β_{17}	β_{18}	β_{19}	β_{20}
5,027e+00	1,512e+00	5,254e-01	1,555e+00	1,779e+00	2,586e+00	1,527e+00
β_{21}	β_{22}	β_{23}	β_{24}	β_{25}	β_{26}	β_{27}
-7,437e+00	-1,375e+01	5,333e-02	-4,947e-01	-4,951e-01	3,529e-01	6,560e-02

A Tabela 8 apresenta o resumo estatístico do modelo LR.

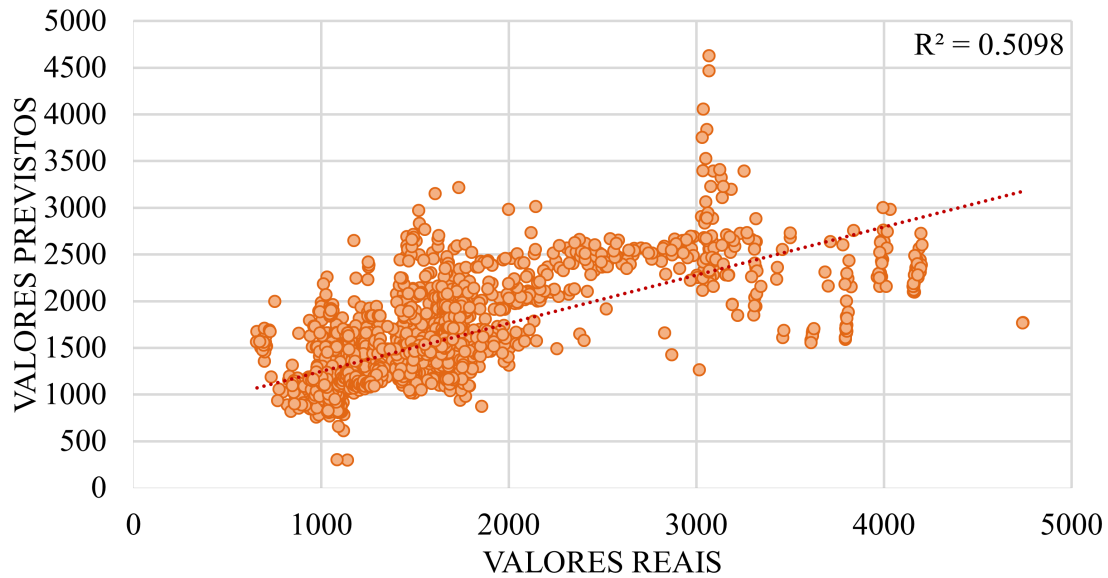
Tabela 8 – Resumo estatístico do modelo LR

Resumo estatístico	
Coefficiente de correlação	0,4821
Erro padrão residual	447,4
Valor p (< 0,001)	<2,2e-16

Fonte: Próprio autor (2020).

Além disso, o modelo LR é testado, e, na Figura 36 é apresentado o gráfico dos valores reais com os valores previstos obtidos com o modelo de previsão, demonstrando assim a relação baixa entre os valores.

Figura 36 – Diagrama de dispersão dos dados reais e os dados previstos do modelo LR

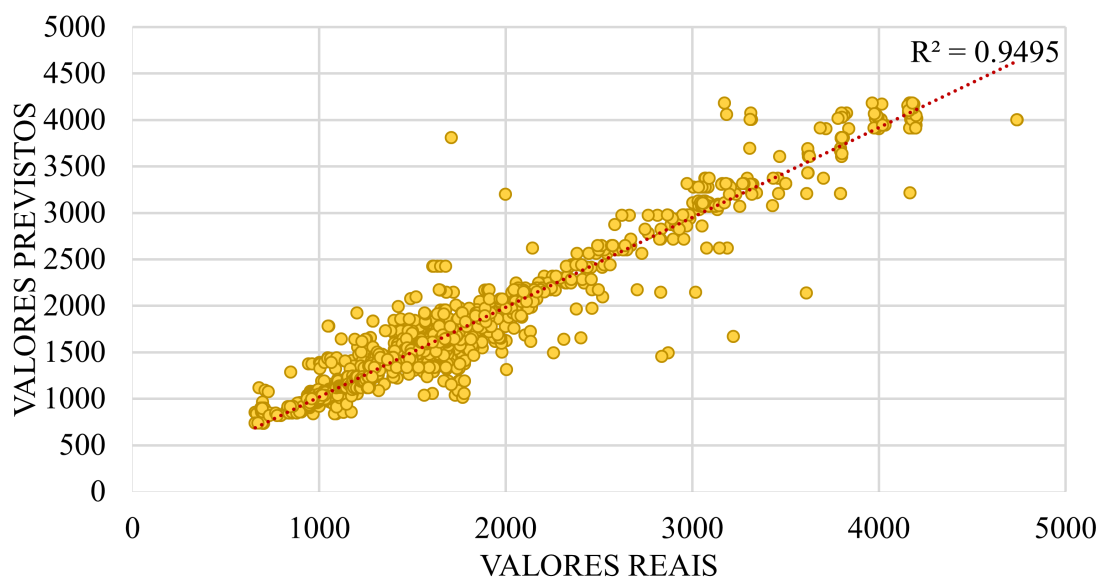


Fonte: Próprio autor (2020)

- Árvore de decisão CART:

O modelo CART possui uma estrutura de 597 folhas de previsão, o algoritmo cria um modelo com valores fixos nas folhas. A Figura 37 ilustra o gráfico de dispersão dos valores reais com os valores previstos do modelo, demonstrando assim a relação entre esses valores. Por outro lado, embora a relação mostrada no gráfico seja boa, o tamanho da árvore ainda é muito grande, o que leva a um sobreajuste.

Figura 37 – Diagrama de dispersão dos dados reais e os dados previstos do modelo CART

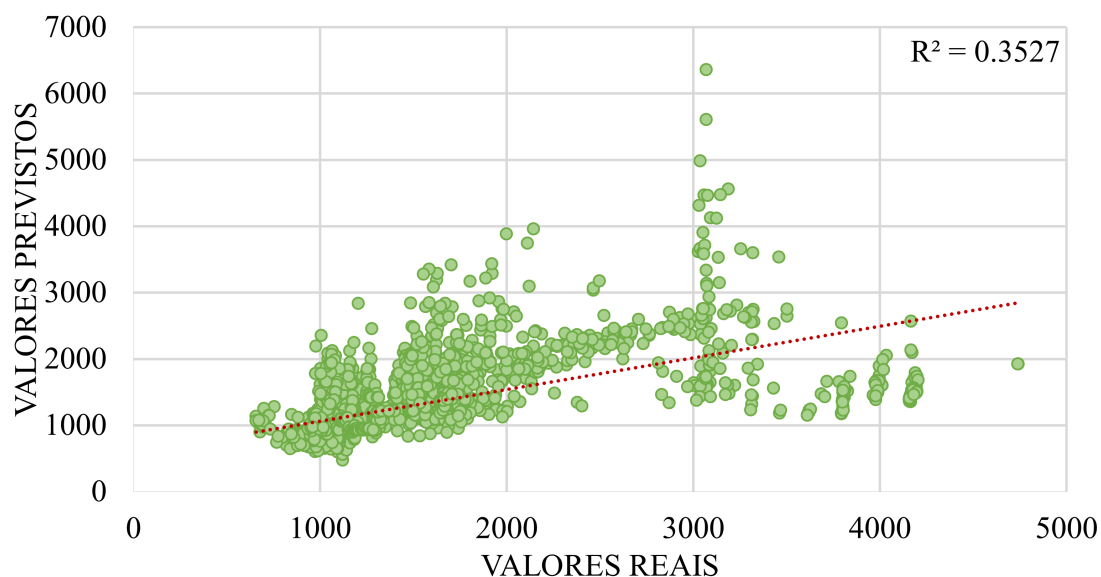


Fonte: Próprio autor (2020)

- Modelo de máquina de vetor de suporte:

O modelo SVM é do tipo regressão com um custo de 10, uma função de perda insensível igual 0,1 e com 1242 vectores de suporte. Depois de testar o modelo, na Figura 38 é apresentado o gráfico da relação entre os valores reais e os dados previstos do modelo, sendo possível observar que há um baixo relacionamento entre os valores.

Figura 38 – Diagrama de dispersão dos dados reais e os dados previstos do modelo SVM



Fonte: Próprio autor (2020)

- Modelo de Redes Neurais (NN):

O modelo NN têm uma estrutura de 27 neurônios na camada de entrada, duas camadas ocultas de 55 e 15 neurônios, na camada de saída um neurônio e uma perda de peso de 0,5. A Tabela 9 apresenta o resumo estatístico das métricas MAE, RMSE e R^2 , obtido com o *software* R.

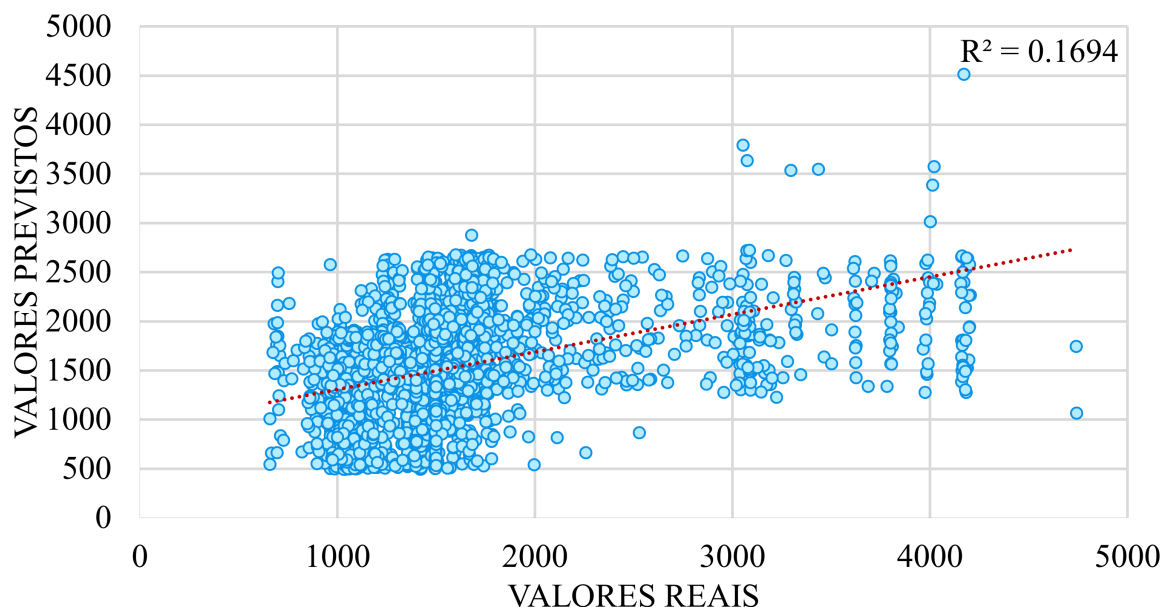
Tabela 9 – Resumo estatístico do modelo NN

Resumo estatístico	
Coefficiente de determinação	0,2868
Raiz do erro quadrático médio	470,1446
Erro médio absoluto	583,8121

Fonte: Próprio autor (2020).

Além disso, a Figura 39 apresenta o gráfico de dispersão dos valores reais e previstos, mostrando assim que o modelo têm um ajuste baixo e que há uma relação baixa entre os valores.

Figura 39 – Diagrama de dispersão dos dados reais e os dados previstos do modelo NN



Fonte: Próprio autor (2020)

Por outro lado, para demonstrar que a divisão de dados de 75%/25% gera um modelo de ML com os melhores resultados, o modelo M5P é selecionado para testar diferentes partições do conjunto de dados. A Tabela 10 apresenta o desempenho das diferentes divisões feitas com os dados para treinamento e teste para o modelo de previsão. Note-se, embora a variação entre os valores seja mínima, a divisão B é a melhor de acordo com as seis métricas, tendo os valores mais baixos nas medidas de erro (MAE, RMSE, RAE e RRSE) e os valores mais altos no coeficiente de determinação e correlação (R^2 e r). Imediatamente, a divisão D é o segundo melhor, seguido por E, C e A.

Tabela 10 – Desempenho das diferentes divisões do conjunto de dados

	Treinamento (%)	Teste (%)	MAE	RMSE	RAE	RRSE	r	R^2
A	70%	30%	72,27	138,28	0,189	0,0525	0,9734	0,94754
B	75%	25%	67,72	118,71	0,176	0,0395	0,9800	0,96049
C	80%	20%	68,65	124,97	0,178	0,0427	0,9784	0,95735
D	85%	15%	67,80	120,19	0,177	0,0398	0,9799	0,96024
E	90%	10%	67,75	120,72	0,177	0,0399	0,9799	0,96021

Fonte: Próprio autor (2020).

Deste jeito, a divisão B para a criação do modelo preditivo é a melhor divisão. Nesta sequência, usando o conjunto de dados de teste, a Tabela 11 mostra o análise das estatísticas básicas das previsões dos modelos testados. De acordo com os valores apresentados na tabela, pode-se observar que os valores estatísticos obtidos com o modelo M5P são os mais próximos dos valores reais, diferentemente dos valores calculados pelos outros modelos que apresentam maior diferença.

Tabela 11 – Estatísticas básicas das previsões dos modelos testados

	M5P	LR	CART	SVM	NN	Atual
Média	1501,76	1506,49	1504,21	1303,66	1499,16	1501,56
Desvio padrão	582,16	431,35	592,48	478,82	553,89	597,11
Mínimo	708,80	295,74	732,18	475,23	496,22	657,77
Quartil 1	1068,12	1153,21	1065,53	930,71	1066,22	1089,34
Quartil 2	1442,34	1463,07	1446,91	1216,84	1492,25	1420,80
Quartil 3	1652,79	1676,14	1641,21	1470,11	1849,79	1635,38
Máxima	4177,57	4624,38	4181,48	6356,04	4514,25	4742,26

Fonte: Próprio autor (2020).

Além disso, os modelos são avaliados por meio das métricas estatísticas de MAE, RMSE, RAE, RRSE, R^2 e r , que são utilizadas para analisar o desempenho e medir a taxa de erro das previsões, como apresentado na Tabela 12.

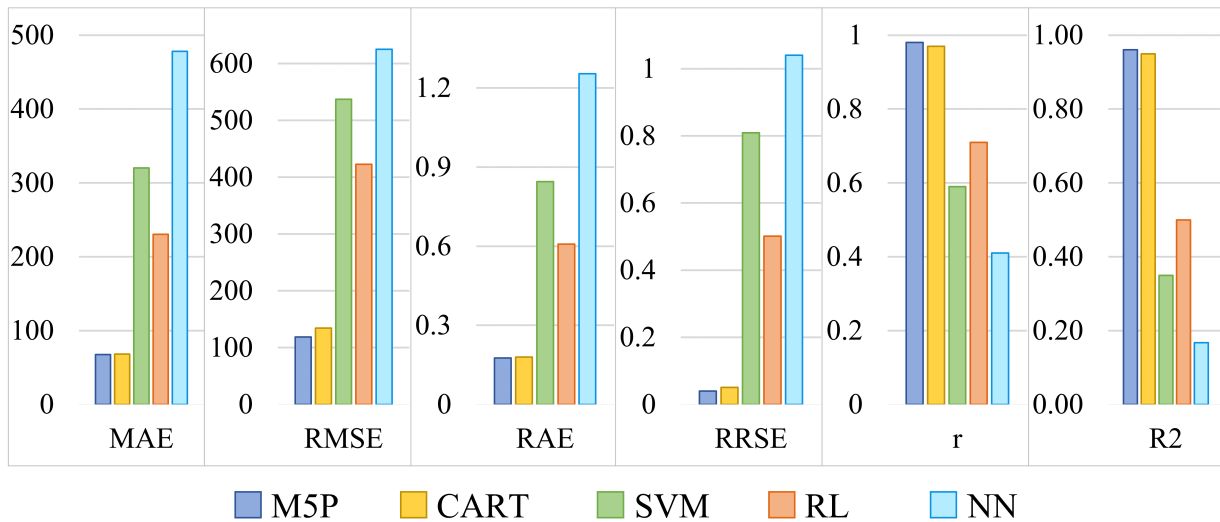
Tabela 12 – Métricas de avaliação dos modelos testados

	MAE	RMSE	RAE	RRSE	r	R^2
M5P	67,72	118,71	0,176	0,04	0,98	0,96
CART	68,09	134,57	0,180	0,05	0,97	0,95
SVM	320,25	537,04	0,845	0,81	0,59	0,35
LR	230,34	422,68	0,608	0,50	0,71	0,50
NN	477,99	625,28	1,254	1,04	0,41	0,17

Fonte: Próprio autor (2020).

Desse modo, a Figura 40 representa graficamente os valores das métricas de avaliação apresentadas na Tabela 12, com a finalidade de uma melhor visualização e comparação dos valores obtidos. Assim, constata-se que de acordo com as métricas avaliadas, os modelos de NN, SVM e LR, nessa ordem, têm os valores mais altos nas métricas de erro e os valores mais baixos na análise de r e R^2 . O modelo de NN é o que têm os piores valores nas métricas, indicando que é um modelo com baixo ajuste e baixa correlação entre os dados reais e previstos. O modelo de LR, em comparação dos modelos NN e SVM, obteve baixos valores nas métricas de erro, com grau de ajuste R^2 de 0,50 e r de 0,71, indicando a existência de uma correlação média entre os valores reais e os previstos pelo modelo.

Figura 40 – Valores das métricas de avaliação dos modelos testados

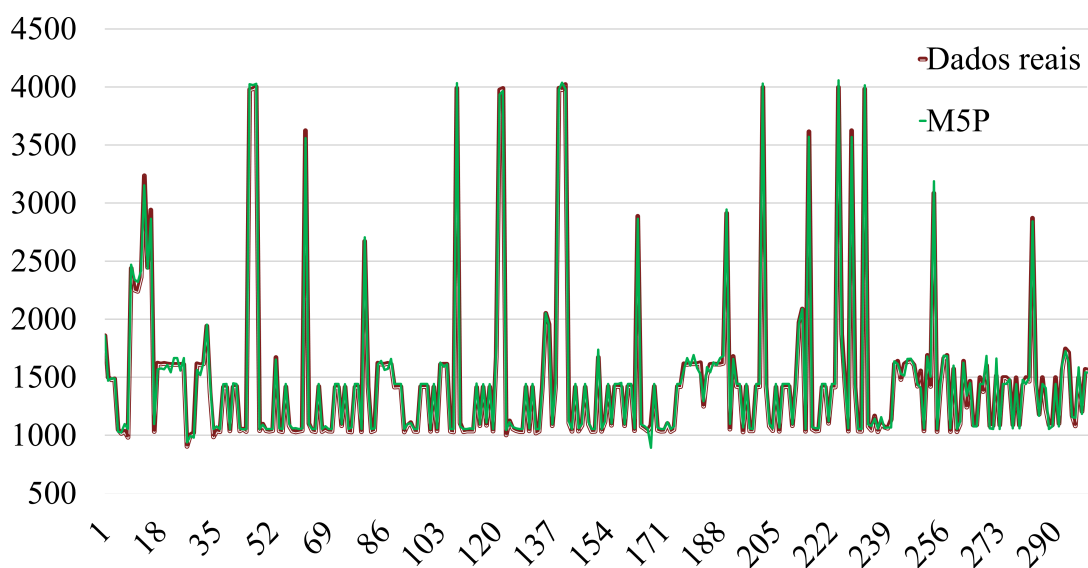


Fonte: Próprio autor (2020)

No entanto, embora os dados com o modelo LR tenham um desempenho considerável demonstrando o comportamento linear entre as variáveis de entrada e saída. Os modelos DT obtiveram melhores resultados nas taxas de erro, sendo quase metade dos obtidos com LR, por outro lado, o grau de ajuste obtido é muito próximo de 1. Portanto, embora nas métricas dos modelos M5P e CART, seus valores sejam muito próximos, o modelo com um ajuste maior é o M5P com as menores taxas de erro e um valor R^2 de 0,96 e r de 0,98, indicando uma forte correlação. Por outro lado, analisando la estructura de DT de ambos modelos, o CART possui um tamanho maior, com 597 folhas de previsão, indicando um sobreajuste, ao contrário do M5P, que possui apenas uma estrutura com 78 folhas.

Na Figura 41 é apresentado o gráfico de linhas com trezentos dados selecionados aleatoriamente, a fim de comparar os dados reais com os dados previstos com o modelo M5P. Nessa figura, é possível observar a proximidade dos dados reais dos dados previstos pelo modelo M5P.

Figura 41 – Gráfico de linhas de dados reais e previstos com M5P

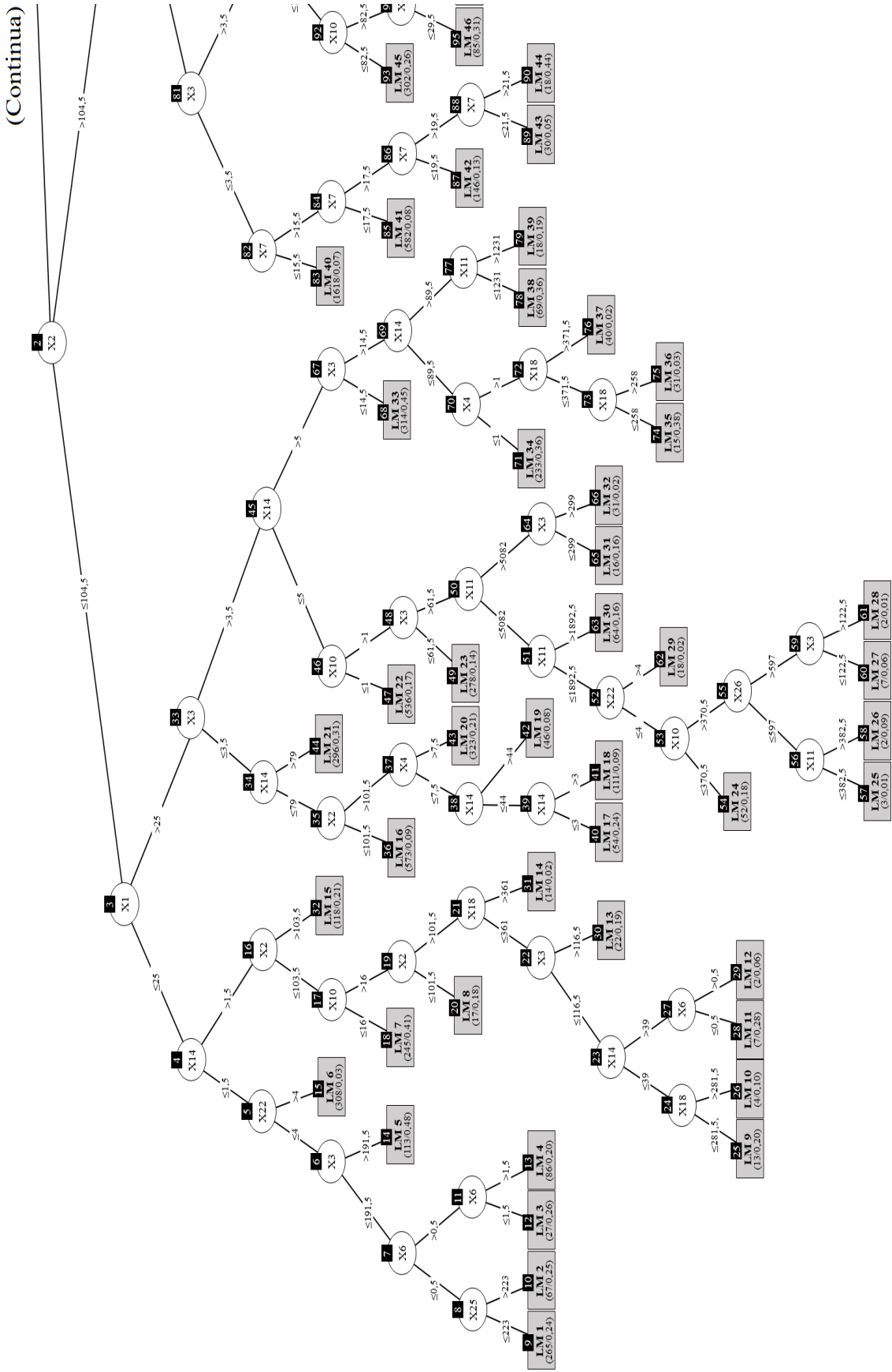


Fonte: Próprio autor (2020)

Por outro lado, conforme descrito na Seção 2.4, na SL os modelos de aprendizagem podem ser categorizados em modelos de caixa branca e caixa preta. Os modelos de DT são considerados de caixa branca, especificamente nos modelos criados com o algoritmo M5P, é possível analisar a estrutura interna do modelo e as funções geradas, sendo esta uma vantagem para a sistemática proposta; além de que para o desenvolvimento da próxima etapa, são muito importantes as funções de previsão geradas em cada uma das folhas do modelo.

Desse jeito, a Figura 42 é gerado com o *software* R, e ilustra a árvore construída com o algoritmo M5P (o Apêndice B apresenta o código da estrutura da árvore), com um total de 78 equações de regressão localizadas em cada uma das folhas de previsão. Além disso, no Apêndice C são apresentadas as 78 equações lineares das folhas do modelo. Na Figura 42, o termo LM mostrado significa modelo linear (*Linear model*) e é a folha de previsão que contém a equação.

Figura 42 – Árvore gerada com o algoritmo M5P



Fonte: R Core Team (2020)

ETAPA 3: OTIMIZAÇÃO

Na etapa de otimização, é utilizado o modelo de inferência gerado da árvore M5P e as equações de previsão do modelo. Para esse processo, é necessário modelar o problema com as restrições, portanto, a Tabela 13 mostra os valores que serão usados como restrições. Nesta tabela, pode-se observar que, embora a empresa tenha contrato com 6 empresas de transporte diferentes, essas transportadoras não realizam esse trabalho para todos os destinos (10), ou seja, por exemplo, as empresas 10, 20 e 40 transportam para o destino 107 e, as empresas 10, 20, 30 e 60 o transportam para o destino 103.

Tabela 13 – Dados para restrições

Destino	Empresas de transporte					
	x_2	x_1				
101	10	20	30	—	—	60
102	10	20	—	—	—	60
103	10	20	30	—	—	60
104	10	20	30	—	—	60
105	10	20	—	—	—	—
106	10	20	—	—	—	60
107	10	20	—	40	—	—
108	10	20	—	—	—	60
109	20	—	—	—	50	—
110	10	20	—	—	—	60

Fonte: Próprio autor (2020).

Por outro lado, vale a pena detalhar que nem sempre uma empresa de transporte leva itens da mesma linha para um destino. Por exemplo, a Tabela 14 mostra alguns dos embarques efetuados onde é possível observar que as mesmas linhas de produtos nem sempre são transportadas para um destino específico.

Tabela 14 – Detalhe de alguns itens enviados

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}
30	101								✓	✓																
30	101							✓				✓													✓	
30	101											✓														
30	101	✓							✓	✓			✓													
30	101	✓							✓		✓														✓	
30	101	✓						✓					✓													
30	101	✓							✓		✓															
30	101	✓										✓														
30	101	✓							✓	✓			✓		✓											
30	101	✓							✓	✓	✓															
20	101	✓							✓	✓															✓	
20	101	✓							✓				✓													
20	101	✓							✓				✓												✓	
20	101	✓							✓	✓																
20	101	✓							✓	✓	✓														✓	
20	101	✓							✓	✓			✓													
60	101	✓							✓	✓	✓		✓												✓	
60	101	✓							✓			✓													✓	
60	101	✓							✓	✓	✓		✓												✓	
60	101	✓							✓	✓	✓		✓		✓										✓	
60	101	✓							✓	✓	✓		✓		✓										✓	
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
60	101											✓													✓	

Fonte: Próprio autor (2020).

Desta forma, por meio do *software* R, com o pacote *Shiny*, é desenvolvida uma interface de usuário para tomada de decisão (Apêndice D, E e F). A lógica do processamento de dados da interface do usuário têm a estrutura de um processo automático, que integra o modelo de ML, as funções de previsão empírica e as restrições, o último mencionado é para o processo de busca de valores ótimos para a tomada de decisão.

No processo tradicional de busca dos valores obtidos, é necessário ter uma função objetivo (para minimizar ou maximizar) e as restrições que interferem no processo. Nesta sistemática, com o modelo de previsão do M5P são geradas “ n ” funções objetivo (equações de previsão), ou seja, dependendo de quais valores assumam algumas variáveis de entrada, a saída é calculada por uma equação diferente. Assim, o problema pode ser modelado conforme as expressões (3.3.1) e (3.3.2).

Função objetivo:

$$\min \quad f_h(x) = \beta_{h0} + \sum_{i=1}^d \beta_{hi}x_i \quad i = (1, 2, \dots, d), \quad h = (1, 2, \dots, n) \quad (3.3.1)$$

Sujeito às restrições::

$$\left\{ \begin{array}{l} a_1 \leq x_2 \leq a_{10} \\ (x_2 = a_k \Rightarrow x_1 \in Q_k) \\ x_i \geq 0 \\ x_i \text{ integer} \\ \forall(a_k) \text{ integer} \\ \mathbf{x} \in \mathbf{X} \end{array} \right. , a_k \neq Q_k, Q_k \in \Delta, k = (1, 2, \dots, 10) \quad (3.3.2)$$

Sendo $\mathbf{x} = (x_1, x_2, \dots, x_d)$ o conjunto de valores das variáveis de entrada, que pertencem ao vetor de decisão \mathbf{X} , e $f_h(x)$ é o conjunto de funções que foram geradas com o modelo M5P. O problema de estudo tem um único objetivo, que é minimizar o custo, e foram geradas 78 equações de acordo com a expressão (3.3.3).

$$f_h(x) = \{f_1(x), f_2(x), f_3(x), \dots, f_h(x)\} \quad , \quad h = (1, 2, \dots, 78) \quad (3.3.3)$$

Dado que o modelo possui 78 equações, a estimativa de um novo valor y é calculada por uma das n equações existentes, ou seja, dependendo dos valores assumidos pelo conjunto \mathbf{x} das variáveis de entrada, uma $f_h(x)$ calcula o novo \hat{y} . Em outras palavras, o valor de \hat{y} é estimado por meio de um processo de inferência de “Se ‘ \mathbf{x} ’ então ‘ $f_m(x)$ ’ caso contrário ‘ $f_{n \neq m}(x)$ ’”, como apresentado na expressão (3.3.4)

$$\left[\mathbf{x}_m \Rightarrow f_m(x) \right] \wedge \left[\mathbf{x}_{n \neq m} \Rightarrow f_{n \neq m}(x) \right] \quad , \quad \forall(m, n) \in h \quad (3.3.4)$$

Assim, De acordo com as expressões (3.3.1) e (3.3.2), o problema é um modelo de otimização discreta, porque cada iteração é baseada na seleção de um valor que pertence a um conjunto discreto de valores de uma variável com restrições. O x_2 é uma variável inteira, condicionada aos 10 destinos a a_k com $k = (1, 2, \dots, 10)$. No entanto, a variável x_1 está condicionado aos valores do vetor Q_k e ao valor

assumido por x_2 , sendo Q_k um dos elementos do conjunto de vetores Δ . O conjunto Δ é composto por 10 vetores Q_k que estão vinculados aos valores assumidos por a_k da variável x_2 , conforme apresentado na expressão (3.3.5).

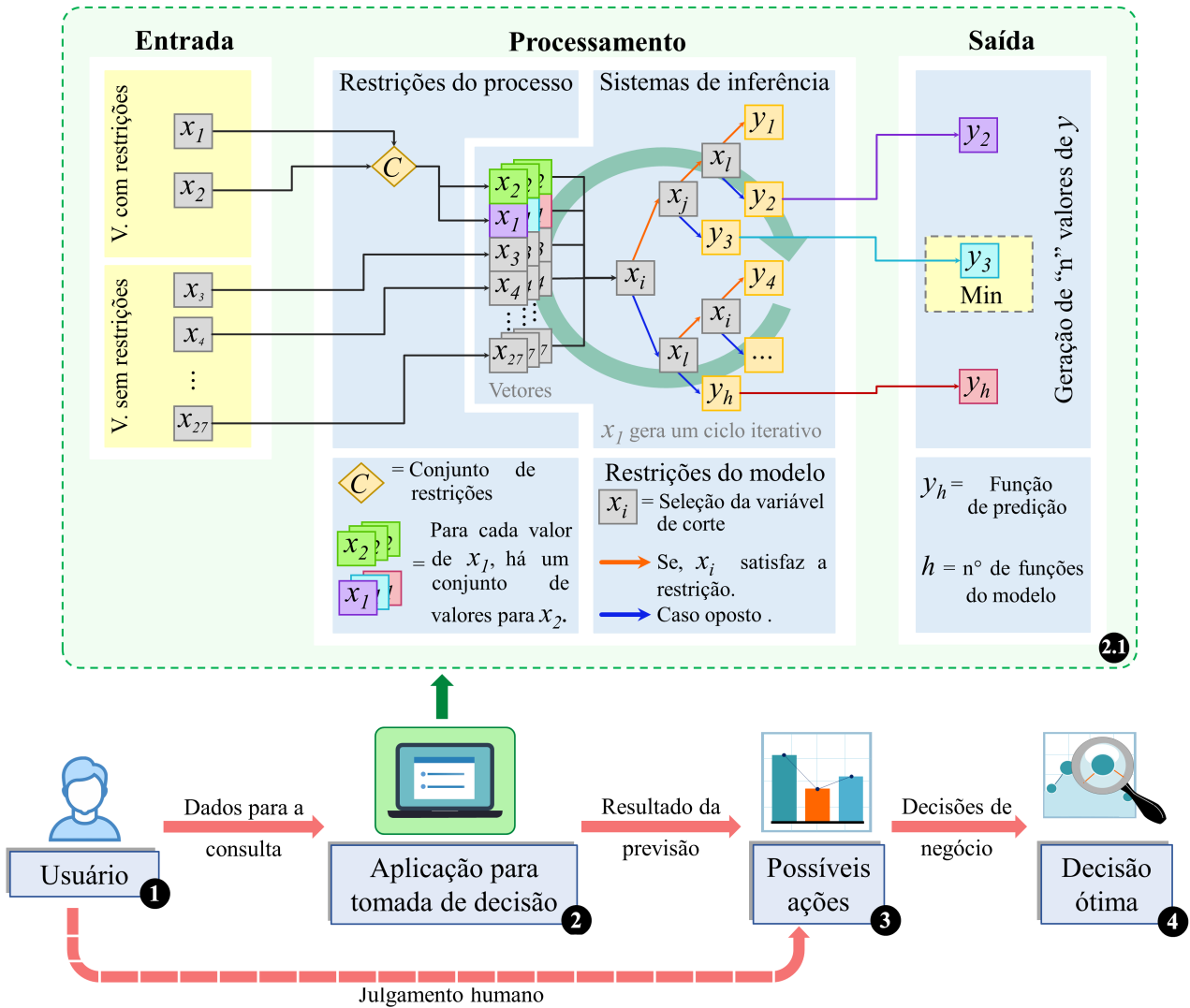
$$(x_2 = a_k \Rightarrow x_1 \in Q_k) \left\{ \begin{array}{l} x_2 = a_1 \Rightarrow x_1 = Q_1 \quad , \quad Q_1 = (q_{1,1}, q_{1,2}, q_{1,3}, q_{1,4}) \\ x_2 = a_2 \Rightarrow x_1 = Q_2 \quad , \quad Q_2 = (q_{2,1}, q_{2,2}, q_{2,3}) \\ x_2 = a_3 \Rightarrow x_1 = Q_3 \quad , \quad Q_3 = (q_{3,1}, q_{3,2}, q_{3,3}, q_{3,4}) \\ x_2 = a_4 \Rightarrow x_1 = Q_4 \quad , \quad Q_4 = (q_{4,1}, q_{4,2}, q_{4,3}, q_{4,4}) \\ x_2 = a_5 \Rightarrow x_1 = Q_5 \quad , \quad Q_5 = (q_{5,1}, q_{5,2}) \\ x_2 = a_6 \Rightarrow x_1 = Q_6 \quad , \quad Q_6 = (q_{6,1}, q_{6,2}, q_{6,3}) \\ x_2 = a_7 \Rightarrow x_1 = Q_7 \quad , \quad Q_7 = (q_{7,1}, q_{7,2}, q_{7,3}) \\ x_2 = a_8 \Rightarrow x_1 = Q_8 \quad , \quad Q_8 = (q_{8,1}, q_{8,2}, q_{8,3}) \\ x_2 = a_9 \Rightarrow x_1 = Q_9 \quad , \quad Q_9 = (q_{9,1}, q_{9,2}) \\ x_2 = a_{10} \Rightarrow x_1 = Q_{10} \quad , \quad Q_{10} = (q_{10,1}, q_{10,2}, q_{10,3}) \end{array} \right. \quad (3.3.5)$$

sendo q um elemento do vetor Q_k relacionados às restrições das transportadoras apresentadas na Tabela 13. Todo esse processo gera um ciclo iterativo que resulta em um conjunto de valores calculados de \hat{y}_h pelas funções empíricas do modelo M5P.

Por outro lado, tudo o que foi mencionado acima é modelado e integrado à interface do usuário por meio do pacote *Shiny* do R, a fim de obter uma visualização mais rápida da aplicação da sistemática ao problema aplicado. A Figura 3 da seção 1.3 resume em detalhes todo o processo interno do fluxo de dados da interface do usuário para tomada de decisão.

Assim, a Figura 3 é modificada e adaptada ao modelo da árvore de decisão, conforme apresentado na Figura 43, sendo o fluxo interno do aplicativo dividido nas três atividades principais de entrada, processamento e saída. Por conseguinte, para a busca dos valores ótimos, o usuário digita os dados necessários para realizar a consulta no aplicativo, dessa forma, os dados inseridos na interface são o destino (x_2) e as quantidades dos itens que serão enviados (x_3, x_4, x_5, \dots ou x_{27}). A variável x_1 é uma variável com restrição e corresponde às empresas de transporte, essa variável assume um valor calculado pelo sistema de tomada de decisão como o valor ideal para obter o custo mínimo de transporte.

Figura 43 – Diagrama do fluxo de dados no sistema de tomada de decisão com M5P

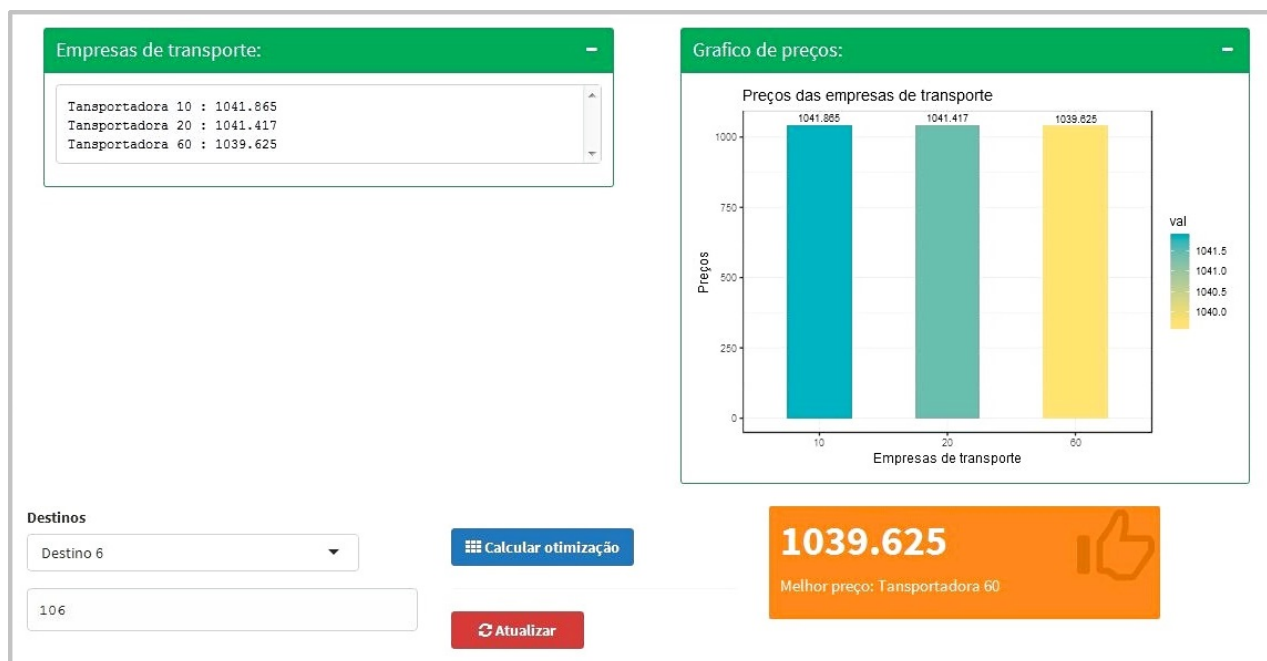


Fonte: Próprio autor (2020)

Por exemplo, considerando a Figura 43 como guia, a empresa precisa enviar produtos para o destino 106 ($x_2 = a_6$), e, envia as quantidades de 1021 da linha 7 (x_7), 2106 da linha 8 (x_8), 13212 da linha 9 (x_9), 16 da linha 11 (x_{11}) e 258 da linha 23 (x_{23}). O gerente da área sabe que, para esse destino, existem três empresas de transporte ($Q_6 = '10, 20 \text{ e } 60'$) que cobrem as rotas, mas, antes de solicitar que uma dessas empresas faça o trabalho, o aplicativo desenvolvido ajudará a ter uma visão mais ampla dos possíveis custos.

Nessa sequência, os dados mencionados acima são inseridos na interface desenvolvida, e, a previsão de custos para cada uma das empresas de transporte é obtida rapidamente, conforme a Figura 44. Na figura, o processo interno do aplicativo gera um ciclo iterativo em que uma previsão é avaliada com base nos possíveis valores que a variável x_1 pode assumir. Além disso, a interface exibe o valor ideal para a empresa e mostra graficamente cada um dos preços gerados no processo de busca pelo valor ideal; este último visa ajudar a identificar outras opções possíveis no caso de que a empresa de transporte com o valor ideal não esteja disponível.

Figura 44 – Resultado na interface de usuário



Fonte: R Core Team (2020)

Desta forma, o gerente da área pode basear suas decisões levando em consideração o resultado da previsão, concluindo assim a seleção de uma transportadora que gere um custo menor. A Tabela 15 apresenta o custo previsto de outros itens transportados, para serem comparados posteriormente com o custo real que a empresa pagou.

Tabela 15 – Custos obtidos com o aplicativo

Destino	Itens	Transportadora	Valor previsto
101	$x_3 = 56, x_{11} = 893, x_{12} = 110,$ $x_{14} = 18$ e $x_{26} = 100$	10	1489,02
		20	1335,58
		30	1298,96
		60	1420,35
110	$x_3 = 87, x_5 = 110,$ $x_{19} = 59$ e $x_{24} = 70$	10	3648,17
		20	3703,00
		60	3922,34

Fonte: Próprio autor (2020).

Como dados reais tem-se que as empresas de transporte contratadas são a transportadora 60 com um custo de 1409,42 para o destino 101 e a transportadora 20 com um custo de 3714,62 para o destino 110. Nesta continuidade, considerando os dados mostrados na Tabela 15 e os custos reais, a Tabela 16 apresenta uma comparação entre o preço real pago e o custo obtido com o aplicativo para a tomada de decisão.

Tabela 16 – Comparação de custos

Destino	Transportadora	Valor previsto	Valor real	Diferença
101	10	1489,02	-	-
	20	1335,58	-	73,84
	30	1298,96	-	110,46
	60	1420,35	1409,42	-
110	10	3648,17	-	66,45
	20	3703,00	3714,62	-
	60	3922,34	-	-

Fonte: Próprio autor (2020).

Na Tabela 16, para o destino 101 as transportadoras 20 e 30 gerariam um custo menor e para o destino 110 a transportadora 10, gerando assim uma diferença considerável com o valor real. Esses custos, a longo prazo, se destaca em um valor mais considerável quando é adicionado a outras diferenças monetárias por um período mais longo de tempo.

Dessa forma, o que é mostrado na Tabela 15 e 16 ajuda a validar o modelo desenvolvido, mostrando alguns dos pagamentos feitos pela empresa pelo transporte de produtos para um destino e comparando-os com os valores obtidos com a interface. Assim, pode-se destacar que existe uma variação entre os possíveis custos gerados por cada transportador, e que a ferramenta apresenta uma contribuição modesta para a tomada de decisão. Em outras palavras, o sistema de tomada de decisão têm a vantagem de fornecer previsões que serão avaliadas pelo especialista em um curto espaço de tempo, o que resulta em transações mais eficientes e eficazes, e que, em períodos de maior tempo, gera economia considerável para a empresa. Os resultados apresentados no estudo apontam que o modelo têm uma grande capacidade para prever o objetivo, então, suas previsões ficarão muito próximas dos valores reais.

Finalmente, a integração do ML como ferramenta de apoio à tomada de decisão gerou bons resultados e tornou-se uma ferramenta eficiente, embora não haja muitas aplicações na literatura para problemas de tomada de decisão na área de logística ou cadeia de suprimentos. Por outro lado, embora atualmente para a empresa onde a sistemática proposta foi aplicada, sua realidade atual na tomada de decisões não se baseia no julgamento de uma previsão de um modelo inteligente, o especialista responsável pela área acredita que o modelo proposto poderia gerar uma vantagem competitiva, pois pode ser usado como um instrumento que ajuda a identificar que existem outras soluções possíveis para um problema.

4 CONCLUSÕES E RECOMENDAÇÕES PARA FUTURAS PESQUISAS

O processo de tomada de decisão é muito importante para uma empresa, porque essas decisões podem levar a organização ao seu sucesso ou fracasso. Assim, com o avanço da tecnologia, as empresas industriais viram a necessidade de aplicar algumas das técnicas emergentes aos seus processos; por exemplo, para tomar a decisão de comprar ou produzir, quem contratar, quanto investir, qual opção gera menos custos, etc.

Nesse sentido, o presente trabalho descreve uma abordagem inovadora para modelar um problema de tomada de decisão, que visa usar técnicas de ML, especificamente foi usado o modelo M5P para obter as funções de previsão, que serão utilizadas em um processo iterativo de otimização para selecionar os melhores valores que satisfaçam um conjunto de restrições e gerem um resultado ótimo. Além disso, a metodologia desta pesquisa consistiu em três etapas principais (exploração de problemas, modelo de ML e otimização), e cada etapa é composta de um conjunto de tarefas que podem ser aplicadas a diferentes problemas de tomada de decisão em uma organização. No entanto, de acordo com as pesquisas realizadas na base de dados de Scopus, não foram encontrados artigos publicados que aplicam a sistemática proposta e que utiliza técnicas de ML, como o algoritmo M5P, a problemas de otimização para tomada de decisão em processos industriais.

Por outro lado, embora, para a segurança da empresa, houvesse várias limitações quanto à disponibilidade do conjunto de dados utilizado para aplicar o sistema, de acordo com especialistas e gerentes da empresa, Os resultados obtidos são satisfatórios e consistentes quando são comparados a uma situação real (tabela de custos reais e os custos gerados pela interface).

Passa-se a responder às questões de pesquisa, apresentados na Seção 1.2:

- Como implementar um modelo de previsão para a tomada de decisão utilizando técnicas de SL? Foi descrito, neste trabalho, o procedimento para implementação de um modelo de ML, O mais importante em todo o processo é a robustez dos dados que serão usados para o modelo preditivo, pois depende deles que o modelo final seja eficiente ou não. As técnicas de SL, para a previsão de resultados contínuos, têm como base principal, em muitos casos, o método tradicional de regressão, sendo uma vantagem para o desenvolvimento e implementação de um modelo preditivo. Por outro lado, existem diferentes ferramentas computacionais que podem ser usadas para o desenvolvimento de um modelo de ML, um dos quais é o *software* R, oferece uma ampla variedade de técnicas estatísticas. Dessa forma, foi possível atender ao primeiro objetivo, pois, neste trabalho, foi utilizada uma técnica SL de caixa branca para implementar o modelo de previsão.
- Que técnica de SL utilizar para desenvolver funções empíricas? Também atendendo ao segundo objetivo, foi utilizada a técnica de árvore de decisão M5P, porque esse modelo gera uma estrutura de uma árvore, na qual as previsões são geradas por equações lineares contidas em suas folhas finais. Essa técnica, provou ser eficiente quando comparada a outras técnicas de SL e, também gerou uma estrutura mais simples quando comparada à técnica de árvore de decisão CART.

- Como modelar o processo de otimização considerando as equações do modelo de previsão? Aproveitando as funções de previsão geradas com a árvore de decisão do M5P e considerando as restrições de algumas das variáveis, foi possível modelar o processo de otimização. Esse processo de integração das restrições que envolve algumas das variáveis estudadas com o modelo de previsão criado foi realizado no pacote *shiny* do *software* R, com o desenvolvimento de uma interface de usuário. Internamente, no aplicativo é gerado um ciclo iterativo, no qual as restrições estipuladas devem ser cumpridas, o que resulta em possíveis valores para as variáveis de decisão, que levará a uma melhor análise para a tomada de decisão. Dessa forma, o terceiro e o quarto conjunto de objetivos também foram cumpridos.
- Como validar os resultados? Atendendo aos dois últimos objetivos, a sistemática proposto foi validado comparando-se o valor pago pela empresa pelo transporte de alguns itens com os custos avaliados pelo modelo proposto. Por conseguinte, e de acordo com o especialista, embora os sistemas inteligentes para tomada de decisão não sejam uma realidade da empresa, o modelo proposto é uma contribuição para a área de estudo e pode gerar uma vantagem competitiva por ser utilizado como um instrumento de apoio à tomada de decisão. Gerando, que o usuário possa avaliar outros custos em pouco tempo, apenas inserindo alguns dados necessários na interface.

De maneira destacada, vale ressaltar que o uso de modelos de previsão para tomada de decisão gera uma vantagem competitiva que gera eficiência e eficácia em uma organização, pois, com esta ferramenta é possível avaliar rapidamente possíveis cenários e ser capaz de tomar melhores decisões, analisando cada um dos resultados obtidos. Sendo assim, a sistemática foi aplicada a uma empresa do segmento de autopeças, especificamente com os dados da área de logística, a fim de validar a abordagem proposta. Além disso, no processo de criação e teste do modelo ML foram utilizadas técnicas estatísticas para sua avaliação, essas técnicas estatísticas aplicadas validaram o modelo de previsão demonstrando que o modelo com M5P é eficiente, gerando também taxas de erro mais baixas do que os outros modelos avaliados.

Finalmente, vale a pena ressaltar que a aplicação de técnicas de ML para análise e tomada de decisão em processos industriais não se destina a substituir pessoas, já que seu objetivo principal é ser uma ferramenta de suporte, trazendo vantagens consideráveis que ajudam o gerente ou analista a tomar melhores decisões em menos tempo. No entanto, a abordagem descrita aqui pode ser aplicada a diferentes problemas de tomada de decisão nas organizações, e, é muito importante que na etapa de criação do modelo ML com M5P sejam utilizados dados de qualidade, pois bem, o poder preditivo e a precisão do modelo dependem desses dados.

Complementando, como propostas para pesquisas futuras sugere-se testar a sistemática com um modelo de otimização com variáveis contínuas, por outro lado, uma segunda proposta é utilizar a sistemática proposta com lógica fuzzy para tomada de decisão.

4.1 PROPOSTA PARA FUTURAS PESQUISAS

- Aplicar o sistemática proposta a outros tipos de negócios.
- Aplicar o sistema a outros tipos de problemas, como, tomada de decisão na área de recursos humanos, finanças, mesmo na mesma área, mas outros problemas.
- Sugere-se testar a sistemática com um modelo de otimização com variáveis contínuas.
- Para o modelo de aprendizagem de máquina, sugere-se testar o algoritmo de árvore de decisão *fuzzy*.
- Para a otimização, sugere-se tentar modelar dados estocásticos avaliados com Monte Carlos para obter os valores ótimos.

REFERÊNCIAS

- ACHARYA, A. et al. Big data, knowledge co-creation and decision making in fashion industry. **International Journal of Information Management**, v. 42, n. May, p. 90–101, oct 2018.
- ADETILOYE, T.; AWASTHI, A. **Chapter 8 - Predicting Short-Term Congested Traffic Flow on Urban Motorway Networks**. 1. ed. [S.l.]: Academic Press, 2017. 145 - 165 p. ISBN 978-0-12-811318-9.
- AGARWAL, R.; DHAR, V. Editorial big data, data science, and analytics: The opportunity and challenge for is research. **Information Systems Research**, v. 25, p. 443–448, 09 2014.
- AKTER, S. et al. Analytics-based decision-making for service systems: A qualitative study and agenda for future research. **International Journal of Information Management**, v. 48, p. 85 – 95, 2019. ISSN 0268-4012.
- AL-JARRAH, O. Y. et al. Efficient machine learning for big data: A review. **Big Data Research**, v. 2, n. 3, p. 87 – 93, 2015. ISSN 2214-5796. Big Data, Analytics, and High-Performance Computing.
- ALPAYDIM, E. **Introduction to Machine Learning**. Cambridge, Massachusetts, London, England: The MIT Press, 2004. 1-433 p. ISBN 0-262-01211-1.
- AURUM, A.; WOHLIN, C. The fundamental nature of requirements engineering activities as a decision-making process. **Information and Software Technology**, v. 45, n. 14, p. 945–954, 2003. ISSN 0950-5849.
- BAG, S. et al. Big data analytics as an operational excellence approach to enhance sustainable supply chain performance. **Resources, Conservation and Recycling**, v. 153, p. 104559, 2020. ISSN 0921-3449.
- BALI, R. et al. **R: Unleash Machine Learning Techniques**. 35 Livery Street, Birmingham B3 2PB, UK.: Packt Publishing Ltd., 2016. 1-1123 p. ISBN 978-1-78712-734-0.
- BASHEER, I.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods**, v. 43, n. 1, p. 3 – 31, 2000. ISSN 0167-7012. Neural Computing in Micrbiology.
- BAZARGAN-LARI, M. R. Real-time response to contamination emergencies of urban water networks. **Iranian Journal of Science and Technology, Transactions of Civil Engineering**, v. 42, n. 1, p. 73–83, Mar 2018. ISSN 2364-1843.
- BELL, E.; BRYMAN, A. The Ethics of Management Research: An Exploratory Content Analysis. **British Journal of Management**, v. 18, n. 1, p. 63–77, mar 2007. ISSN 1045-3172.
- BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modeling. **International Journal of Operations & Production Management**, v. 22, n. 2, p. 241–264, feb 2002. ISSN 0144-3577.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. Cambridge, U.K.: Springer Science and Business Media, LLC, 2006. 1-749 p. ISBN 0-387-31073-8.
- BOHANEK, M.; ROBNIK-SIKONJA, M.; BORSTNAR, M. K. Decision-making framework with double-loop learning through interpretable black-box machine learning models. **Industrial Management and Data Systems**, v. 117, n. 7, p. 1389–1406, 2017. ISSN 0263-5577.

BRADLEY, A.; MEER, R. van der; MCKAY, C. Personalized pancreatic cancer management: A systematic review of how machine learning is supporting decision-making. **Pancreas**, v. 48, n. 5, p. 598–604, 2019. ISSN 0885-3177.

BREIMAN, L. et al. **Classification and Regression Trees**. [S.l.]: Chapman and Hall/CRC, 1998. 1-366 p. ISBN 0-412-04841-8.

BUSONI, L.; BABUSKA, R.; SCHUTTER, B. D. A comprehensive survey of multiagent reinforcement learning. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, v. 38, n. 2, p. 156–172, March 2008. ISSN 1094-6977.

BUTZ, M. V.; SIGAUD, O.; GERARD, P. Anticipatory behavior: Exploiting knowledge about the future to improve current behavior. In: **Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems**. [S.l.]: Springer, Berlin, Germany, 2003. p. 1–10. ISBN 3540404295. ISSN 0302-9743.

CANTAMESSA, M. An empirical perspective upon design research. **Journal of Engineering Design**, v. 14, n. 1, p. 1–15, mar 2003. ISSN 0954-4828.

CARRASCO, I. G. et al. Automatic detection of relationships between banking operations using machine learning. **Information Sciences**, v. 485, p. 319 – 346, 2019. ISSN 0020-0255.

CASTILLO, P. A. et al. Applying computational intelligence methods for predicting the sales of newly published books in a real editorial business management environment. **Knowledge-Based Systems**, v. 115, p. 133 – 151, 2017. ISSN 0950-7051.

CHANG, C.-C.; LIN, C.-J. LIBSVM: A Library for Support Vector Machines. **ACM Trans. Intell. Syst. Technol.**, v. 2, may 2011. ISSN 2157-6904.

CHEN, C. P.; ZHANG, C.-Y. Data-intensive applications, challenges, techniques and technologies: A survey on big data. **Information Sciences**, v. 275, p. 314 – 347, 2014. ISSN 0020-0255.

CHO, S. et al. Tool breakage detection using support vector machine learning in a milling process. **International Journal of Machine Tools and Manufacture**, v. 45, n. 3, p. 241 – 249, 2005. ISSN 0890-6955.

CHOI, T.-M.; CHENG, T. C. E.; ZHAO, X. Multi-Methodological Research in Operations Management. **Production and Operations Management**, v. 25, n. 3, p. 379–389, mar 2016. ISSN 10591478.

CHONG, E. K. P. **An Introduction to Optimization**. Second edi. [S.l.]: John Wiley and Sons, 1996. 1-495 p. ISBN 9780471758006.

CHOU, P. A. Optimal partitioning for classification and regression trees. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 13, n. 4, p. 340–354, April 1991. ISSN 0162-8828.

CHOUGRAD, H.; ZOUAKI, H.; ALHEYANE, O. Deep convolutional neural networks for breast cancer screening. **Computer Methods and Programs in Biomedicine**, v. 157, p. 19 – 30, 2018. ISSN 0169-2607.

CRISCI, C.; GHATTAS, B.; PERERA, G. A review of supervised machine learning algorithms and their applications to ecological data. **Ecological Modelling**, v. 240, p. 113 – 122, 2012. ISSN 0304-3800.

CUI, L. Complex industrial automation data stream mining algorithm based on random internet of robotic things. **Automatika**, Taylor and Francis, v. 60, n. 5, p. 570–579, 2019.

- DENG, L.; YU, D. Deep learning: Methods and applications. v. 7, n. 3-4, p. 197–387, 2013.
- DORMANN, C. F. et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. **Ecography**, v. 36, n. 1, p. 27–46, 2013.
- ECK, N. J. V.; WALTMAN, L. Text mining and visualization using VOSviewer. **ISSI Newsletter**, v. 7, n. 3, p. 50–54, 2011. ISSN 1932-6203.
- ETEMAD-SHAHIDI, A.; MAHJOOBI, J. Comparison between m5p model tree and neural networks for prediction of significant wave height in lake superior. **Ocean Engineering**, v. 36, n. 15, p. 1175 – 1181, 2009. ISSN 0029-8018.
- FAVARETTO, M.; CLERCQ, E. D.; BERNICE, S. E. Big data and discrimination: perils, promises and solutions. a systematic review. **Journal of Big Data**, v. 6, n. 1, p. 12, Feb 2019. ISSN 2196-1115.
- FENG, S.; ZHOU, H.; DONG, H. Using deep neural network with small dataset to predict material defects. **Materials and Design**, v. 162, p. 300 – 310, 2019. ISSN 0264-1275.
- FERNÁNDEZ-DELGADO, M. et al. An extensive experimental survey of regression methods. **Neural Networks**, v. 111, p. 11 – 34, 2019. ISSN 0893-6080.
- FIGUEIREDO, M. A. Adaptive sparseness for supervised learning. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 25, n. 9, p. 1150–1159, Sep. 2003. ISSN 0162-8828.
- FILIPIC, B.; JUNKAR, M. Using inductive machine learning to support decision making in machining processes. **Computers in Industry**, v. 43, n. 1, p. 31 – 41, 2000. ISSN 0166-3615.
- FILIPPINI, R. Operations management research: some reflections on evolution, models and empirical studies in OM. **International Journal of Operations & Production Management**, v. 17, n. 7, p. 655–670, jul 1997. ISSN 0144-3577.
- FLATH, C. M.; STEIN, N. Towards a data science toolbox for industrial analytics applications. **Computers in Industry**, v. 94, p. 16 – 25, 2018. ISSN 0166-3615.
- FORBES-INSIGHTS; PENSKE. Logistics, supply chain and transportation 2023: Change at breakneck speed. In: . FORBES INSIGHTS, 2018. p. 11–12. ISBN 212-367-2662. Disponível em: <www.forbes.com/forbes-insights/our-work/transportation-2023/>.
- FOURNIER, D.; CRÉMILLEUX, B. A quality index for decision tree pruning. **Knowledge-Based Systems**, v. 15, n. 1, p. 37–43, 2002. ISSN 0950-7051.
- FRENCH, S.; MAULE, J.; PAPAMICHAIL, N. **Decision Behaviour, Analysis and Support**. Cambridge, United Kingdom: Cambridge University Press, 2009. ISBN 9780511609947.
- FUSE, K. et al. Patient-based prediction algorithm of relapse after allo-hsct for acute leukemia and its usefulness in the decision-making process using a machine learning approach. **Cancer Medicine**, v. 8, p. 5058– 5067, 2019.
- GAMARRA, C.; GUERRERO, J. M.; MONTERO, E. A knowledge discovery in databases approach for industrial microgrid planning. **Renewable and Sustainable Energy Reviews**, v. 60, p. 615 – 630, 2016. ISSN 1364-0321.
- GANDOMI, A.; HAIDER, M. Beyond the hype: Big data concepts, methods, and analytics. **International Journal of Information Management**, v. 35, n. 2, p. 137 – 144, 2015. ISSN 0268-4012.

- GARTNER INC. **Gartner's Top 10 Strategic Technology Trends for 2017**. 2016. Disponível em: <www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017>.
- GARTNER INC. **Gartner's Top 10 Strategic Technology Trends for 2018**. 2017. Disponível em: <www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2018>.
- GE, Z. et al. Data mining and analytics in the process industry: The role of machine learning. **IEEE Access**, v. 5, p. 20590–20616, 2017. ISSN 2169-3536.
- GENRIKHOV, I. E.; DJUKOVA, E. V.; ZHURAVLEV, V. I. On full regression decision trees. **Pattern Recognition and Image Analysis**, v. 27, n. 1, p. 1–7, Jan 2017. ISSN 1555-6212.
- GENRIKHOV, I. E.; DJUKOVA, E. V.; ZHURAVLEV, V. I. On full regression decision trees. **Pattern Recognition and Image Analysis**, v. 27, p. 1–7, 2017. ISSN 1555-6212.
- GHAHRAMANI, Z. **Unsupervised Learning**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. 72-112 p. ISBN 978-3-540-28650-9.
- GOLDSTEIN, B. A.; NAVAR, A. M.; CARTER, R. E. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. **European Heart Journal**, v. 38, n. 23, p. 1805–1814, 07 2016. ISSN 0195-668X.
- GROMPING, U. Variable importance in regression models. **Wiley Interdisciplinary Reviews: Computational Statistics**, v. 7, n. 2, p. 137–152, 2015.
- GULSHAN, V. et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. **JAMA**, v. 316, n. 22, p. 2402–2410, 12 2016. ISSN 0098-7484.
- HAZEN, B. T. et al. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. **International Journal of Production Economics**, v. 154, p. 72–80, 2014. ISSN 0925-5273.
- HAZEN, B. T. et al. Back in business: operations research in support of big data analytics for operations and supply chain management. **Annals of Operations Research**, v. 270, n. 1, p. 201–211, Nov 2018. ISSN 1572-9338.
- HENNING, K.; WOLF-DIETER, L.; WOLFGANG, W. Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. **VDI nachrichten**, v. 13, n. 1, abr 2011.
- HERMANN, M.; PENTEK, T.; OTTO, B. Design Principles for Industrie 4.0 Scenarios. In: **2016 49th Hawaii International Conference on System Sciences (HICSS)**. [S.l.]: IEEE, 2016. v. 2016-March, p. 3928–3937. ISBN 978-0-7695-5670-3. ISSN 15301605.
- Huang, G. et al. Semi-supervised and unsupervised extreme learning machines. **IEEE Transactions on Cybernetics**, v. 44, n. 12, p. 2405–2417, Dec 2014. ISSN 2168-2267.
- JAIN, A. K. Data clustering: 50 years beyond k-means. **Pattern Recognition Letters**, v. 31, n. 8, p. 651 – 666, 2010. ISSN 0167-8655.
- JIANG, J. et al. Learning and inference in knowledge-based probabilistic model for medical diagnosis. **Knowledge-Based Systems**, v. 138, p. 58 – 68, 2017. ISSN 0950-7051.
- JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. **Science**, v. 349, p. 255–260, 2015. ISSN 1095-9203.

JUNG, N. chung. **Eco-hydraulic Modelling of Eutrophication for Reservoir Management**. [S.l.]: CRC Press, 2009. 1-247 p. ISBN 978-0-415-57382-5,0415573823.

KAELBLING, L. P.; LITTMAN, M. L.; MOORE, A. W. Reinforcement learning: A survey. **Journal of Artificial Intelligence Research**, v. 4, p. 237–285, 1996. ISSN 1076 - 9757.

KARABADJI, N. E. I. et al. An evolutionary scheme for decision tree construction. **Knowledge-Based Systems**, v. 119, p. 166 – 177, 2017. ISSN 0950-7051.

KARTAL, H. et al. An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. **Computers and Industrial Engineering**, v. 101, p. 599 – 613, 2016. ISSN 0360-8352.

KESHTEGAR, B.; KISI, O. Rm5tree: Radial basis m5 model tree for accurate structural reliability analysis. **Reliability Engineering and System Safety**, v. 180, p. 49 – 61, 2018. ISSN 0951-8320.

KOLBIN, V. **Decision making and programming**. St Petersburg University, Russia: World Scientific, 2003. 1-746 p. ISBN 9812383794.

KONONENKO, I.; KUKAR, M. **Machine Learning and Data Mining: Introduction to Principles and Algorithms**. Coll House, Westergate, Chichester, West Sussex, UK: Woodhead Publishing, 2007. 1-475 p. ISBN 190427521-4.

KOTU, V.; DESHPANDE, B. Chapter 8 - model evaluation. In: KOTU, V.; DESHPANDE, B. (Ed.). **Data Science (Second Edition)**. Second edition. [S.l.]: Morgan Kaufmann, 2019. p. 263 – 279. ISBN 978-0-12-814761-0.

KRAWCZYK, B. et al. Ensemble learning for data stream analysis: A survey. **Information Fusion**, v. 37, p. 132 – 156, 2017. ISSN 1566-2535.

KWOK, L.; YU, B. Spreading social media messages on facebook: An analysis of restaurant business to consumer communications. **Cornell Hospitality Quarterly**, v. 54, n. 1, p. 84–94, 2013.

LANGLEY, P.; SIMON, H. A. Applications of machine learning and rule induction. **Communications of the ACM**, v. 38, p. 54–64, 1995. ISSN 0001-0782.

LANTZ, B. **Machine Learning with R**. Second edition. 35 Livery Street, Birmingham B3 2PB, UK.: Packt Publishing Ltd., 2015. 1-452 p. ISBN 978-1-78439-390-8.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, p. 436–444, 2015. ISSN 1476-4687.

LEE, H. et al. Feature selection practice for unsupervised learning of credit card fraud detection. v. 96, n. 2, p. 408–417, 2018.

LI, G.; SHI, J. On comparing three artificial neural networks for wind speed forecasting. **Applied Energy**, v. 87, n. 7, p. 2313 – 2320, 2010. ISSN 0306-2619.

LI, W. et al. Transfer learning-based default prediction model for consumer credit in china. **The Journal of Supercomputing**, v. 75, n. 2, p. 862–884, 2019. ISSN 1573-0484.

LI, X.; ZHAO, H.; ZHU, W. A cost sensitive decision tree algorithm with two adaptive mechanisms. **Knowledge-Based Systems**, v. 88, p. 24–33, 2015. ISSN 0950-7051.

LOYER, J.-L. et al. Comparison of machine learning methods applied to the estimation of manufacturing cost of jet engine components. **International Journal of Production Economics**, v. 178, p. 109 – 119, 2016. ISSN 0925-5273.

- LU, Y. Industry 4.0: A survey on technologies, applications and open research issues. **Journal of Industrial Information Integration**, v. 6, p. 1–10, 2017. ISSN 2452-414X.
- MANDAL, S. The influence of big data analytics management capabilities on supply chain preparedness, alertness and agility: An empirical investigation. **Information Technology and People**, v. 32, n. 2, p. 297–318, 2019. ISSN 0959-3845.
- MARSLAND, S. **Machine Learning: An Algorithmic Perspective**. 6000 Broken Sound Parkway NW, Suite 300, Boca Raton: Taylor and Francis Group, LLC, 2015. 1-452 p. ISBN 978-1-4665-8333-7.
- MAY, R.; MAIER, H.; DANDY, G. Data splitting for artificial neural networks using som-based stratified sampling. **Neural Networks**, v. 23, n. 2, p. 283 – 294, 2010. ISSN 0893-6080.
- MEISEL, S. **Anticipatory Optimization for Dynamic Decision Making**. New York, USA: Springer, 2011. v. 51. 1-197 p. ISBN 9781461405054.
- METZGER, A. et al. Comparing and combining predictive business process monitoring techniques. **IEEE Transactions on Systems, Man, and Cybernetics: Systems**, v. 45, n. 2, p. 276–290, Feb 2015. ISSN 2168-2216.
- MICHALSKI, R.; CARBONELL, J.; MITCHELL, T. **Machine Learning An Artificial Intelligence Approach**. [S.l.]: Springer Science and Business Media, 2013. 1-588 p. ISBN 3662124076.
- MIGUEL, P. A. C. Estudo de caso na engenharia de produção: estruturação e recomendações para sua condução. **Production**, v. 17, n. 1, p. 216–229, apr 2007. ISSN 0103-6513.
- MJOLSNESS, E.; DECOSTE, D. Machine learning for science: State of the art and future prospects. **Science**, v. 293, p. 2051–2055, 2001. ISSN 1095-9203.
- MOKTADIR, M. A. et al. Barriers to big data analytics in manufacturing supply chains: A case study from bangladesh. **Computers and Industrial Engineering**, v. 128, p. 1063 – 1075, 2019. ISSN 0360-8352.
- MONTGOMERY, D. C.; RUNGER, G. C. **Applied Statistics and Probability for Engineers**. [S.l.]: John Wiley and Sons, 2003. 1-720 p. ISBN 978-0471381815.
- MORTENSON, M. J.; DOHERTY, N. F.; ROBINSON, S. Operational research from taylorism to terabytes: A research agenda for the analytics age. **European Journal of Operational Research**, v. 241, n. 3, p. 583 – 595, 2015. ISSN 0377-2217.
- MUN, J. **Chapter 3 - A Primer on Quantitative Risk Analysis**. 1. ed. [S.l.]: Academic Press, 2014. 63 - 118 p. ISBN 978-0-12-401690-3.
- MURTY, K. G. **Optimization Models For Decision Making**. University of Michigan, Ann Arbor, USA: University of Michigan, 2003. 1-640 p. ISBN 1234567.
- NAGI, J. et al. A computational intelligence scheme for the prediction of the daily peak load. **Applied Soft Computing**, v. 11, n. 8, p. 4773 – 4788, 2011. ISSN 1568-4946.
- NAJAFI, I. Assessment and Modeling of Decision-Making Process for e-Commerce Trust Based on Machine Learning Algorithms. In: **Lecture Notes in Electrical Engineering**. [S.l.]: Springer, 2019. v. 480, p. 969–986.
- NIETO, Y. et al. Usage of machine learning for strategic decision making at higher educational institutions. **IEEE Access**, v. 7, n. 1, p. 75008–750017, April 2019.

- O'DONOVAN, P. et al. Big data in manufacturing: a systematic mapping study. **Journal of Big Data**, v. 2, n. 1, p. 20, dec 2015. ISSN 2196-1115.
- OJA, E. Unsupervised learning in neural computation. **Theoretical Computer Science**, v. 287, n. 1, p. 187 – 207, 2002. ISSN 0304-3975. Natural Computing.
- OUYANG, Y. et al. Applying regression models to query-focused multi-document summarization. **Information Processing and Management**, v. 47, n. 2, p. 227 – 237, 2011. ISSN 0306-4573.
- PORTUGAL, I.; ALENCAR, P.; COWAN, D. The use of machine learning algorithms in recommender systems: A systematic review. **Expert Systems with Applications**, v. 97, p. 205–227, 2018. ISSN 0957-4174.
- POURNARAS, E.; PILGERSTORFER, P.; ASIKIS, T. Decentralized collective learning for self-managed sharing economies. **ACM Transactions on Autonomous and Adaptive Systems**, v. 13, n. 2, 2018. ISSN 1556-4665.
- POZZOLO, A. D. et al. Learned lessons in credit card fraud detection from a practitioner perspective. **Expert Systems with Applications**, v. 41, n. 10, p. 4915 – 4928, 2014. ISSN 0957-4174.
- QIU, D. et al. Sample selection according to categorical sensitivity for building cross-sell model. **Journal of Computational Information Systems**, v. 4, n. 2, p. 583–588, 2008. ISSN 1553-9105.
- QIU, J. et al. A survey of machine learning for big data processing. **EURASIP Journal on Advances in Signal Processing**, v. 2016, n. 1, p. 67, May 2016. ISSN 1687-6180.
- QUINLAN, J. R. Learning with continuous classes. In: **In Proceedings AI'92**. [S.l.: s.n.], 1992.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2020. Disponível em: <<http://www.R-project.org/>>.
- R-PROJECT. **Introduction to R. R Project**. 2018. Disponível em: <<https://www.r-project.org/about.html>>.
- RANASINGHE, R. et al. Application of artificial neural networks for predicting the impact of rolling dynamic compaction using dynamic cone penetrometer test results. **Journal of Rock Mechanics and Geotechnical Engineering**, v. 9, n. 2, p. 340 – 349, 2017. ISSN 1674-7755.
- RARDIN, R. L. **Optimization in Operations Research**. Second edi. [S.l.]: Pearson, 2015. 1-1187 p. ISBN 978-0134384559.
- RODRIGUEZ-PADIAL, N.; MARIN, M.; DOMINGO, R. An approach to integrating tactical decision-making in industrial maintenance balance scorecards using principal components analysis and machine learning. **Complexity**, v. 2017, p. 15, 2017. ISSN 1076-2787.
- ROKACH, L. Ensemble-based classifiers. **Artificial Intelligence Review**, v. 33, n. 1, p. 1–39, Feb 2010. ISSN 1573-7462.
- ROSS, S. M. **CHAPTER 12 - Linear Regression**. Third edition. [S.l.]: Academic Press, 2010. 537 - 604 p. ISBN 978-0-12-374388-6.
- RUDIN, C. et al. Machine learning for the new york city power grid. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 34, n. 2, p. 328–345, Feb 2012. ISSN 0162-8828.
- SAMUEL, A. L. Some studies in machine learning using the game of checkers. **IBM Journal of Research and Development**, v. 3, n. 3, p. 210–229, 1959. ISSN 210-229.

- SEGHIER, M. E. A. B. et al. Reliability analysis based on hybrid algorithm of m5 model tree and monte carlo simulation for corroded pipelines: Case of study x60 steel grade pipes. **Engineering Failure Analysis**, v. 97, p. 793 – 803, 2019. ISSN 1350-6307.
- SHAI, S.-S.; SHAI, B.-D. **Understanding Machine Learning: From Theory to Algorithms**. Avenue of the Americas, New York, USA: Cambridge University Press, 2014. 1-416 p. ISBN 978-1-107-05713-5.
- SHCHERBAKOV, M. et al. A survey of forecast error measures. **World Applied Sciences Journal**, v. 24, p. 171–176, 01 2013.
- SICILIANO, R.; MOLA, F. Multivariate data analysis and modeling through classification and regression trees. **Computational Statistics and Data Analysis**, v. 32, n. 3, p. 285 – 301, 2000. ISSN 0167-9473.
- SMOLA, A. J.; SCHOLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, v. 14, p. pages199–222, August 2004. ISSN 0960-3174.
- STRANDHAGEN, J. W. et al. The fit of Industry 4.0 applications in manufacturing logistics: a multiple case study. **Advances in Manufacturing**, v. 5, n. 4, p. 344–358, dec 2017.
- VERCELLIS, C. **Business Intelligence: Data Mining and Optimization for Decision Making**. Politecnico di Milano, Italy: John Wiley, 2009. 420 p. ISBN 9780470511381.
- VILLE, B. D. Decision trees. **Wiley Interdisciplinary Reviews: Computational Statistics**, v. 5, n. 6, p. 448–455, 2013.
- VOYANT, C. et al. Machine learning methods for solar radiation forecasting: A review. **Renewable Energy**, v. 105, p. 569 – 582, 2017. ISSN 0960-1481.
- WANG, G. et al. Two credit scoring models based on dual strategy ensemble trees. **Knowledge-Based Systems**, v. 26, p. 61–68, 2012. ISSN 0950-7051.
- WANG, Y.; WITTEN, I. H. Inducing model trees for continuous classes. In: **Proceedings of the poster papers of the European Conference on Machine Learning**. [S.l.: s.n.], 1997.
- WITTEN, I. H.; FRANK, E. **Data Mining: Practical Machine Learning Tools and Techniques**. Second edi. [S.l.]: ELSEIVER, 2005. 558 p. ISBN 0120884070.
- YAO, L.; GE, Z. Big data quality prediction in the process industry: A distributed parallel modeling framework. **Journal of Process Control**, v. 68, p. 1–13, aug 2018. ISSN 09591524.
- YUE, X. et al. Cloud-assisted industrial cyber-physical systems: An insight. **Microprocessors and Microsystems**, v. 39, n. 8, p. 1262–1270, nov 2015. ISSN 01419331.
- ZHAN, C.; GAN, A.; HADI, M. Prediction of lane clearance time of freeway incidents using the m5p tree algorithm. **IEEE Transactions on Intelligent Transportation Systems**, 2011.
- ZHANG, D.; TSAI, J. J. **Advances in Machine Learning Applications in Software Engineering**. 3 Henrietta Street, Covent Garden, London WC2E 8LU: Idea Group Publishing, 2007. 1-499 p. ISBN 1-59140-943-8.
- ZHANG, X. et al. Predicting customer churn through interpersonal influence. **Knowledge-Based Systems**, v. 28, p. 97 – 104, 2012. ISSN 0950-7051.

ZHONG, R. Y. et al. Big data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. **Computers and Industrial Engineering**, v. 101, p. 572–591, 2016. ISSN 0360-8352.

ZHOU, L. et al. Machine learning on big data: Opportunities and challenges. **Neurocomputing**, v. 237, n. January, p. 350–361, may 2017. ISSN 09252312.

ZHU, X.; GOLDBERG, A. **Introduction to Semi-Supervised Learning**. Madison, Wisconsin: Morgan and Claypool, 2009. v. 1. 1-130 p. ISBN 9781598295481.

APÊNDICE A – VALORES DA ANÁLISE DE CORRELAÇÃO DAS VARIÁVEIS

Tabela 17 – Valores da análise de correlação das variáveis

(continua)

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
x_1	1	-0,39	-0,06	0,23	-0,06	-0,11	-0,26	-0,03	-0,22	-0,06	0,15	0,19
x_2	-0,39	1	-0,14	-0,14	0,05	-0,10	0,03	0,03	0,60	0,00	-0,18	-0,23
x_3	-0,06	-0,14	1	-0,12	-0,05	-0,03	-0,27	0,05	-0,19	0,19	0,36	0,36
x_4	0,23	-0,14	-0,12	1	-0,02	-0,02	-0,10	-0,03	-0,09	-0,06	-0,07	-0,07
x_5	-0,06	0,05	-0,05	-0,02	1	-0,01	-0,05	0,02	-0,05	0,01	-0,02	0,00
x_6	-0,11	-0,10	-0,03	-0,02	-0,01	1	-0,05	-0,01	-0,04	-0,03	-0,03	0,03
x_7	-0,26	0,03	-0,27	-0,10	-0,05	-0,05	1	-0,06	-0,20	-0,16	-0,15	-0,15
x_8	-0,03	0,03	0,05	-0,03	0,02	-0,01	-0,06	1	-0,05	0,49	0,26	0,16
x_9	-0,22	0,60	-0,19	-0,09	-0,05	-0,04	-0,20	-0,05	1	-0,14	-0,13	-0,13
x_{10}	-0,06	0,00	0,19	-0,06	0,01	-0,03	-0,16	0,49	-0,14	1	0,42	0,20
x_{11}	0,15	-0,18	0,36	-0,07	-0,02	-0,03	-0,15	0,26	-0,13	0,42	1	0,68
x_{12}	0,19	-0,23	0,36	-0,07	0,00	-0,03	-0,15	0,16	-0,13	0,20	0,68	1
x_{13}	-0,17	-0,17	0,33	-0,04	-0,02	0,02	-0,08	-0,02	-0,07	0,11	-0,01	-0,05
x_{14}	0,30	-0,32	-0,14	-0,08	-0,05	-0,03	-0,21	-0,06	-0,18	-0,03	-0,05	-0,08
x_{15}	-0,06	0,01	0,20	-0,02	-0,01	-0,01	-0,05	-0,01	-0,05	-0,04	-0,04	-0,04
x_{16}	0,14	0,22	0,00	-0,03	-0,02	-0,02	-0,05	-0,02	-0,06	-0,04	-0,04	-0,04
x_{17}	0,03	0,01	0,11	-0,02	-0,01	-0,01	-0,04	-0,01	-0,04	-0,02	0,13	0,17
x_{18}	0,00	-0,06	-0,03	0,19	-0,01	-0,01	-0,04	-0,01	-0,04	-0,02	-0,03	-0,03
x_{19}	-0,03	0,03	0,03	-0,01	-0,01	-0,01	-0,03	0,63	-0,03	0,33	0,23	0,09
x_{20}	0,16	0,23	-0,05	-0,03	-0,01	-0,01	-0,06	-0,02	-0,05	-0,04	-0,04	-0,04
x_{21}	-0,03	0,05	0,10	-0,02	-0,01	-0,01	-0,05	-0,01	-0,05	0,51	0,32	0,04
x_{22}	-0,14	-0,15	0,16	-0,03	-0,02	-0,02	-0,07	-0,02	-0,06	0,05	-0,05	-0,05
x_{23}	-0,02	0,06	0,43	-0,03	-0,02	-0,02	-0,07	-0,02	-0,06	-0,05	-0,05	-0,04
x_{24}	-0,05	0,05	0,38	-0,03	-0,01	-0,01	-0,06	-0,02	-0,05	-0,05	-0,04	-0,04
x_{25}	-0,11	0,07	0,67	-0,05	-0,03	-0,03	-0,12	-0,03	-0,10	-0,09	-0,08	-0,08
x_{26}	0,14	0,06	0,04	-0,04	-0,02	-0,02	-0,11	0,04	-0,09	0,17	0,06	0,02
x_{27}	0,06	0,09	-0,02	-0,01	-0,01	-0,01	-0,03	-0,01	-0,02	-0,02	-0,02	-0,02
y	0,26	0,09	-0,11	0,08	-0,08	0,02	-0,34	-0,05	-0,04	-0,13	-0,02	0,01

Fonte: Próprio autor (2020).

Tabela 13 - Valores da análise de correlação das variáveis

(continua)

	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}	x_{24}
x_1	-0,17	0,30	-0,06	0,14	0,03	0,00	-0,03	0,16	-0,03	-0,14	-0,02	-0,05
x_2	-0,17	-0,32	0,01	0,22	0,01	-0,06	0,03	0,23	0,05	-0,15	0,06	0,05
x_3	0,33	-0,14	0,20	0,00	0,11	-0,03	0,03	-0,05	0,10	0,16	0,43	0,38
x_4	-0,04	-0,08	-0,02	-0,03	-0,02	0,19	-0,01	-0,03	-0,02	-0,03	-0,03	-0,03
x_5	-0,02	-0,05	-0,01	-0,02	-0,01	-0,01	-0,01	-0,01	-0,01	-0,02	-0,02	-0,01
x_6	0,02	-0,03	-0,01	-0,02	-0,01	-0,01	-0,01	-0,01	-0,01	-0,02	-0,02	-0,01
x_7	-0,08	-0,21	-0,05	-0,05	-0,04	-0,04	-0,03	-0,06	-0,05	-0,07	-0,07	-0,06
x_8	-0,02	-0,06	-0,01	-0,02	-0,01	-0,01	0,63	-0,02	-0,01	-0,02	-0,02	-0,02
x_9	-0,07	-0,18	-0,05	-0,06	-0,04	-0,04	-0,03	-0,05	-0,05	-0,06	-0,06	-0,05
x_{10}	0,11	-0,03	-0,04	-0,04	-0,02	-0,02	0,33	-0,04	0,51	0,05	-0,05	-0,05
x_{11}	-0,01	-0,05	-0,04	-0,04	0,13	-0,03	0,23	-0,04	0,32	-0,05	-0,05	-0,04
x_{12}	-0,05	-0,08	-0,04	-0,04	0,17	-0,03	0,09	-0,04	0,04	-0,05	-0,04	-0,04
x_{13}	1	-0,07	-0,02	-0,03	-0,02	-0,02	-0,01	-0,02	-0,02	0,37	-0,03	-0,02
x_{14}	-0,07	1	-0,05	-0,07	0,08	0,01	-0,03	-0,06	0,06	-0,07	0,06	-0,06
x_{15}	-0,02	-0,05	1	-0,02	-0,01	-0,01	-0,01	-0,02	-0,01	-0,02	-0,02	-0,01
x_{16}	-0,03	-0,07	-0,02	1	-0,01	-0,01	-0,01	0,68	-0,02	-0,02	-0,02	-0,01
x_{17}	-0,02	0,08	-0,01	-0,01	1	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01
x_{18}	-0,02	0,01	-0,01	-0,01	-0,01	1	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01
x_{19}	-0,01	-0,03	-0,01	-0,01	-0,01	-0,01	1	-0,01	-0,01	-0,01	-0,01	-0,01
x_{20}	-0,02	-0,06	-0,02	0,68	-0,01	-0,01	-0,01	1	-0,02	-0,02	-0,02	-0,02
x_{21}	-0,02	0,06	-0,01	-0,02	-0,01	-0,01	-0,01	-0,02	1	-0,02	-0,02	-0,02
x_{22}	0,37	-0,07	-0,02	-0,02	-0,01	-0,01	-0,01	-0,02	-0,02	1	-0,02	-0,02
x_{23}	-0,03	-0,06	-0,02	-0,02	-0,01	-0,01	-0,01	-0,02	-0,02	-0,02	1	0,53
x_{24}	-0,02	-0,06	-0,01	-0,01	-0,01	-0,01	-0,01	-0,02	-0,02	-0,02	0,53	1
x_{25}	-0,04	-0,11	-0,03	0,04	-0,02	-0,02	-0,02	-0,03	-0,03	-0,04	0,60	0,53
x_{26}	0,07	-0,01	-0,03	-0,03	0,00	-0,02	0,01	-0,03	0,09	-0,04	-0,03	-0,03
x_{27}	-0,01	-0,03	0,02	-0,01	-0,01	-0,01	0,00	-0,01	-0,01	-0,01	-0,01	-0,01
y	-0,04	0,34	0,10	0,21	0,08	0,11	-0,04	0,26	-0,08	-0,14	-0,09	-0,08

Fonte: Próprio autor (2020).

Tabela 13 - Valores da análise de correlação das variáveis

	(conclusão)			
	x_{25}	x_{26}	x_{27}	y
x_1	-0,11	0,14	0,06	0,26
x_2	0,07	0,06	0,09	0,09
x_3	0,66	0,04	-0,02	-0,11
x_4	-0,05	-0,04	-0,01	0,08
x_5	-0,03	-0,02	-0,01	-0,08
x_6	-0,03	-0,02	-0,01	0,02
x_7	-0,12	-0,11	-0,03	-0,34
x_8	-0,03	0,04	-0,01	-0,05
x_9	-0,10	-0,09	-0,02	-0,04
x_{10}	-0,09	0,17	-0,02	-0,13
x_{11}	-0,08	0,06	-0,02	-0,02
x_{12}	-0,08	0,02	-0,02	0,1
x_{13}	-0,04	0,07	-0,01	-0,04
x_{14}	-0,11	-0,01	-0,03	0,34
x_{15}	-0,03	-0,03	0,02	0,10
x_{16}	0,04	-0,03	-0,01	0,21
x_{17}	-0,02	0,00	-0,01	0,08
x_{18}	-0,02	-0,02	-0,01	0,11
x_{19}	-0,02	0,01	0,00	-0,04
x_{20}	-0,03	-0,03	-0,01	0,26
x_{21}	-0,03	0,09	-0,01	-0,08
x_{22}	-0,04	-0,04	-0,01	0,14
x_{23}	0,60	-0,03	-0,01	-0,09
x_{24}	0,53	-0,03	-0,01	-0,08
x_{25}	1	-0,06	-0,01	-0,15
x_{26}	-0,06	1	0,20	0,26
x_{27}	-0,01	0,20	1	0,14
y	-0,15	0,26	0,14	1

Fonte: Próprio autor (2020).

APÊNDICE B – ÁRVORE DE MODELO M5P

A árvore do modelo M5P é gerada usando o *software* R, nas linhas a seguir é apresentada a estrutura interna do modelo de previsão. Pode-se observar que o modelo apresentado é um sistema de inferência que contém um conjunto de regras internas das variáveis de entrada e, quando cada uma das regras é cumprida, a previsão é calculada por meio de uma equação de regressão linear. No código, LM significa modelo linear (*Linear model*) e, neste caso, há 78 modelos lineares gerados. Além disso, também é possível observar o número de instâncias usadas para a geração do modelo linear e o RMSE do nó que contém o referido modelo.

```

X2 <= 106,5 :
|  X2 <= 104,5 :
|  |  X1 <= 25 :
|  |  |  X14 <= 1,5 :
|  |  |  |  X22 <= 4 :
|  |  |  |  |  X3 <= 191,5 :
|  |  |  |  |  |  X6 <= 0,5 :
|  |  |  |  |  |  |  X25 <= 223 : LM1 ( Instancias= 265/ RMSE= 0,24)
|  |  |  |  |  |  |  X25 > 223 : LM2 ( Instancias= 67/ RMSE= 0,25)
|  |  |  |  |  |  |  X6 > 0,5 :
|  |  |  |  |  |  |  |  X6 <= 1,5 : LM3 ( Instancias= 27/ RMSE= 0,26)
|  |  |  |  |  |  |  |  X6 > 1,5 : LM4 ( Instancias= 86/ RMSE= 0,20)
|  |  |  |  |  |  |  X3 > 191,5 : LM5 ( Instancias= 113/ RMSE= 0,48)
|  |  |  |  |  |  X22 > 4 : LM6 ( Instancias= 308/ RMSE= 0,03)
|  |  |  |  X14 > 1,5 :
|  |  |  |  |  X2 <= 103,5 :
|  |  |  |  |  |  X10 <= 16 : LM7 ( Instancias= 245/ RMSE= 0,41)
|  |  |  |  |  |  X10 > 16 :
|  |  |  |  |  |  |  X2 <= 101,5 : LM8 ( Instancias= 17/ RMSE= 0,18)
|  |  |  |  |  |  |  X2 > 101,5 :
|  |  |  |  |  |  |  |  X18 <= 361 :
|  |  |  |  |  |  |  |  |  X3 <= 116,5 :
|  |  |  |  |  |  |  |  |  |  X14 <= 39 :
|  |  |  |  |  |  |  |  |  |  |  X18 <= 281,5 : LM9 ( Instancias= 13/ RMSE= 0,20)
|  |  |  |  |  |  |  |  |  |  |  X18 > 281,5 : LM10 ( Instancias= 4/ RMSE= 0,10)
|  |  |  |  |  |  |  |  |  |  X14 > 39 :
|  |  |  |  |  |  |  |  |  |  |  |  X6 <= 0,5 : LM11 ( Instancias= 7/ RMSE= 0,28)
|  |  |  |  |  |  |  |  |  |  |  |  X6 > 0,5 : LM12 ( Instancias= 2/ RMSE= 0,06)
|  |  |  |  |  |  |  |  |  |  |  X3 > 116,5 : LM13 ( Instancias= 22/ RMSE= 0,19)
|  |  |  |  |  |  |  |  |  |  X18 > 361 : LM14 ( Instancias= 14/ RMSE= 0,02)
|  |  |  |  |  |  X2 > 103,5 : LM15 ( Instancias= 118/ RMSE= 0,21)
|  |  |  X1 > 25 :
|  |  |  |  X3 <= 3,5 :
|  |  |  |  |  X14 <= 79 :
|  |  |  |  |  |  X2 <= 101,5 : LM16 ( Instancias= 573/ RMSE= 0,09)
|  |  |  |  |  |  X2 > 101,5 :
|  |  |  |  |  |  |  X4 <= 7,5 :
|  |  |  |  |  |  |  |  X14 <= 44 :
|  |  |  |  |  |  |  |  |  X14 <= 3 : LM17 ( Instancias= 54/ RMSE= 0,24)
|  |  |  |  |  |  |  |  |  X14 > 3 : LM18 ( Instancias= 111/ RMSE= 0,09)
|  |  |  |  |  |  |  |  X14 > 44 : LM19 ( Instancias= 46/ RMSE= 0,08)
|  |  |  |  |  |  |  X4 > 7,5 : LM20 ( Instancias= 323/ RMSE= 0,21)
|  |  |  |  |  X14 > 79 : LM21 ( Instancias= 296/ RMSE= 0,31)
|  |  |  |  X3 > 3,5 :
|  |  |  |  |  X14 <= 5 :
|  |  |  |  |  |  X10 <= 1 : LM22 ( Instancias= 536/ RMSE= 0,17)

```



```

| | | | | X10 > 1 :
| | | | | | X3 <= 61,5 : LM23 ( Instancias= 278/ RMSE= 0,14)
| | | | | | X3 > 61,5 :
| | | | | | | X11 <= 5082 :
| | | | | | | | X11 <= 1892,5 :
| | | | | | | | | X22 <= 4 :
| | | | | | | | | | X10 <= 370,5 : LM24 ( Instancias= 52/ RMSE= 0,18)
| | | | | | | | | | X10 > 370,5 :
| | | | | | | | | | | X26 <= 597 :
| | | | | | | | | | | | X11 <= 382,5 : LM25 ( Instancias= 3/ RMSE= 0,01)
| | | | | | | | | | | | X11 > 382,5 : LM26 ( Instancias= 2/ RMSE= 0,09)
| | | | | | | | | | | | X26 > 597 :
| | | | | | | | | | | | X3 <= 122,5 : LM27 ( Instancias= 7/ RMSE= 0,06)
| | | | | | | | | | | | X3 > 122,5 : LM28 ( Instancias= 2/ RMSE= 0,01)
| | | | | | | | | | | | X22 > 4 : LM29 ( Instancias= 18/ RMSE= 0,02)
| | | | | | | | | | | | X11 > 1892,5 : LM30 ( Instancias= 64/ RMSE= 0,16)
| | | | | | | | | | | | X11 > 5082 :
| | | | | | | | | | | | X3 <= 299 : LM31 ( Instancias= 16/ RMSE= 0,16)
| | | | | | | | | | | | X3 > 299 : LM32 ( Instancias= 31/ RMSE= 0,02)
| | | | | X14 > 5 :
| | | | | | X3 <= 14,5 : LM33 ( Instancias= 314/ RMSE= 0,45)
| | | | | | X3 > 14,5 :
| | | | | | | X14 <= 89,5 :
| | | | | | | | X4 <= 1 : LM34 ( Instancias= 233/ RMSE= 0,36)
| | | | | | | | X4 > 1 :
| | | | | | | | | X18 <= 371,5 :
| | | | | | | | | | X18 <= 258 : LM35 ( Instancias= 15/ RMSE= 0,88)
| | | | | | | | | | X18 > 258 : LM36 ( Instancias= 31/ RMSE= 0,03)
| | | | | | | | | | X18 > 371,5 : LM37 ( Instancias= 40/ RMSE= 0,02)
| | | | | | | | | | X14 > 89,5 :
| | | | | | | | | | | X11 <= 1231 : LM38 ( Instancias= 69/ RMSE= 0,36)
| | | | | | | | | | | X11 > 1231 : LM39 ( Instancias= 18/ RMSE= 0,19)
| X2 > 104,5 :
| | X3 <= 142,5 :
| | | X3 <= 3,5 :
| | | | X7 <= 15,5 : LM40 ( Instancias= 1618/ RMSE= 0,07)
| | | | X7 > 15,5 :
| | | | | X7 <= 17,5 : LM41 ( Instancias= 582/ RMSE= 0,08)
| | | | | X7 > 17,5 :
| | | | | | X7 <= 19,5 : LM42 ( Instancias= 146/ RMSE= 0,13)
| | | | | | X7 > 19,5 :
| | | | | | | X7 <= 21,5 : LM43 ( Instancias= 30/ RMSE= 0,05)
| | | | | | | X7 > 21,5 : LM44 ( Instancias= 18/ RMSE= 0,44)
| | | | X3 > 3,5 :
| | | | | X12 <= 190,5 :
| | | | | | X10 <= 82,5 : LM45 ( Instancias= 302/ RMSE= 0,26)
| | | | | | X10 > 82,5 :
| | | | | | | X3 <= 29,5 : LM46 ( Instancias= 85/ RMSE= 0,31)
| | | | | | | X3 > 29,5 : LM47 ( Instancias= 63/ RMSE= 0,16)
| | | | | | X12 > 190,5 :
| | | | | | | X8 <= 0,5 :
| | | | | | | | X21 <= 0,5 : LM48 ( Instancias= 143/ RMSE= 0,08)
| | | | | | | | X21 > 0,5 : LM49 ( Instancias= 143/ RMSE= 0,04)
| | | | | | | X8 > 0,5 : LM50 ( Instancias= 170/ RMSE= 0,03)
| | X3 > 142,5 :
| | | X25 <= 224,5 :
| | | | X2 <= 105,5 : LM51 ( Instancias= 149/ RMSE= 0,01)
| | | | X2 > 105,5 : LM52 ( Instancias= 71/ RMSE= 0,29)
| | | | X25 > 224,5 :
| | | | | X24 <= 17 :
| | | | | | X25 <= 848,5 : LM53 ( Instancias= 300/ RMSE= 0,06)
| | | | | | X25 > 848,5 :

```

```

| | | | | | | X16 <= 48 :
| | | | | | | | X3 <= 323,5 : LM54 ( Instancias= 24/ RMSE= 0,05)
| | | | | | | | X3 > 323,5 : LM55 ( Instancias= 34/ RMSE= 0,04)
| | | | | | | | X16 > 48 : LM56 ( Instancias= 18/ RMSE= 0,14)
| | | | | X24 > 17 : LM57 ( Instancias= 239/ RMSE= 0,10)
X2 > 106,5 :
| | X9 <= 23 :
| | | X2 <= 107,5 :
| | | | X14 <= 45 : LM58 ( Instancias= 54/ RMSE= 0,52)
| | | | X14 > 45 :
| | | | | X10 <= 37 :
| | | | | | X14 <= 128 : LM59 ( Instancias= 110/ RMSE= 0,14)
| | | | | | X14 > 128 : LM60 ( Instancias= 47/ RMSE= 0,20)
| | | | | | X10 > 37 : LM61 ( Instancias= 56/ RMSE= 0,26)
| | | X2 > 107,5 :
| | | | X2 <= 109,5 :
| | | | | X14 <= 2 :
| | | | | | X3 <= 14,5 : LM62 ( Instancias= 38/ RMSE= 0,27)
| | | | | | X3 > 14,5 :
| | | | | | | X12 <= 12 : LM63 ( Instancias= 37/ RMSE= 0,52)
| | | | | | | X12 > 12 : LM64 ( Instancias= 114/ RMSE= 0,51)
| | | | | | X14 > 2 :
| | | | | | | X17 <= 14,5 : LM65 ( Instancias= 204/ RMSE= 0,26)
| | | | | | | X17 > 14,5 : LM66 ( Instancias= 61/ RMSE= 0,01)
| | | | X2 > 109,5 :
| | | | | X26 <= 254 :
| | | | | | X20 <= 219,5 :
| | | | | | | X20 <= 155,5 : LM67 ( Instancias= 26/ RMSE= 0,27)
| | | | | | | X20 > 155,5 : LM68 ( Instancias= 100/ RMSE= 0,01)
| | | | | | X20 > 219,5 :
| | | | | | | X20 <= 275 : LM69 ( Instancias= 70/ RMSE= 0,08)
| | | | | | | X20 > 275 :
| | | | | | | | X20 <= 567,5 : LM70 ( Instancias= 38/ RMSE= 0,17)
| | | | | | | | X20 > 567,5 :
| | | | | | | | | X20 <= 680,5 :
| | | | | | | | | | X10 <= 11,5 : LM71 ( Instancias= 6/ RMSE= 0,04)
| | | | | | | | | | X10 > 11,5 :
| | | | | | | | | | | X16 <= 81 : LM72 ( Instancias= 3/ RMSE= 0,03)
| | | | | | | | | | | X16 > 81 : LM73 ( Instancias= 3/ RMSE= 0,14)
| | | | | | | | | | | X20 > 680,5 : LM74 ( Instancias= 5/ RMSE= 0,04)
| | | | | | X26 > 254 :
| | | | | | | X3 <= 1,5 : LM75 ( Instancias= 104/ RMSE= 0,06)
| | | | | | | X3 > 1,5 :
| | | | | | | | X26 <= 1531 : LM76 ( Instancias= 21/ RMSE= 0,32)
| | | | | | | | X26 > 1531 : LM77 ( Instancias= 34/ RMSE= 0,11)
| | X9 > 23 : LM78 ( Instancias= 1684/ RMSE= 0,07)

```

APÊNDICE C – EQUAÇÕES DE REGRESSÃO DA ÁRVORE DE MODELO M5P

Como a estrutura da árvore M5P consiste em uma série de regras SE-ENTÃO, para calcular uma previsão, o modelo processa as informações de entrada usando essas regras para finalmente produzir o valor previsto utilizando a equação de predição. Assim, como são gerados 78 modelos lineares (equações de previsão), para uma ordem maior, cada y_i corresponde a um modelo linear, onde $i = 1, 2, 3, \dots, 78$; ou seja, para LM 1, sua equação é y_1 .

Modelo linear 1 com instâncias = 265 e RMSE = 0,24:

$$y_1 = 44,256 \times x_1 + 45,5459 \times x_2 + 0,0719 \times x_3 - 1,4562 \times x_4 - 0,1404 \times x_5 + 9,5628 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,0201 \times x_{10} + 0,0009 \times x_{11} - 0,032 \times x_{12} + 0,0099 \times x_{13} + 0,0763 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,2241 \times x_{17} + 0,0369 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1585 \times x_{22} + 0,0004 \times x_{23} + 0,1321 \times x_{25} + 0,008 \times x_{26} + 0,0058 \times x_{27} - 3773,1859$$

Modelo linear 2 com instâncias = 67 e RMSE = 0,25:

$$y_2 = 12,2614 \times x_1 + 143,827 \times x_2 + 0,0719 \times x_3 - 1,4562 \times x_4 - 0,1404 \times x_5 + 9,5628 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,0201 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,0688 \times x_{13} + 0,0763 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,2241 \times x_{17} + 0,0369 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1585 \times x_{22} + 0,0004 \times x_{23} + 1,7582 \times x_{25} + 0,0151 \times x_{26} + 0,0058 \times x_{27} - 13807,1451$$

Modelo linear 3 com instâncias = 27 e RMSE = 0,26:

$$y_3 = 7,5779 \times x_1 + 19,1347 \times x_2 + 5,6515 \times x_3 - 1,4562 \times x_4 - 0,1404 \times x_5 + 21,5398 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,0201 \times x_{10} - 0,0007 \times x_{11} + 0,0014 \times x_{12} - 0,0145 \times x_{13} + 0,0763 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,2241 \times x_{17} + 0,0369 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1585 \times x_{22} + 0,0004 \times x_{23} + 0,1382 \times x_{25} + 0,005 \times x_{26} + 0,0058 \times x_{27} - 641,2526$$

Modelo linear 4 com instâncias = 86 e RMSE = 0,20:

$$y_4 = 7,5779 \times x_1 + 19,1347 \times x_2 + 0,6542 \times x_3 - 1,4562 \times x_4 - 0,1404 \times x_5 + 17,5394 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,0201 \times x_{10} - 0,0007 \times x_{11} + 0,0014 \times x_{12} - 0,0145 \times x_{13} + 0,0763 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,2241 \times x_{17} + 0,0369 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1585 \times x_{22} + 0,0004 \times x_{23} + 0,1382 \times x_{25} + 0,005 \times x_{26} + 0,0058 \times x_{27} - 521,9067$$

Modelo linear 5 com instâncias = 113 e RMSE = 0,48:

$$y_5 = 5,8422 \times x_1 + 19,265 \times x_2 + 1,8534 \times x_3 - 1,4562 \times x_4 - 0,1404 \times x_5 + 204,6475 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,0459 \times x_{10} + 0,0019 \times x_{11} + 0,0014 \times x_{12} - 0,4768 \times x_{13} + 0,0763 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,2241 \times x_{17} + 0,0369 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1585 \times x_{22} + 0,0004 \times x_{23} + 0,121 \times x_{25} + 0,005 \times x_{26} + 0,0058 \times x_{27} - 855,4358$$

Modelo linear 6 com instâncias = 308 e RMSE = 0,03:

$$y_6 = 1,8539 \times x_1 + 12,183 \times x_2 + 2,3095 \times x_3 - 1,4562 \times x_4 - 0,1404 \times x_5 + 5,9447 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 2,2793 \times x_{10} + 0,0028 \times x_{11} + 0,0014 \times x_{12} + 2,1952 \times x_{13} + 0,0763 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,3411 \times x_{17} + 0,0369 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,4284 \times x_{22} + 0,0004 \times x_{23} + 0,0521 \times x_{25} + 0,0086 \times x_{26} + 0,0058 \times x_{27} - 1103,3559$$

Modelo linear 7 com instâncias = 245 e RMSE = 0,41:

$$y_7 = 52,996 \times x_1 - 32,0402 \times x_2 + 0,1145 \times x_3 - 6,8768 \times x_4 - 0,1404 \times x_5 + 125,9708 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,002 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,0479 \times x_{13} - 0,7059 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 0,1705 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,0188 \times x_{26} + 0,0058 \times x_{27} + 4226,4381$$

Modelo linear 8 com instâncias = 17 e RMSE = 0,18:

$$y_8 = 11,8058 \times x_1 + 73,6335 \times x_2 - 0,5941 \times x_3 - 28,9707 \times x_4 - 0,1404 \times x_5 + 325,8786 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,3249 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,5214 \times x_{13} + 0,7771 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 1,2146 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,2245 \times x_{26} + 0,0058 \times x_{27} - 6240,7998$$

Modelo linear 9 com instâncias = 13 e RMSE = 0,20:

$$y_9 = 11,8058 \times x_1 - 7,9809 \times x_2 + 1,3687 \times x_3 - 4,0469 \times x_4 - 0,1404 \times x_5 + 447,6158 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,1339 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,0727 \times x_{13} + 2,326 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 1,2825 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,1152 \times x_{26} + 0,0058 \times x_{27} + 2188,7136$$

Modelo linear 10 com instâncias = 4 e RMSE = 0,10:

$$y_{10} = 11,8058 \times x_1 - 7,9809 \times x_2 + 1,3687 \times x_3 - 4,8148 \times x_4 - 0,1404 \times x_5 + 447,6158 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,1339 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,0727 \times x_{13} + 2,326 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 1,2825 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,1152 \times x_{26} + 0,0058 \times x_{27} + 2215,8049$$

Modelo linear 11 com instâncias = 7 e RMSE = 0,28:

$$y_{11} = 11,8058 \times x_1 - 7,9809 \times x_2 + 1,6956 \times x_3 - 7,6149 \times x_4 - 0,1404 \times x_5 + 526,8807 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,1339 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} - 0,0079 \times x_{13} + 2,8183 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 1,386 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,1152 \times x_{26} + 0,0058 \times x_{27} + 2190,8386$$

Modelo linear 12 com instâncias = 2 e RMSE = 0,06:

$$y_{12} = 11,8058 \times x_1 - 7,9809 \times x_2 + 1,6956 \times x_3 - 7,6149 \times x_4 - 0,1404 \times x_5 + 541,5822 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,1339 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} - 0,0079 \times x_{13} + 2,8183 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 1,386 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,1152 \times x_{26} + 0,0058 \times x_{27} + 2206,6412$$

Modelo linear 13 com instâncias = 22 e RMSE = 0,19:

$$y_{13} = 11,8058 \times x_1 - 14,3556 \times x_2 + 0,1409 \times x_3 - 7,6149 \times x_4 - 0,1404 \times x_5 + 371,7012 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,1339 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,3146 \times x_{13} + 0,9317 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 1,0013 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,1256 \times x_{26} + 0,0058 \times x_{27} + 3103,6065$$

Modelo linear 14 com instâncias = 14 e RMSE = 0,02:

$$y_{14} = 11,8058 \times x_1 + 50,9846 \times x_2 + 1,8655 \times x_3 + 2,2846 \times x_4 - 0,1404 \times x_5 + 363,1627 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 - 0,1339 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} + 0,3146 \times x_{13} + 0,0835 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 2,141 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,0188 \times x_{26} + 0,0058 \times x_{27} - 4048,7676$$

Modelo linear 15 com instâncias = 118 e RMSE = 0,21:

$$y_{15} = 163,2288 \times x_1 + 60,5575 \times x_2 + 14,0017 \times x_3 - 13,3191 \times x_4 - 0,1404 \times x_5 + 11,8997 \times x_6 - 0,1241 \times x_7 - 0,0156 \times x_8 - 0,0181 \times x_9 + 0,002 \times x_{10} + 0,0009 \times x_{11} + 0,0014 \times x_{12} - 0,0198 \times x_{13} - 4,3316 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,073 \times x_{17} + 0,1195 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1947 \times x_{22} + 0,0004 \times x_{23} + 0,0119 \times x_{25} + 0,3554 \times x_{26} + 0,0058 \times x_{27} - 5998,2779$$

Modelo linear 16 com instâncias = 573 e RMSE = 0,09:

$$y_{16} = 2,2334 \times x_1 - 1,6416 \times x_2 - 14,1047 \times x_3 + 0,8543 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0045 \times x_8 - 0,0181 \times x_9 + 0,3244 \times x_{10} + 0,0234 \times x_{11} + 0,0017 \times x_{12} - 0,0233 \times x_{13} + 1,6059 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,036 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,3955 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} - 0,0088 \times x_{26} + 0,0058 \times x_{27} + 1593,1071$$

Modelo linear 17 com instâncias = 54 e RMSE = 0,24:

$$y_{17} = 0,5374 \times x_1 - 46,1902 \times x_2 + 21,6566 \times x_3 + 31,4852 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0045 \times x_8 - 0,0181 \times x_9 + 1,2296 \times x_{10} + 0,0006 \times x_{11} + 0,0017 \times x_{12} - 0,0233 \times x_{13} + 1,7066 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,036 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,3955 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} - 0,5469 \times x_{26} + 0,0058 \times x_{27} + 6070,7566$$

Modelo linear 18 com instâncias = 111 e RMSE = 0,09:

$$y_{18} = 0,3897 \times x_1 - 46,1902 \times x_2 + 2,2974 \times x_3 + 9,6436 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0045 \times x_8 - 0,0181 \times x_9 + 0,2264 \times x_{10} + 0,0006 \times x_{11} + 0,0017 \times x_{12} - 0,0233 \times x_{13} + 2,3406 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,036 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,3955 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} - 0,0915 \times x_{26} + 0,0058 \times x_{27} + 6272,4622$$

Modelo linear 19 com instâncias = 46 e RMSE = 0,08:

$$y_{19} = 0,4046 \times x_1 - 304,0194 \times x_2 + 0,9213 \times x_3 + 4,4554 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0045 \times x_8 - 0,0181 \times x_9 + 0,2621 \times x_{10} + 0,0006 \times x_{11} + 0,0017 \times x_{12} - 0,0233 \times x_{13} + 4,6482 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,036 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,3955 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} - 0,1325 \times x_{26} + 0,0058 \times x_{27} + 32697,3752$$

Modelo linear 20 com instâncias = 323 e RMSE = 0,21:

$$y_{20} = 0,094 \times x_1 - 15,8809 \times x_2 + 52,0296 \times x_3 + 52,1298 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0045 \times x_8 - 0,0181 \times x_9 + 0,0572 \times x_{10} + 0,0006 \times x_{11} + 0,0017 \times x_{12} - 0,0233 \times x_{13} + 0,308 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,036 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,3955 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} - 0,0095 \times x_{26} + 0,0058 \times x_{27} + 2672,0887$$

Modelo linear 21 com instâncias = 296 e RMSE = 0,31:

$$y_{21} = 0,0577 \times x_1 - 44,5255 \times x_2 - 0,5029 \times x_3 + 0,834 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0045 \times x_8 - 0,0181 \times x_9 + 0,0183 \times x_{10} + 0,0017 \times x_{11} + 0,0017 \times x_{12} - 0,0233 \times x_{13} - 0,9846 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,036 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,3955 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0018 \times x_{26} + 0,0058 \times x_{27} + 6266,1536$$

Modelo linear 22 com instâncias = 536 e RMSE = 0,17:

$$y_{22} = 4,4644 \times x_1 + 0,7441 \times x_2 + 0,6915 \times x_3 + 1,8666 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,0473 \times x_8 - 0,0181 \times x_9 + 0,038 \times x_{10} + 0,0079 \times x_{11} + 0,0334 \times x_{12} - 0,0836 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 1,1768 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0871 \times x_{26} + 0,0058 \times x_{27} + 1026,8596$$

Modelo linear 23 com instâncias = 278 e RMSE = 0,14:

$$y_{23} = 2,8856 \times x_1 + 31,514 \times x_2 + 2,5921 \times x_3 + 26,5565 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,8823 \times x_8 - 0,0181 \times x_9 + 0,4994 \times x_{10} + 0,0354 \times x_{11} + 0,0128 \times x_{12} - 0,1966 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 2,0137 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0084 \times x_{26} + 0,0058 \times x_{27} - 2002,071$$

Modelo linear 24 com instâncias = 52 e RMSE = 0,18:

$$y_{24} = 1,4134 \times x_1 + 0,7441 \times x_2 + 0,4976 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 1,1355 \times x_{10} + 0,0587 \times x_{11} + 0,057 \times x_{12} - 0,5924 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 7,8153 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0094 \times x_{26} + 0,0058 \times x_{27} + 1121,038$$

Modelo linear 25 com instâncias = 3 e RMSE = 0,01:

$$y_{25} = 1,4134 \times x_1 + 0,7441 \times x_2 - 2,2118 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 1,0931 \times x_{10} + 0,4318 \times x_{11} + 0,057 \times x_{12} - 0,5924 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 7,8153 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0728 \times x_{26} + 0,0058 \times x_{27} + 1378,7264$$

Modelo linear 26 com instâncias = 2 e RMSE = 0,09:

$$y_{26} = 1,4134 \times x_1 + 0,7441 \times x_2 - 2,2118 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 1,0931 \times x_{10} + 0,435 \times x_{11} + 0,057 \times x_{12} - 0,5924 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 7,8153 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0728 \times x_{26} + 0,0058 \times x_{27} + 1391,6946$$

Modelo linear 27 com instâncias = 7 e RMSE = 0,06:

$$y_{27} = 1,4134 \times x_1 + 0,7441 \times x_2 - 2,189 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 1,0931 \times x_{10} + 0,3309 \times x_{11} + 0,057 \times x_{12} - 0,5924 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 7,8153 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0622 \times x_{26} + 0,0058 \times x_{27} + 1380,7667$$

Modelo linear 28 com instâncias = 2 e RMSE = 0,01:

$$y_{28} = 1,4134 \times x_1 + 0,7441 \times x_2 - 2,2959 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 1,0931 \times x_{10} + 0,3309 \times x_{11} + 0,057 \times x_{12} - 0,5924 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 7,8153 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0622 \times x_{26} + 0,0058 \times x_{27} + 1387,7281$$

Modelo linear 29 com instâncias = 18 e RMSE = 0,02:

$$y_{29} = 1,4134 \times x_1 + 0,7441 \times x_2 + 0,8744 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 1,8455 \times x_{10} + 0,0776 \times x_{11} + 0,057 \times x_{12} + 0,3683 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 11,3019 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0094 \times x_{26} + 0,0058 \times x_{27} + 630,7401$$

Modelo linear 30 com instâncias = 64 e RMSE = 0,16:

$$y_{30} = 5,0226 \times x_1 + 0,7441 \times x_2 + 0,1056 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 0,4143 \times x_{10} + 0,0006 \times x_{11} + 0,1269 \times x_{12} - 0,6411 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 5,8952 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0895 \times x_{26} + 0,0058 \times x_{27} + 1012,5722$$

Modelo linear 31 com instâncias = 16 e RMSE = 0,16:

$$y_{31} = 2,6575 \times x_1 + 0,7441 \times x_2 + 0,6868 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 + 0,3815 \times x_{10} + 0,0129 \times x_{11} + 0,0572 \times x_{12} - 0,6643 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 5,5091 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0094 \times x_{26} + 0,0058 \times x_{27} + 1136,9508$$

Modelo linear 32 com instâncias = 31 e RMSE = 0,02:

$$y_{32} = 2,2725 \times x_1 + 0,7441 \times x_2 + 0,525 \times x_3 + 4,3946 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,1237 \times x_8 - 0,0181 \times x_9 - 0,0334 \times x_{10} + 0,1006 \times x_{11} + 0,1454 \times x_{12} - 0,6643 \times x_{13} + 0,081 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,0573 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 5,5091 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,3932 \times x_{26} + 0,0058 \times x_{27} + 518,9739$$

Modelo linear 33 com instâncias = 314 e RMSE = 0,45:

$$y_{33} = -0,0735 \times x_1 - 0,9605 \times x_2 - 19,5786 \times x_3 + 3,8195 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 0,0669 \times x_{10} - 0,0015 \times x_{11} + 0,6247 \times x_{12} - 0,0189 \times x_{13} + 0,6586 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,1302 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0097 \times x_{26} + 0,0058 \times x_{27} + 1803,5433$$

Modelo linear 34 com instâncias = 233 e RMSE = 0,36:

$$y_{34} = -0,3245 \times x_1 - 0,588 \times x_2 + 0,0086 \times x_3 + 7,7128 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 0,43 \times x_{10} + 0,0457 \times x_{11} + 0,0556 \times x_{12} - 0,0189 \times x_{13} + 2,9692 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,2644 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0155 \times x_{26} + 0,0058 \times x_{27} + 1507,7277$$

Modelo linear 35 com instâncias = 15 e RMSE = 0,88:

$$y_{35} = 19,97 \times x_1 - 0,588 \times x_2 + 0,0086 \times x_3 + 26,0897 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 0,2044 \times x_{10} + 0,0069 \times x_{11} + 0,0556 \times x_{12} - 0,0189 \times x_{13} + 0,8223 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,8528 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,5504 \times x_{26} + 0,0058 \times x_{27} + 939,0137$$

Modelo linear 36 com instâncias = 31 e RMSE = 0,03:

$$y_{36} = 2,8723 \times x_1 - 0,588 \times x_2 + 0,0086 \times x_3 + 26,0897 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 0,2044 \times x_{10} + 0,0069 \times x_{11} + 0,0556 \times x_{12} - 0,0189 \times x_{13} + 0,8223 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 2,7773 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,5504 \times x_{26} + 0,0058 \times x_{27} + 953,1351$$

Modelo linear 37 com instâncias = 40 e RMSE = 0,02:

$$y_{37} = -0,3245 \times x_1 - 0,588 \times x_2 + 0,0086 \times x_3 + 30,3688 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 0,2044 \times x_{10} + 0,0069 \times x_{11} + 0,0556 \times x_{12} - 0,0189 \times x_{13} + 0,8223 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 3,6848 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,6088 \times x_{26} + 0,0058 \times x_{27} + 765,221$$

Modelo linear 38 com instâncias = 69 e RMSE = 0,36:

$$y_{38} = -0,9381 \times x_1 - 0,588 \times x_2 + 0,3223 \times x_3 + 9,3289 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 0,4852 \times x_{10} - 0,001 \times x_{11} + 0,2572 \times x_{12} - 0,0189 \times x_{13} + 0,5069 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,3252 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0318 \times x_{26} + 0,0058 \times x_{27} + 1704,9835$$

Modelo linear 39 com instâncias = 18 e RMSE = 0,19:

$$y_{39} = -45,6469 \times x_1 - 0,588 \times x_2 + 6,8566 \times x_3 + 9,3289 \times x_4 - 0,1404 \times x_5 + 0,2652 \times x_6 - 0,1241 \times x_7 + 0,025 \times x_8 - 0,0181 \times x_9 + 4,3129 \times x_{10} - 0,001 \times x_{11} + 0,4745 \times x_{12} - 0,0189 \times x_{13} + 0,5069 \times x_{14} + 0,0087 \times x_{15} + 0,0018 \times x_{16} + 0,0445 \times x_{17} + 0,3252 \times x_{18} + 0,0027 \times x_{20} - 0,0221 \times x_{21} - 0,9032 \times x_{22} + 0,0004 \times x_{23} - 0,0012 \times x_{25} + 0,0318 \times x_{26} + 0,0058 \times x_{27} + 3297,4943$$

Modelo linear 40 com instâncias = 1618 e RMSE = 0,07:

$$\begin{aligned}
 y_{40} = & 0,0118 \times x_1 + 1,2939 \times x_2 - 0,0027 \times x_3 + 0,0991 \times x_4 + 1,5416 \times x_5 + \\
 & 0,2747 \times x_6 - 0,0179 \times x_7 + 0,0036 \times x_8 - 0,0181 \times x_9 - 0,0031 \times x_{10} + \\
 & 0 \times x_{11} + 0,0015 \times x_{12} + 0,0032 \times x_{13} - 8,811 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,3725 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0155 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,0648 \times x_{25} + 0,0018 \times x_{26} + \\
 & 0,0176 \times x_{27} + 918,3949
 \end{aligned}$$

Modelo linear 41 com instâncias = 582 e RMSE = 0,08:

$$\begin{aligned}
 y_{41} = & 0,0118 \times x_1 + 1,2939 \times x_2 - 0,0027 \times x_3 + 0,0991 \times x_4 - 0,0327 \times x_5 + \\
 & 0,2747 \times x_6 + 0,461 \times x_7 + 0,0036 \times x_8 - 0,0181 \times x_9 - 0,0031 \times x_{10} + \\
 & 0 \times x_{11} + 0,0015 \times x_{12} + 0,0032 \times x_{13} - 0,1171 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,022 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0155 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,0027 \times x_{25} + 0,0018 \times x_{26} + \\
 & 0,0176 \times x_{27} + 917,089
 \end{aligned}$$

Modelo linear 42 com instâncias = 146 e RMSE = 0,13:

$$\begin{aligned}
 y_{42} = & 0,0118 \times x_1 + 1,2939 \times x_2 - 0,0027 \times x_3 + 0,0991 \times x_4 - 0,0327 \times x_5 + \\
 & 0,2747 \times x_6 + 4,5794 \times x_7 + 0,0036 \times x_8 - 0,0181 \times x_9 - 0,0031 \times x_{10} + \\
 & 0 \times x_{11} + 0,0015 \times x_{12} + 0,0032 \times x_{13} - 0,1171 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,022 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0155 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,0027 \times x_{25} + 0,0018 \times x_{26} + \\
 & 0,0176 \times x_{27} + 858,0169
 \end{aligned}$$

Modelo linear 43 com instâncias = 30 e RMSE = 0,05:

$$\begin{aligned}
 y_{43} = & 0,0118 \times x_1 + 1,2939 \times x_2 - 0,0027 \times x_3 + 0,0991 \times x_4 - 0,0327 \times x_5 + \\
 & 0,2747 \times x_6 + 35,2202 \times x_7 + 0,0036 \times x_8 - 0,0181 \times x_9 - 0,0031 \times x_{10} + \\
 & 0 \times x_{11} + 0,0015 \times x_{12} + 0,0032 \times x_{13} - 0,1171 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,022 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0155 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,0027 \times x_{25} + 0,0018 \times x_{26} + \\
 & 0,0176 \times x_{27} + 219,6133
 \end{aligned}$$

Modelo linear 44 com instâncias = 18 e RMSE = 0,44:

$$\begin{aligned}
 y_{44} = & 0,0118 \times x_1 + 1,2939 \times x_2 - 0,0027 \times x_3 + 0,0991 \times x_4 - 0,0327 \times x_5 + \\
 & 0,2747 \times x_6 + 127,9411 \times x_7 + 0,0036 \times x_8 - 0,0181 \times x_9 - 0,0031 \times x_{10} + \\
 & 0 \times x_{11} + 0,0015 \times x_{12} + 0,0032 \times x_{13} - 0,1171 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,022 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0155 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,0027 \times x_{25} + 0,0018 \times x_{26} + \\
 & 0,0176 \times x_{27} - 1822,5831
 \end{aligned}$$

Modelo linear 45 com instâncias = 302 e RMSE = 0,26:

$$\begin{aligned}
 y_{45} = & 8,0755 \times x_1 + 169,6828 \times x_2 - 1,5769 \times x_3 + 0,0991 \times x_4 - 0,1208 \times x_5 + \\
 & 0,2747 \times x_6 + 0,0449 \times x_7 + 0,0459 \times x_8 - 0,0181 \times x_9 + 0,7862 \times x_{10} + \\
 & 0 \times x_{11} + 0,0046 \times x_{12} + 0,0032 \times x_{13} - 0,0091 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,0163 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0091 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,004 \times x_{25} + 0,0022 \times x_{26} + \\
 & 0,0176 \times x_{27} - 17062,5071
 \end{aligned}$$

Modelo linear 46 com instâncias = 85 e RMSE = 0,31:

$$\begin{aligned}
 y_{46} = & 0,4003 \times x_1 + 39,3293 \times x_2 + 7,1861 \times x_3 + 0,0991 \times x_4 + 45,925 \times x_5 + \\
 & 0,2747 \times x_6 + 0,0449 \times x_7 + 0,1816 \times x_8 - 0,0181 \times x_9 - 0,0028 \times x_{10} - \\
 & 0,0018 \times x_{11} + 0,0046 \times x_{12} + 0,0032 \times x_{13} - 0,0091 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,0163 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0091 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,004 \times x_{25} + 0,0022 \times x_{26} + \\
 & 0,0176 \times x_{27} - 3420,5895
 \end{aligned}$$

Modelo linear 47 com instâncias = 63 e RMSE = 0,16:

$$\begin{aligned}
 y_{47} = & 0,4003 \times x_1 + 39,3293 \times x_2 + 0,3602 \times x_3 + 0,0991 \times x_4 + 7,8964 \times x_5 + \\
 & 0,2747 \times x_6 + 0,0449 \times x_7 + 0,5189 \times x_8 - 0,0181 \times x_9 + 0,0971 \times x_{10} - \\
 & 0,0082 \times x_{11} + 0,0046 \times x_{12} + 0,0032 \times x_{13} - 0,0091 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,0163 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0091 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,004 \times x_{25} - 0,1404 \times x_{26} + \\
 & 0,0176 \times x_{27} - 3160,6691
 \end{aligned}$$

Modelo linear 48 com instâncias = 143 e RMSE = 0,08:

$$\begin{aligned}
 y_{48} = & 0,0249 \times x_1 + 12,5595 \times x_2 + 0,1876 \times x_3 + 0,0991 \times x_4 - 0,1208 \times x_5 + \\
 & 0,2747 \times x_6 + 0,0449 \times x_7 + 0,0255 \times x_8 - 0,0181 \times x_9 - 0,0084 \times x_{10} - \\
 & 0,01 \times x_{11} + 0,0094 \times x_{12} + 0,0032 \times x_{13} - 0,0091 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,0163 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0335 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,004 \times x_{25} + 0,0022 \times x_{26} + \\
 & 0,0176 \times x_{27} - 229,1597
 \end{aligned}$$

Modelo linear 49 com instâncias = 143 e RMSE = 0,04:

$$\begin{aligned}
 y_{49} = & 0,0249 \times x_1 + 12,5595 \times x_2 - 0,0113 \times x_3 + 0,0991 \times x_4 - 0,1208 \times x_5 + \\
 & 0,2747 \times x_6 + 0,0449 \times x_7 + 0,0255 \times x_8 - 0,0181 \times x_9 - 0,0084 \times x_{10} - \\
 & 0,0008 \times x_{11} + 0,0094 \times x_{12} + 0,0032 \times x_{13} - 0,0091 \times x_{14} + 0,0209 \times x_{15} + \\
 & 0,0163 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0335 \times x_{21} - \\
 & 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,004 \times x_{25} + 0,5428 \times x_{26} + \\
 & 0,0176 \times x_{27} - 474,1948
 \end{aligned}$$

Modelo linear 50 com instâncias = 170 e RMSE = 0,03:

$$y_{50} = 0,0249 \times x_1 + 12,5595 \times x_2 + 0,0097 \times x_3 + 0,0991 \times x_4 - 0,1208 \times x_5 + 0,2747 \times x_6 + 0,0449 \times x_7 + 0,0327 \times x_8 - 0,0181 \times x_9 - 0,0096 \times x_{10} - 0,0004 \times x_{11} + 0,0068 \times x_{12} + 0,0032 \times x_{13} - 0,0091 \times x_{14} + 0,0209 \times x_{15} + 0,0163 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0487 \times x_{21} - 0,0972 \times x_{22} + 0,0011 \times x_{23} + 0,003 \times x_{24} + 0,004 \times x_{25} + 0,0022 \times x_{26} + 0,0176 \times x_{27} - 176,7043$$

Modelo linear 51 com instâncias = 149 e RMSE = 0,01:

$$y_{51} = -0,0448 \times x_1 + 52,5652 \times x_2 + 2,8786 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,0271 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 3,1864 \times x_{15} + 0,0082 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} + 0,01 \times x_{23} + 0,0579 \times x_{24} + 0,006 \times x_{25} + 0,0353 \times x_{26} + 3,1809 \times x_{27} - 5444,6914$$

Modelo linear 52 com instâncias = 71 e RMSE = 0,29:

$$y_{52} = -0,0448 \times x_1 + 82,2528 \times x_2 + 0,3819 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,1631 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 0,768 \times x_{15} + 0,0082 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} + 0,01 \times x_{23} + 0,0579 \times x_{24} + 0,006 \times x_{25} + 0,0503 \times x_{26} + 0,7621 \times x_{27} - 7791,6081$$

Modelo linear 53 com instâncias = 300 e RMSE = 0,06:

$$y_{53} = -0,0451 \times x_1 + 8,7672 \times x_2 + 0,0217 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,0065 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 0,1299 \times x_{15} - 0,0714 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} - 0,2183 \times x_{23} + 0,0586 \times x_{24} + 0,0162 \times x_{25} + 0,0085 \times x_{26} + 0,1253 \times x_{27} + 131,1871$$

Modelo linear 54 com instâncias = 24 e RMSE = 0,05:

$$y_{54} = -0,0451 \times x_1 + 8,7672 \times x_2 + 0,3875 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,0065 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 0,1299 \times x_{15} + 0,0585 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} - 0,0604 \times x_{23} + 0,0586 \times x_{24} + 0,2723 \times x_{25} + 0,0085 \times x_{26} + 0,1253 \times x_{27} - 222,0479$$

Modelo linear 55 com instâncias = 34 e RMSE = 0,04:

$$y_{55} = -0,0451 \times x_1 + 8,7672 \times x_2 + 0,3367 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,0065 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 0,1299 \times x_{15} + 0,0585 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} - 0,057 \times x_{23} + 0,0586 \times x_{24} + 0,0342 \times x_{25} + 0,0085 \times x_{26} + 0,1253 \times x_{27} + 27,0638$$

Modelo linear 56 com instâncias = 18 e RMSE = 0,14:

$$y_{56} = -0,0451 \times x_1 + 8,7672 \times x_2 + 0,2806 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,0065 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 0,1299 \times x_{15} - 0,0665 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} - 0,0661 \times x_{23} + 0,0586 \times x_{24} + 0,0342 \times x_{25} + 0,0085 \times x_{26} + 0,1253 \times x_{27} + 109,5258$$

Modelo linear 57 com instâncias = 239 e RMSE = 0,10:

$$y_{57} = -0,879 \times x_1 + 8,7672 \times x_2 + 0,3859 \times x_3 + 0,0991 \times x_4 - 0,1245 \times x_5 + 0,2747 \times x_6 + 0,0217 \times x_7 + 0,0044 \times x_8 - 0,0181 \times x_9 - 0,0034 \times x_{10} + 0 \times x_{11} + 0,0065 \times x_{12} + 0,0032 \times x_{13} - 0,0065 \times x_{14} + 0,1299 \times x_{15} + 0,0082 \times x_{16} + 0,0256 \times x_{17} + 0,01 \times x_{18} + 0,0027 \times x_{20} - 0,0095 \times x_{21} - 0,0972 \times x_{22} + 0,2893 \times x_{23} + 0,0746 \times x_{24} + 0,0127 \times x_{25} + 0,0085 \times x_{26} + 0,1253 \times x_{27} + 29,6099$$

Modelo linear 58 com instâncias = 54 e RMSE = 0,52:

$$y_{58} = -97,2843 \times x_1 - 20,8653 \times x_2 - 2,4428 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,1508 \times x_{10} + 0,0041 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 13,3377 \times x_{14} + 0,008 \times x_{15} + 0,0557 \times x_{16} - 1,2219 \times x_{17} + 0,0094 \times x_{18} + 0,1111 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0208 \times x_{26} + 0,0063 \times x_{27} + 7542,5519$$

Modelo linear 59 com instâncias = 110 e RMSE = 0,14:

$$y_{59} = -5,1846 \times x_1 - 20,8653 \times x_2 - 0,6131 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,1137 \times x_{10} - 0,0014 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,656 \times x_{14} + 0,008 \times x_{15} + 0,0557 \times x_{16} - 1,2219 \times x_{17} + 0,0094 \times x_{18} + 0,1111 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0208 \times x_{26} + 0,0063 \times x_{27} + 6313,6395$$

Modelo linear 60 com instâncias = 47 e RMSE = 0,20:

$$y_{60} = -5,1846 \times x_1 - 20,8653 \times x_2 - 0,6131 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,1137 \times x_{10} - 0,0014 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,5767 \times x_{14} + 0,008 \times x_{15} + 0,0557 \times x_{16} - 1,2219 \times x_{17} + 0,0094 \times x_{18} + 0,1111 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,348 \times x_{26} + 0,0063 \times x_{27} + 6255,2625$$

Modelo linear 61 com instâncias = 56 e RMSE = 0,26:

$$y_{61} = -5,1846 \times x_1 - 20,8653 \times x_2 + 0,0787 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 0,5686 \times x_{10} + 0,0327 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,6597 \times x_{14} + 0,008 \times x_{15} + 0,0557 \times x_{16} - 1,2219 \times x_{17} + 0,0094 \times x_{18} + 0,1111 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0996 \times x_{26} + 0,0063 \times x_{27} + 6149,5578$$

Modelo linear 62 com instâncias = 38 e RMSE = 0,27:

$$y_{62} = 4,2996 \times x_1 - 259,3779 \times x_2 + 0,0669 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 1,1386 \times x_{10} - 0,0185 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,7519 \times x_{14} + 0,008 \times x_{15} + 0,0065 \times x_{16} + 0,1843 \times x_{17} + 0,0094 \times x_{18} + 0,0554 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 1,8739 \times x_{26} + 0,0031 \times x_{27} + 29061,6772$$

Modelo linear 63 com instâncias = 37 e RMSE = 0,52:

$$y_{63} = 14,1897 \times x_1 - 1134,3834 \times x_2 - 0,1381 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,8075 \times x_{10} - 0,0236 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,7519 \times x_{14} + 0,008 \times x_{15} + 0,0065 \times x_{16} + 0,1843 \times x_{17} + 0,0094 \times x_{18} + 0,0554 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0126 \times x_{26} + 0,0031 \times x_{27} + 123856,6733$$

Modelo linear 64 com instâncias = 114 e RMSE = 0,51:

$$y_{64} = 3,4567 \times x_1 - 210,2138 \times x_2 - 0,1381 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,0108 \times x_{10} - 0,0157 \times x_{11} + 0,1121 \times x_{12} + 0,0059 \times x_{13} + 0,7519 \times x_{14} + 0,008 \times x_{15} + 0,0065 \times x_{16} + 0,1843 \times x_{17} + 0,0094 \times x_{18} + 0,0554 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0126 \times x_{26} + 0,0031 \times x_{27} + 23863,3073$$

Modelo linear 65 com instâncias = 204 e RMSE = 0,26:

$$y_{65} = 1,0048 \times x_1 - 29,5296 \times x_2 - 1,1529 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,0108 \times x_{10} - 0,0055 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,8291 \times x_{14} + 0,008 \times x_{15} + 0,0065 \times x_{16} + 0,567 \times x_{17} + 0,0094 \times x_{18} + 0,0554 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0126 \times x_{26} + 0,0031 \times x_{27} + 4881,0155$$

Modelo linear 66 com instâncias = 61 e RMSE = 0,01:

$$y_{66} = 1,0048 \times x_1 - 29,5296 \times x_2 + 5,4168 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 5,9166 \times x_{10} + 5,9219 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 7,0555 \times x_{14} + 0,008 \times x_{15} + 0,0065 \times x_{16} + 7,4667 \times x_{17} + 0,0094 \times x_{18} + 0,0554 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0126 \times x_{26} + 0,0031 \times x_{27} + 3642,3028$$

Modelo linear 67 com instâncias = 26 e RMSE = 0,27:

$$y_{67} = 5,4834 \times x_1 + 14,6164 \times x_2 + 12,6702 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 74,7103 \times x_{10} - 0,0031 \times x_{11} - 2,4682 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} - 0,0206 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} - 1,183 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} + 912,6583$$

Modelo linear 68 com instâncias = 100 e RMSE = 0,01:

$$y_{68} = 0,9339 \times x_1 + 14,6164 \times x_2 + 7,5472 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 7,5375 \times x_{10} - 0,0031 \times x_{11} + 4,3137 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} + 0,174 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 6,3462 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} - 813,4898$$

Modelo linear 69 com instâncias = 70 e RMSE = 0,08:

$$y_{69} = 0,6301 \times x_1 + 14,6164 \times x_2 + 4,9261 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 6,6817 \times x_{10} - 0,0031 \times x_{11} + 7,352 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} + 0,3032 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 6,1739 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} - 657,5125$$

Modelo linear 70 com instâncias = 38 e RMSE = 0,17:

$$y_{70} = 0,6301 \times x_1 + 14,6164 \times x_2 - 0,2343 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 6,3716 \times x_{10} - 0,0031 \times x_{11} - 0,3051 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} - 0,2445 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} - 0,6872 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} + 1677,6353$$

Modelo linear 71 com instâncias = 6 e RMSE = 0,04:

$$y_{71} = 0,6301 \times x_1 + 14,6164 \times x_2 - 0,2343 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 9,096 \times x_{10} - 0,0031 \times x_{11} - 0,3051 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} + 1,5357 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 0,7092 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} + 359,395$$

Modelo linear 72 com instâncias = 3 e RMSE = 0,03:

$$y_{72} = 0,6301 \times x_1 + 14,6164 \times x_2 - 0,2343 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 8,692 \times x_{10} - 0,0031 \times x_{11} - 0,3051 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} + 1,7767 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 0,7092 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} + 349,367$$

Modelo linear 73 com instâncias = 3 e RMSE = 0,14:

$$y_{73} = 0,6301 \times x_1 + 14,6164 \times x_2 - 0,2343 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 8,692 \times x_{10} - 0,0031 \times x_{11} - 0,3051 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} + 1,7767 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 0,7092 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} + 357,1193$$

Modelo linear 74 com instâncias = 5 e RMSE = 0,04:

$$y_{74} = 0,6301 \times x_1 + 14,6164 \times x_2 - 0,2343 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 + 5,6313 \times x_{10} - 0,0031 \times x_{11} - 0,3051 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} + 1,1016 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 1,7449 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0153 \times x_{26} + 0,0065 \times x_{27} - 159,6571$$

Modelo linear 75 com instâncias = 104 e RMSE = 0,06:

$$y_{75} = 2,2547 \times x_1 + 14,6164 \times x_2 - 0,2343 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,0108 \times x_{10} - 0,0031 \times x_{11} - 2,0293 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} - 0,0338 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 0,0752 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0068 \times x_{26} + 0,0031 \times x_{27} + 1308,773$$

Modelo linear 76 com instâncias = 21 e RMSE = 0,32:

$$y_{76} = 3,0088 \times x_1 + 14,6164 \times x_2 - 0,9262 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,0108 \times x_{10} - 0,0031 \times x_{11} - 2,0293 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} - 0,0338 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 0,0752 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0054 \times x_{26} - 0,0018 \times x_{27} + 1401,5923$$

Modelo linear 77 com instâncias = 34 e RMSE = 0,11:

$$y_{77} = 1,2885 \times x_1 + 14,6164 \times x_2 - 0,4245 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 - 0,1341 \times x_9 - 0,0108 \times x_{10} - 0,0031 \times x_{11} - 2,0293 \times x_{12} + 0,0059 \times x_{13} + 0,4664 \times x_{14} + 0,008 \times x_{15} - 0,0338 \times x_{16} - 0,2837 \times x_{17} + 0,0094 \times x_{18} + 0,0752 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0079 \times x_{26} - 0,0005 \times x_{27} + 1406,7971$$

Modelo linear 78 com instâncias = 1684 e RMSE = 0,07:

$$y_{78} = 0,1792 \times x_1 - 2,5934 \times x_2 - 0,0692 \times x_3 + 0,0828 \times x_4 - 0,2676 \times x_5 + 0,4096 \times x_6 - 0,1978 \times x_7 + 0,0013 \times x_8 + 0,2967 \times x_9 - 0,0089 \times x_{10} - 0,0004 \times x_{11} + 0,0006 \times x_{12} + 0,0059 \times x_{13} + 0,1072 \times x_{14} + 0,008 \times x_{15} + 0,0102 \times x_{16} - 0,1633 \times x_{17} + 0,0094 \times x_{18} + 0,0223 \times x_{20} - 0,0399 \times x_{21} - 0,0728 \times x_{22} - 0,0027 \times x_{25} + 0,0045 \times x_{26} + 0,0012 \times x_{27} + 1707,7452$$

APÊNDICE D – APLICATIVO FEITO EM SHINY PARA TOMADA DE DECISÃO - TELA 1

Figura 45 – Interface de aplicação feito em Shiny para tomada de decisão - Tela 1

CUSTO DE TRANSPORTE

Informação

Calcular custo

O que você encontrará aqui

Uma ferramenta interativa para apoiar a tomada de decisões, para ajudá-lo a obter mais rapidamente o custo de transporte de uma certa quantidade de itens para um determinado destino.

Fonte: R Core Team (2020)

APÊNDICE E – APLICATIVO FEITO EM SHINY PARA TOMADA DE DECISÃO - TELA 2

Figura 46 – Interface de aplicação feito em Shiny para tomada de decisão - Tela 2

The interface is titled "CUSTO DE TRANSPORTE" and features a sidebar with "Informação" and "Calcular custo" options. The main content area includes:

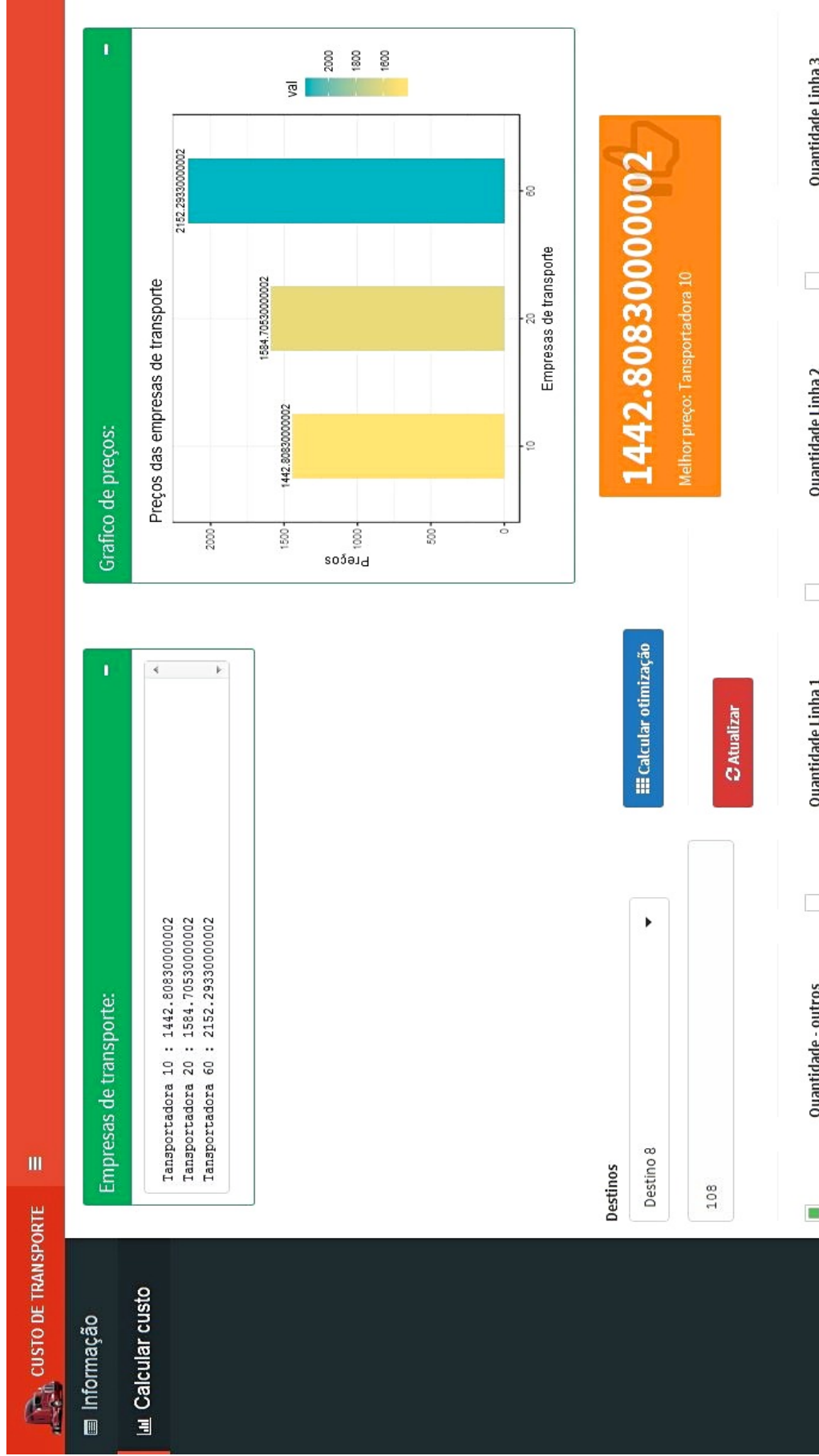
- Empresas de transporte:** A green bar with a minus sign (-) and a "Gráfico de preços:" label with a plus sign (+).
- Destinos:** A dropdown menu with "Choose" and a right-pointing arrow, followed by an empty text input field.
- Buttons:** "Calcular otimização" (blue) and "Atualizar" (red).
- Quantity Inputs:** A grid of 15 input fields, each with a checkbox and a label:

<input type="checkbox"/>	Quantidade - outros	<input type="checkbox"/>	Quantidade Linha 1	<input type="checkbox"/>	Quantidade Linha 2	<input type="checkbox"/>	Quantidade Linha 3
<input type="checkbox"/>	Quantidade Linha 4	<input type="checkbox"/>	Quantidade Linha 5	<input type="checkbox"/>	Quantidade Linha 6	<input type="checkbox"/>	Quantidade Linha 7
<input type="checkbox"/>	Quantidade Linha 8	<input type="checkbox"/>	Quantidade Linha 9	<input type="checkbox"/>	Quantidade Linha 10	<input type="checkbox"/>	Quantidade Linha 11
<input type="checkbox"/>	Quantidade Linha 12	<input type="checkbox"/>	Quantidade Linha 13	<input type="checkbox"/>	Quantidade Linha 14	<input type="checkbox"/>	Quantidade Linha 15

Fonte: R Core Team (2020)

APÊNDICE F – APLICATIVO FEITO EM SHINY PARA TOMADA DE DECISÃO - TELA 3

Figura 47 – Interface de aplicação feito em Shiny para tomada de decisão - Tela 3



Fonte: R Core Team (2020)