

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO DE JOINVILLE
CURSO DE ENGENHARIA MECATRÔNICA

GUILHERME CERVI GARCIA

ANÁLISE DE MÉTODOS DE MACHINE LEARNING NA PREDIÇÃO DO
COMPORTAMENTO DE UM BUFFER DE TEMPERATURA.

Joinville
2022

GUILHERME CERVI GARCIA

ANÁLISE DE MÉTODOS DE MACHINE LEARNING NA PREDIÇÃO DO
COMPORTAMENTO DE UM BUFFER DE TEMPERATURA.

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de bacharel em Engenharia Mecatrônica no curso de Engenharia Mecatrônica, da Universidade Federal de Santa Catarina, Centro Tecnológico de Joinville.

Orientador: Dr. João Carlos Xavier

Joinville
2022

Dedico este trabalho a minha família, em especial, ao meu pai Paulo e minha mãe
Gicele.

AGRADECIMENTOS

Agradeço a todas as pessoas que contribuíram, com suas potências, neste trabalho. À minha família, primeiramente aos meus pais, Paulo e Gicele, que sempre me incentivaram e apoiaram tanto nos momentos fáceis quanto nos momentos difíceis desta jornada. Também aos meus irmãos, Linny e Rick, e a Graciele e ao Paulo, meus padrinhos, por toda a estrutura que me proporcionaram.

A todos os amigos que a universidade me proporcionou, por me acompanharem neste movimento que é a universidade. Nela muitas aprendizagens, inúmeros desafios, especialmente, com uma pandemia. Pela ajuda nos estudos e por todas as nossas boas lembranças. A Magali, uma pessoa especial, que conheci neste final de estudos. Um obrigado especial a meu professor orientador Dr. João Carlos Xavier pela orientação e oportunidade. Ao Programa de Educação Tutorial por ter me apresentado ao professor e contribuído para o início desta pesquisa. E ao meu companheiro de pesquisa Thiago, figura essencial para o desenvolvimento deste trabalho.

Aos professores Dr. Diogo Lôndero da Silva e Dr. Benjamin Grando Moreira por aceitarem fazerem parte desta banca. À Universidade Federal de Santa Catarina que me proporcionou todas as oportunidades necessárias para esta conclusão.

RESUMO

As vacinas representam um grande avanço para a humanidade. São responsáveis por combater e erradicar diversas doenças que causaram grandes perdas. Recentemente o mundo foi lembrado da sua importância pela pandemia do COVID-19. Por serem materiais termolábeis os imunobiológicos precisam de uma rede de frio bem estruturada para manter sua eficiência. Afim de aumentar a segurança da temperatura das vacinas foi desenvolvido um sistema embarcado para monitoramento de temperatura. O projeto foi realizado por estudantes da Universidade Federal de Santa Catarina do campus de Joinville. O sistema embarcado utiliza um amortecedor de temperatura físico. Entretanto o Manual de Rede de Frio proíbe a inserção dos termômetros dentro de qualquer recipiente nas geladeiras de vacinas. De forma a obedecer o Manual foi projetado substituí-lo por uma técnica de inteligência artificial. Assim, o objetivo do trabalho é treinar e comparar diferentes métodos de inteligência artificial com a finalidade de resolver este problema. Foram desenvolvidos algoritmos de árvores de decisão, florestas aleatórias, máquinas de vetores de suporte e redes neurais. Em seguida, foram executados uma série de treinamentos variando as entradas e parâmetros de treino. A métrica escolhida para medir a acurácia de cada algoritmo foi o erro médio quadrático. A partir dos resultados obtidos foram gerados gráficos com o objetivo de identificar o impacto de cada variação dos parâmetros. Analisando os resultados foi constatado que as florestas aleatórias obtiveram o melhor desempenho dentre os testes realizados, com um erro quadrático médio de 0,304.

Palavras-chave: Temperatura. Inteligência Artificial. Vacinas. Rede de Frio. Aprendizado de Máquina.

ABSTRACT

The vaccines represent a big breakthrough to humanity. They are responsible for fight and eradicate various diseases that caused huge losses. Recently the world remembered the importance of the vaccines due to COVID-19 pandemic. As they are thermolabile material the immunobiological need a cold chain well structured to keep their efficiency. To increase the safety of their temperature a embedded system for temperature monitoring was developed. The project was accomplished by students of the Universidade Federal de Santa Catarina of the Joinville campus. The embedded system uses a physical temperature buffer. Nonetheless the Manual de Rede de Frio forbids the insertion of thermometers inside any kind of container inside vaccines refrigerators. To comply with the Manual was intended to replace it with an artificial intelligence algorithm. Therefore, the purpose of these work is to train and compare different methods of artificial intelligence with the goal of solving this problem. Were developed algorithms of decision threes, random forests, support vector machines and neural networks. Next, there were executed a series of trainings varying their entries and training parameters. The metric chosen to measure the precision of each algorithm was the mean squared error. From the results graphics were generated with the intent to identify the impact of each variation in the parameters. And with the analysis of the graphics, it was found that the random forests got the best results within the tests.

Keywords: Temperature. Artificial Intelligence. Vaccines. Cold Chain. Machine Learning.

LISTA DE FIGURAS

Figura 1 – Sistema embarcado para monitoramento de temperatura.	16
Figura 2 – Exemplo de uma árvore de decisão para regressão.	18
Figura 3 – Exemplo de uma floresta aleatória.	19
Figura 4 – Exemplo de uma SVM de regressão.	20
Figura 5 – Funções de ativação e suas derivadas.	21
Figura 6 – Função ELU.	22
Figura 7 – Gráfico dos dados de treinamento.	25
Figura 8 – Gráfico dos dados padronizados de acordo com a equação 2.	26
Figura 9 – Gráfico dos dados normalizados de acordo com a equação 3.	27
Figura 10 – Gráfico do MSE de acordo com a profundidade.	32
Figura 11 – Gráfico do MSE da árvore para os conjuntos 1 a 3.	32
Figura 12 – Gráfico do MSE da árvore para os conjuntos 4 a 6.	33
Figura 13 – Gráfico do MSE da árvore para os conjuntos 7 a 9.	33
Figura 14 – Gráfico do MSE da floresta com dados normalizados para os conjuntos 1 a 9.	34
Figura 15 – Gráfico do MSE da floresta com dados padronizados para os conjuntos 1 a 9.	34
Figura 16 – Gráfico do MSE da SVM com dados normalizados para a função gaussiana.	35
Figura 17 – Gráfico do MSE da SVM com dados normalizados para a função polinomial.	36
Figura 18 – Gráfico do MSE da SVM com dados padronizados para a função gaussiana.	36
Figura 19 – Gráfico do MSE da SVM com dados padronizados para a função polinomial.	37
Figura 20 – Gráfico do MSE em relação as épocas da RNA.	38
Figura 21 – Gráfico do MSE em relação aos conjuntos de treinamento da RNA.	38
Figura 22 – Gráfico de previsões e valores esperados da Árvore de Decisão.	40
Figura 23 – Gráfico de previsões e valores esperados da Floresta Aleatória.	40
Figura 24 – Gráfico de previsões e valores esperados da Máquina de Vetores de Suporte.	41
Figura 25 – Gráfico de previsões e valores esperados das Redes Neurais Artificiais.	41

LISTA DE TABELAS

Tabela 1 – Definição dos conjuntos de treinamento	27
Tabela 2 – Modelo 1	28
Tabela 3 – Modelo 2	29
Tabela 4 – Modelo 3	29
Tabela 5 – Modelo 4	29
Tabela 6 – Modelo 5	30
Tabela 7 – Modelo 6	30
Tabela 8 – Tabela de MSE dos modelos de RNAs para os dados padronizados	39
Tabela 9 – Tabela de MSE dos modelos de RNAs para os dados normalizados	39
Tabela 10 – Tabela dos resultados finais	42

LISTA DE ABREVIATURAS E SIGLAS

ELU	Unidade Exponencial Linear
GPU	Unidade de processamento Gráfico
GVE	Grupo de Vigilância Epidemiológica
IA	Inteligência Artificial
MLP	Perceptron Multicamada
MSE	Erro Quadrático Médio
OMS	Organização Mundial da Saúde
PNI	Programa Nacional de Imunizações
ReLU	Unidade Linear Retificada
RNA	Rede Neural Artificial
SUS	Sistema Único de Saúde
SVM	Máquinas de Vetores de Suporte

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Objetivos	12
1.1.1	Objetivo Geral	12
1.1.2	Objetivos Específicos	12
2	FUNDAMENTAÇÃO TEÓRICA	13
2.0.1	Rede de Frio	13
2.0.1.1	A cadeia de frio	14
2.0.2	O dispositivo de Monitoramento de Temperatura	15
2.0.3	Inteligência Artificial	16
2.0.4	Árvores de Decisão	18
2.0.5	Florestas Aleatórias	18
2.0.6	Máquinas de Vetores de Suporte	19
2.0.7	Redes Neurais Artificiais	20
2.0.7.1	Funções de ativação	21
2.0.7.2	Inicializadores	22
2.0.7.3	Otimizadores	22
2.0.8	Métricas de Desempenho	23
3	MÉTODO	24
3.0.1	Dados de Treinamento	24
3.0.2	Árvores de Regressão	27
3.0.3	Florestas Aleatórias	28
3.0.4	Máquina de Vetores	28
3.0.5	Redes Neurais Artificiais	28
3.0.5.1	Modelos	28
3.0.6	Outros Parâmetros	29
4	ANÁLISE DE DADOS	31
4.0.1	Árvores de Decisão	31
4.0.1.1	Profundidade	31
4.0.1.2	Conjuntos de treinamento	31
4.0.2	Florestas Aleatórias	32
4.0.3	Máquinas de Vetores de Suporte	35
4.0.4	Redes Neurais Artificiais	37
4.0.4.1	Número de Épocas	37
4.0.4.2	Conjuntos de Treinamento	37

4.0.4.3	Modelos	39
4.0.5	Comparação Entre as Técnicas	39
5	CONCLUSÕES	43
	REFERÊNCIAS	45

1 INTRODUÇÃO

A imunização da população é uma das mais importantes intervenções da humanidade na área da saúde. As vacinas são responsáveis por salvar milhões de vidas todos os anos. São responsáveis também por ter erradicado a varíola e pelo processo de erradicação da poliomielite. Para que as vacinas cheguem à população com a potência necessária, uma rede de frio com bom funcionamento é parte fundamental do processo (HANSON et al., 2017).

Por serem produtos sensíveis à ação da temperatura, para evitar deterioração, os imunizantes precisam ser armazenados em temperaturas controladas e ideais para manter sua capacidade imunogênica. Esta prática exige cuidado e uma boa estrutura. A faixa de temperatura ideal para evitar danos à maioria das vacinas é entre +2°C e +8°C (PATINE et al., 2021).

A perda de imunobiológicos por alteração da temperatura é um problema recorrente nas unidades de vacinação. Dados do Grupo de Vigilância Epidemiológica (GVE) do estado de São Paulo, coletados entre 2010 e 2017, mostram que 2,1% do total das doses aplicadas foram descartadas e 3,0% das doses utilizadas estiveram fora da faixa de temperatura ideal (PATINE et al., 2021). Um fator decisivo para o agravamento deste problema é que apesar de não ser recomendado pelo Manual da Rede de Frio (MINISTÉRIO DA SAÚDE, 2017), os refrigeradores domésticos são realidade em muitas Unidades Básicas de Saúde.

Para amenizar este problema, os membros do projeto de pesquisa intitulado *Monitoramento de Temperatura de Geladeiras Domésticas Utilizadas na Rede Pública de Saúde para Armazenamento de Vacinas* desenvolveram um sistema embarcado o qual permite o monitoramento automatizado da temperatura de geladeiras domésticas utilizadas no armazenamento de vacinas (ROSSIGNATTI et al., 2021). Trata-se de um projeto de baixo custo, que se comunica com o usuário através de um aplicativo para celular. Tem-se como objetivo evitar perdas de vacinas, ocasionadas por uma possível falha da rede de frio. Outro objetivo é acompanhar a temperatura, evitando também a utilização de imunobiológicos que estiveram fora da faixa adequada (ROSSIGNATTI et al., 2021).

O sistema embarcado utiliza um recipiente contendo glicerina para simular a temperatura real das vacinas. Este recipiente é chamado de amortecedor, ou em inglês *buffer*, de temperatura. O Manual da Rede de Frio proíbe a inserção de termômetros em recipientes, com ou sem líquidos dentro (MINISTÉRIO DA SAÚDE, 2017). Para contornarmos este problema foi pensado no desenvolvimento de uma técnica de Inteligência Artificial (IA) para substituir este buffer.

O trabalho foi realizado estudando alguns dos principais métodos de inteligência artificial. Entre estes estão as redes neurais, as máquinas de vetores de suporte, as árvores de decisão e as florestas aleatórias. Estes métodos foram comparados a partir da variação de suas entradas e parâmetros de forma a encontrar o algoritmo cuja resposta mais se encaixe com os valores esperados.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

O objetivo deste trabalho é desenvolver uma ferramenta capaz de simular o comportamento térmico de um *buffer* de temperatura utilizando algoritmos de inteligência artificial.

1.1.2 Objetivos Específicos

- Desenvolver uma ferramenta capaz de melhorar o sistema embarcado;
- Implementar diferentes algoritmos de Inteligência artificial. Eles são: árvores de decisão, florestas aleatórias, máquinas de vetores de suporte e redes neurais artificiais;
- Treinar os métodos de inteligência artificial de forma a conseguirem simular o comportamento térmico de um *buffer* de temperatura;
- Testar e comparar estes algoritmos utilizando diferentes dados de entrada;
- Identificar a técnica de aprendizado de máquina mais eficiente para o problema.

2 FUNDAMENTAÇÃO TEÓRICA

As vacinas representam uma das mais importantes medidas de prevenção de doenças. A primeira vacina foi desenvolvida em 1796 por Edward Jenner após vários anos de pesquisa realizando experimentos com a varíola bovina, o que deu origem ao termo *vaccine* derivado do termo latino *vacca* (APS et al., 2018). A varíola é considerada uma das doenças mais devastadoras da história da humanidade e causou a morte de milhões antes de ser erradicada. O último caso natural da doença foi na Somália em 1977. Em 1980 a varíola foi considerada erradicada pela OMS. Isto se deu devido a um intenso plano de vacinação iniciado em 1967 pela Organização Mundial da Saúde (OMS) (THE WORLD HEALTH ORGANIZATION, 2022).

O Programa Nacional de Imunizações (PNI), criado em 1973 por determinação do Ministério da Saúde, representou um grande avanço para a saúde pública do Brasil. O programa tinha como objetivo coordenar as ações de imunização que possuíam como principais características a descontinuidade, o caráter episódico e a reduzida área de cobertura. A institucionalização do programa se deu em 1975, e foi resultado de fatores de âmbito nacional e internacional que estimulavam e expandiam a utilização destes agentes imunizantes. Por conseguinte o PNI passou a coordenar as atividades de imunização desenvolvidas rotineiramente na rede de serviços (MINISTÉRIO DA SAÚDE, 2022a).

A legislação específica sobre imunizações e vigilância epidemiológica, Lei 6.259 de 30-10-1975 e Decreto 78.231 de 30-12-76, deu ênfase às atividades permanentes de vacinação e contribuiu para fortalecer institucionalmente o Programa (MINISTÉRIO DA SAÚDE, 2022a). Como resultado deste fortalecimento, 19 vacinas recomendadas pela OMS são oferecidas gratuitamente no Sistema Único de Saúde (SUS). Também foi possível erradicar a poliomielite e a transmissão sustentada, transmissão do vírus entre a população do próprio país, da rubéola e do sarampo (APS et al., 2018).

2.0.1 Rede de Frio

A rede de frio envolve toda a estrutura física, técnica e administrativa relativa à vacinação da população. É orientada pela coordenação-geral do PNI e é organizada em três níveis de gestão: nacional, estadual e municipal. O processo logístico que coordena a distribuição, o manuseio e o armazenamento das vacinas é chamada de cadeia de frio (MINISTÉRIO DA SAÚDE, 2022b).

2.0.1.1 A cadeia de frio

Para que a população receba a imunização de forma correta e com seu maior potencial é necessária a adequada manutenção e monitoramento da cadeia de frio de conservação de vacinas. Especificamente, uma cadeia de suprimentos é o sistema de organizações, pessoas, tecnologias, atividades, informações e recursos envolvidos em transportar e armazenar um produto do fornecedor para o consumidor. Vacinas são produtos biológicos que com o tempo e com a temperatura perdem sua capacidade de imunização. Assim, de forma a manter sua eficiência, a maioria das vacinas deve ser armazenada em ambientes refrigerados com temperatura entre 2°C e 8°C (HANSON et al., 2017).

Dentre as cadeias de suprimentos existe a cadeia de frio. Esta possui temperatura controlada, resultado de necessidades relacionadas à distribuição, armazenamento e processamento de produtos sensíveis à temperatura. Como as vacinas são materiais termolábeis, precisam deste cuidado. Devido às regras de agências governamentais, a cadeia de frio tem se tornado uma das principais preocupações dos campos da farmácia e biotecnologia (EMOND, 2008).

Diversos fatores influenciam na temperatura de armazenamento, dentre eles: a quantidade de vacinas armazenadas no refrigerador; o fluxo de ar do ambiente; a posição das vacinas (perto da unidade de refrigeração as vacinas estarão em temperaturas menores se comparar ao centro do refrigerador); possíveis objetos armazenados junto das vacinas; a quantidade de vezes a qual o compartimento é aberto; o tipo do refrigerador (se possui congelador, o tipo de prateleiras, a variação do controle de temperatura). Estes parâmetros têm papel importante na manutenção da temperatura ideal, por isso colocar o refrigerador na faixa de temperatura não é suficiente para garantir que as vacinas mantenham sua potência (TSANG et al., 2018).

O controle da cadeia de frio é importante não apenas para cumprir as regulamentações como também para garantir segurança na sua utilização. Nestas cadeias uma falha no sistema pode acarretar em mudanças bruscas de temperatura, necessitando uma resposta rápida o suficiente para prevenir danos aos produtos durante todos os estágios da rede. Os métodos mais utilizados para previsão de temperatura se baseiam em previsões de longo termo ou dependem de vários sensores em posições determinadas. Intervenções nestes sistemas são requeridas imediatamente, portanto, tem-se a necessidade de um sistema que fornece informações em tempo real (EMOND, 2008).

Devido a grande sensibilidade dos imunobiológicos às alterações na temperatura, aos possíveis desvios de fluxo de atividades de distribuição e armazenamento e os aspectos de segurança e qualidade que este tipo de produto exige, o Ministério da Saúde elaborou um *Manual de Rede de Frio*. O manual tem como

objetivo definir um referencial, teórico e operacional, capaz de produzir uma unidade de procedimentos desde o laboratório produtor até o momento que as vacinas são ofertadas à população brasileira (MINISTÉRIO DA SAÚDE, 2017).

Entre os requisitos e equipamentos necessários para o armazenamento dos imunobiológicos relacionam-se: caixas térmicas, utilizadas para transporte, atividades de rotina e campanhas; Instrumentos para medição de temperatura; bobinas reutilizáveis, recipientes constituídos de material plástico, contendo gel em concentração não tóxica e água ou apenas água; câmaras refrigeradas com faixa de operação de +2°C e +8°C (MINISTÉRIO DA SAÚDE, 2017).

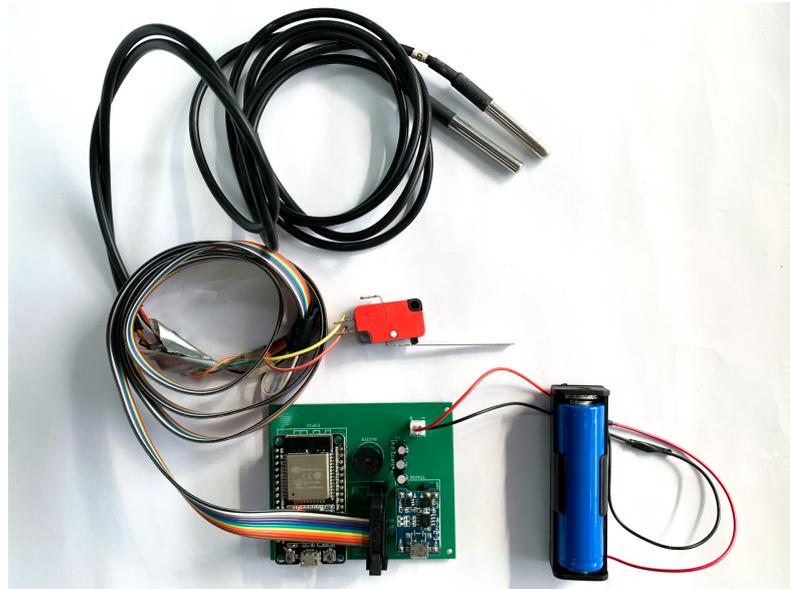
Apesar de ser equipamento necessário para o correto armazenamento das vacinas, as câmaras refrigeradas não são realidade em diversas Unidades Básicas de Saúde. O Manual de Rede de Frio não indica a utilização de refrigeradores domésticos, uma vez que estes não possuem precisão no ajuste da temperatura. Outro importante componente da rede de frio é a caixa térmica, utilizadas no transporte e no uso diário (MINISTÉRIO DA SAÚDE, 2017). Ambos são locais onde o dispositivo pode ser aplicado.

2.0.2 O dispositivo de Monitoramento de Temperatura

Para que não seja necessária a troca de geladeiras domésticas por câmaras refrigeradas, devido ao seu custo elevado, foi produzido um sistema embarcado para monitoramento de temperatura. O protótipo é composto por: um módulo microcontrolador ESP32, responsável pela aquisição e envio dos dados para o usuário; dois sensores de temperatura Ds18b20, para a medição das temperaturas externas e internas; um módulo TP4056, utilizado para carregar a bateria; uma bateria, para manter o sistema funcionando apesar de quedas de energia; um botão, para detectar se a porta está aberta ou fechada; e um *buzzer*, utilizado como alarme sonoro caso a porta fique aberta por um período indesejado.

Durante os testes de aquisição de dados foi constatado que a temperatura interna da geladeira varia de forma mais súbita que a temperatura das vacinas. Para isto foi pensado em um *buffer* de temperatura, um recipiente com um líquido no interior, o qual simularia o efeito amortecedor causado pela vacina. Neste caso, este efeito se refere à desaceleração da queda da temperatura da vacina em relação à queda da temperatura do ar no interior da geladeira. No entanto, o Manual de Rede de frio proíbe a utilização de frascos no sensor, tanto com quanto sem líquido (MINISTÉRIO DA SAÚDE, 2017). Portanto visando uma leitura mais estável, o projeto propõe simular o *buffer* físico.

Figura 1 – Sistema embarcado para monitoramento de temperatura.



Fonte: Elaborado pelo autor.

2.0.3 Inteligência Artificial

Algoritmos de aprendizado de máquina são poderosas ferramentas de generalização e previsão. A crescente facilidade de aquisição e armazenamento de dados proporciona algoritmos cada vez melhores, uma vez que sua acurácia aumenta com uma maior quantidade de informação. O aprendizado de máquina engloba diferentes modelos, que podem ser implementados de diferentes formas. Alguns métodos populares são as redes neurais artificiais, as árvores de decisão, máquina de vetores, florestas aleatórias, entre outros. A escolha do modelo depende da sua aplicação, da precisão e dos recursos disponíveis. Há também a possibilidade de mesclar diferentes métodos (BURRELL, 2016).

Um dos pioneiros no estudo acerca da inteligência artificial foi John McCarthy. McCarthy organizou em 1956 uma pesquisa de dez semanas para a Conferência de Dartmouth. Essa pesquisa foi chamada de *Estudo da Inteligência Artificial* e foi a primeira vez que o termo foi utilizado. Os objetivos do estudo eram ambiciosos. Os pesquisadores entendiam que todos os aspectos do aprendizado e qualquer outro campo da inteligência poderiam ser, a princípio, descritos tão precisamente que uma máquina poderia ser desenvolvida para simular este aprendizado (TAULLI, 2019).

Durante a conferência, Allen Newell, Cliff Shaw e Herbert Simon apresentaram um programa denominado de *Logic Theorist*, ou em português, teórico da lógica, considerado por muitos o primeiro programa de inteligência artificial. Ele tinha como objetivo resolver teoremas matemáticos do livro *Principia Mathematica*, de Alfred North Whitehead e Bertrand Russell.

Entre 1956 e 1974 existiam duas vertentes sobre a IA. Uma era liderada por

Minsky, que acreditava na necessidade de sistemas simbólicos, o que significava que a IA deveria se basear na lógica tradicional de programação, utilizando declarações *se-então-senão*. Por outro lado, Frank Rosenblatt entendia que os sistemas deveriam ser mais próximos de redes neurais artificiais. Este campo do conhecimento foi chamado de *conexionismo* e seriam capazes de aprender conforme dados são absorvidos (APS et al., 2018).

Em 1957, Rosenblatt criou o primeiro programa de computador para isso, chamado de Mark 1 Perceptron. Ele incluía um sistema de câmeras para diferenciar duas imagens de 20x20 pixels. O componente que simula um neurônio no algoritmo foi chamado por Rosenblatt de perceptron. O jornal *The New York Times* inclusive publicou um artigo com a manchete *Psicólogo mostra o embrião de um computador designado para ler e aprender* (THE NEW YORK TIMES, 1958).

Em seguida, essa área passou por um período nos anos 70 conhecido como o primeiro *Inverno da IA*. O entusiasmo com a nova tecnologia diminuiu e os principais investidores congelaram gastos com inteligência artificial por não conseguir as inovações que almejavam. As técnicas não cumpriram os objetivos prometidos, em parte por causa das limitações de *hardware*. Por exemplo o minicomputador PDP-11/45, comum entre as pesquisas de IA, possuía uma memória de até 128 Kwords. Após esse período, Geoffrey Hinton trabalhou para desenvolver a base do que hoje conhecemos como *deep learning*. Em 1986 ele publicou um inovador artigo junto de David Rumelhart e Ronald J. Williams intitulado *Aprendendo representações a partir da retropropagação do erro*. O trabalho apresentou processos chave para a utilização da retropropagação em redes neurais artificiais, resultando em um aumento significativo da acurácia (TAULLI, 2019).

O segundo inverno na área de inteligência artificial veio no fim da década de 1980, quando estes sistemas encontraram novamente uma barreira de *hardware* e perderam o interesse no mundo dos negócios. Este inverno durou até meados da década de 1990. Após este inverno novos conceitos, teorias e modelos foram desenvolvidos, impulsionados principalmente pelo grande crescimento dos conjuntos de dados, sendo a internet fator indispensável para este crescimento. Outro fator foi a nova tecnologia das unidades de processamento gráfico (GPUs), desenvolvidas principalmente para processamento gráfico em jogos. Esta arquitetura se tornou ideal para a inteligência artificial. A capacidade de processamento em paralelo das GPUs possibilitou uma melhora significativa do tempo de computação dos modelos (TAULLI, 2019).

Existem diversos métodos diferentes de inteligência artificial. Porém neste trabalho veremos quatro. Os algoritmos selecionados têm treinamento supervisionado, isto é, eles precisam de dados de treinamento e suas respectivas soluções. Eles são: as árvores de decisão, as florestas aleatórias, as máquinas de vetores de suporte e as

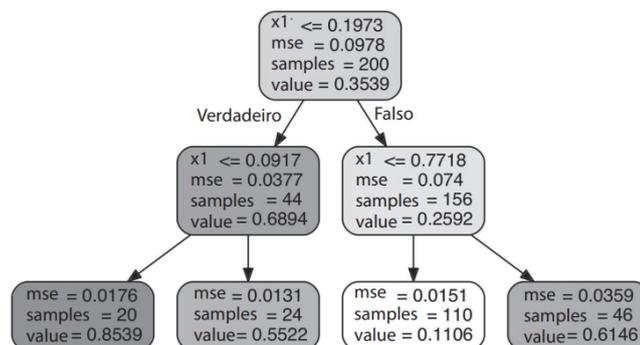
redes neurais artificiais.

2.0.4 Árvores de Decisão

As árvores de decisão são algoritmos poderosos capazes de executar tanto tarefas de classificação quanto de regressão. Elas fazem poucos pressupostos acerca dos dados de treinamento e conseguem se adaptar facilmente aos dados de entrada. Se for deixada sem restrições uma árvore provavelmente ficará sobreajustada. Assim será capaz de se adaptar com muita precisão aos dados de treinamento, mas será ineficaz na previsão de novos resultados (GÉRON, 2019).

Uma árvore de decisão começa em um nó raiz e a partir deste há uma série de nós internos. Em cada nó da árvore acontece um algoritmo de decisão. Após terminarem todas as decisões necessárias o modelo encontra sua previsão. Este método é simples de compreender e tem bom desempenho com grandes conjuntos de dados. Mas quanto mais profunda é a árvore mais complexidade é adicionada e resultará em uma menor performance em relação às complexidades de tempo e espaço (TAULLI, 2019).

Figura 2 – Exemplo de uma árvore de decisão para regressão.



Fonte: (GÉRON, 2019).

2.0.5 Florestas Aleatórias

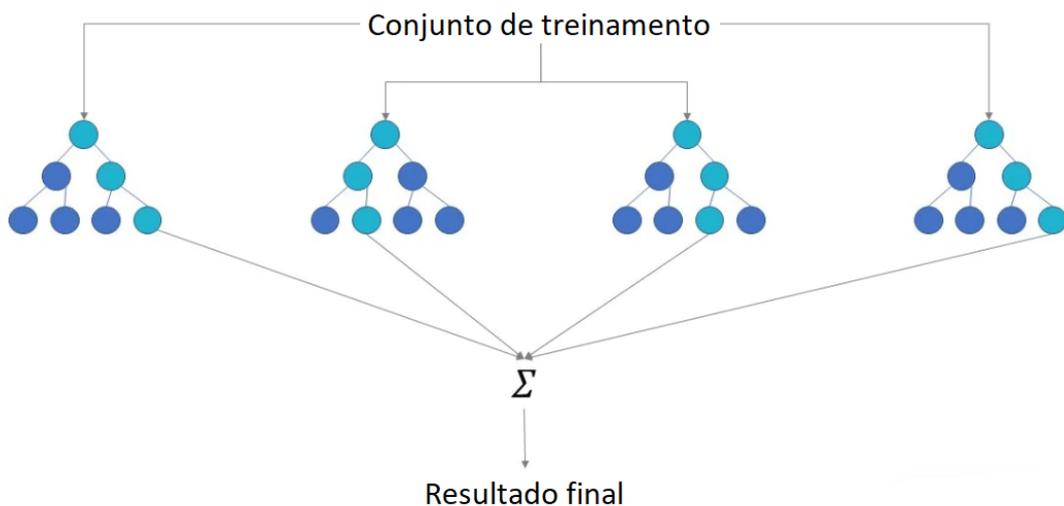
Para ter maior precisão, uma escolha viável é agregar previsões de um conjunto de algoritmos e a partir destas encontrar sua resposta. Um conjunto de previsores é chamado de *ensemble* e um algoritmo que aprende a partir das previsões destes é chamado de *ensemble method* (GÉRON, 2019).

O método de *ensemble* mais conhecido é o *bagging*. Neste uma amostragem aleatória em um conjunto de treinamento é realizada com substituição. Depois de várias amostragens feitas, os previsores são treinados independentemente. E para tarefas de regressão, a solução final é a média dos resultados de cada modelo. O que

contribui com a redução da variância dentro de um conjunto de dados (IBM CLOUD EDUCATION, 2020).

Ao treinar algumas árvores de decisão, cada uma com um conjunto de dados aleatórios. E em seguida, comparar seus resultados para encontrar uma resposta final, temos um *ensemble* de árvores de decisão. Este é chamado de florestas aleatórias (GÉRON, 2019).

Figura 3 – Exemplo de uma floresta aleatória.



Fonte: Adaptado de (IBM CLOUD EDUCATION, 2020).

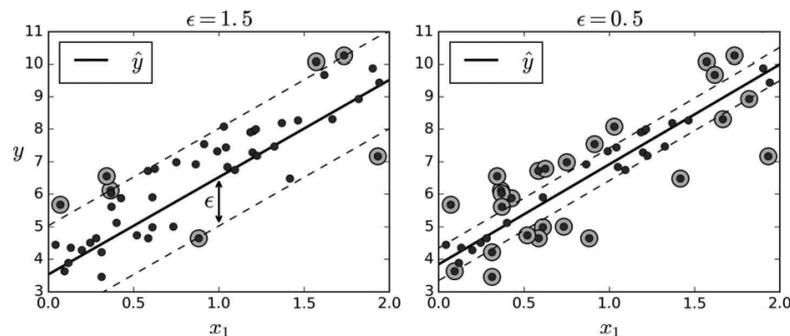
Ao desenvolver as árvores de decisão no seu interior o algoritmo introduz uma aleatoriedade extra. Enquanto as árvores normalmente buscam a melhor característica ao dividir um nó. As florestas buscam entre um subconjunto aleatório das características. Este método resulta em uma maior diversidade de *ensembles*. O que pode produzir um resultado melhor (GÉRON, 2019).

2.0.6 Máquinas de Vetores de Suporte

Assim como as árvores de decisão, as Máquinas de Vetores de Suporte (SVM) são capazes de realizar tanto tarefas de classificação quanto de regressão. Na regressão o método tenta preencher a maior via possível entre duas classes e limita as violações da margem. A largura das margens é definida por um hiperparâmetro ϵ , porém alterar a quantidade de instâncias dentro da margem não afeta o resultado final, assim o modelo é insensível a este hiperparâmetro. O hiperparâmetro de regularização informa o quanto a SVM tentará se aproximar do valor esperado, quanto maior o seu valor mais regularizado será o resultado. A figura 4 mostra um exemplo de regressão linear feita por um algoritmo SVM. Para modelos não lineares é utilizada outra função

de ativação. Entre alguns exemplos de funções estão a polinomial, a sigmoide e a gaussiana (GÉRON, 2019).

Figura 4 – Exemplo de uma SVM de regressão.



Fonte: (GÉRON, 2019).

2.0.7 Redes Neurais Artificiais

As redes neurais artificiais (RNAs) são modelos matemáticos e computacionais baseados em redes neurais biológicas. Consistem em grupos de neurônios artificiais interconectados os quais processam informações. Na maioria dos casos as RNAs são adaptáveis e mudam sua estrutura de acordo com dados internos e externos durante a fase de aprendizado. São ferramentas estatísticas não lineares de modelagem de dados. E podem ser utilizadas para modelar relações complexas entre entradas e saídas com o objetivo de encontrar padrões nestas informações (CHEN; SHAW, 2011).

Existem diversos tipos de RNAs, desde uma das mais simples arquiteturas, como o perceptron, até os modelos mais complexos de aprendizado profundo. Elas são versáteis, escaláveis e poderosas, capazes de lidar com tarefas altamente complexas, tornando-as o principal método de aprendizado de máquina (GÉRON, 2019).

O perceptron foi criado baseado em um neurônio artificial, a diferença entre eles é que as entradas e saídas daquele são números e do neurônio são valores binários (ligado ou desligado). Ele utiliza uma função degrau nas suas entradas ponderadas para encontrar um resultado. O perceptron possui sérias fraquezas. Entretanto, ao serem agregados e empilhados algumas destas limitações são eliminadas e a rede resultante é chamada de Perceptron Multicamada (MLP). Um MLP é composto por uma camada de entrada, camadas ocultas e uma camada de saída. Uma rede neural é considerada profunda quando possui duas ou mais camadas ocultas (GÉRON, 2019). Porém, apesar de muito poderoso um modelo de aprendizado profundo pode facilmente ter milhões de parâmetros dentro de uma série de camadas, o que torna extremamente difícil um entendimento claro do que está acontecendo no seu interior (TAULLI, 2019).

Treinar redes neurais artificiais foi problema até 1986, quando D. E.

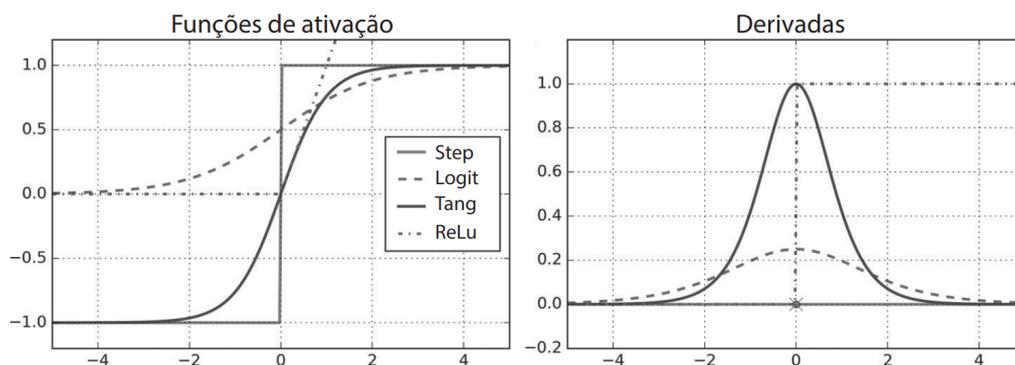
Rumelhart et al. publicaram um artigo apresentando o algoritmo de treinamento de retropropagação. Neste modelo para cada instância de treinamento é feita uma previsão, é medido o seu erro, o algoritmo passa por cada camada no reverso para verificar a contribuição de cada instância no erro e por último seus pesos são ajustados. Estes são os passos que acontecem em cada época de treinamento (RUMELHART; HINTON; WILLIAMS, 1986).

Para que haja um gradiente no ajuste dos erros foi feita uma alteração na arquitetura da rede neural, substituindo a função de ativação. Como a função degrau utilizada por cada perceptron não possui curvas suaves, não é capaz de produzir um gradiente. Portanto foram propostas outras funções de ativação.

2.0.7.1 Funções de ativação

As funções logística e tangente hiperbólica possuem formato parecido, são contínuas e diferenciáveis. A principal diferença é que a primeira varia entre 0 e 1, enquanto a segunda varia entre -1 e 1, resultando nas saídas de cada camada centradas em torno de zero, o que normalmente ajuda na convergência. Outra função comumente utilizada é a unidade linear retificada (ReLU), que apesar de não ser diferenciável em zero, é contínua e sua principal vantagem em relação às anteriores é a rapidez no cálculo (GÉRON, 2019).

Figura 5 – Funções de ativação e suas derivadas.

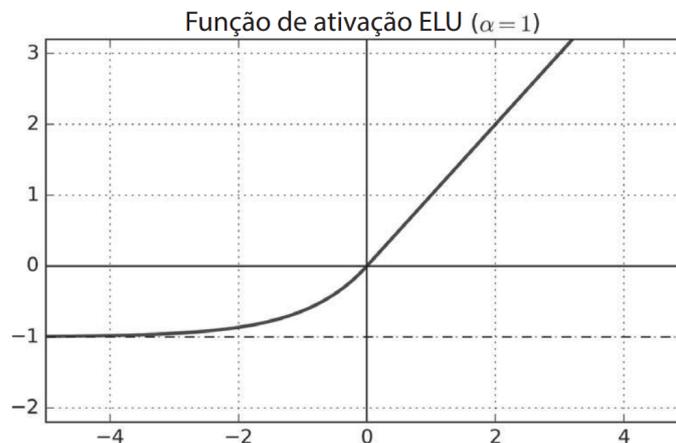


Fonte: (GÉRON, 2019).

Outra função de ativação que obteve destaque foi a unidade exponencial linear (ELU). Ela reduziu o tempo de treinamento e melhorou o resultado das redes em relação ao conjunto de testes. A principal desvantagem desta é a velocidade mais lenta do seu cálculo, porém esta é compensada por uma convergência mais rápida. Como ela assume valores negativos quando sua entrada é menor que zero, ajuda a evitar que os gradientes das camadas inferiores se aproximem muito de zero dificultando a convergência para uma boa solução, efeito chamado de *vanishing gradients* (GÉRON,

2019).

Figura 6 – Função ELU.



Fonte: (GÉRON, 2019).

2.0.7.2 Inicializadores

Ao analisar os extremos das funções de ativação logística, tangente hiperbólica e o extremo negativo da ELU, é possível identificar que quanto maiores ficam as entradas as funções se aproximam de uma assíntota e sua derivada se aproxima de zero. O seu efeito é a diluição do gradiente de retropropagação e consequentemente estagnando o aprendizado do modelo (GÉRON, 2019). A inicialização Glorot reduz este problema. Ela propõe que os pesos de conexão sejam inicializados de forma aleatória a partir de uma distribuição normal centralizada em zero (GLOROT; BENGIO, 2010).

2.0.7.3 Otimizadores

Outra técnica a ser levada em consideração é a otimização. Os otimizadores são utilizados para aumentar a velocidade de treinamento de uma rede neural e normalmente utilizam as derivadas parciais de primeira ordem para calcular seu gradiente. Um dos otimizadores mais simples e mais conhecidos é o Gradiente Descendente. Ele funciona atualizando os pesos subtraindo o gradiente calculado pela função de custo dos pesos de cada neurônio multiplicados pela taxa de aprendizado. Este método não utiliza os gradientes anteriores no seu cálculo, o que pode desacelerar o processo. Outros métodos são mais eficientes, como a otimização *Momentum*, a qual utiliza um *vetor momentum* no seu cálculo, e com isso faz com que o gradiente seja significativamente mais rápido. Há também a otimização RMSProp, que acumula os gradientes mais recentes por meio de um decaimento exponencial. Exceto em problemas simples este é um otimizador que funciona melhor se comparado aos

anteriores. Entretanto, existe o otimizador Adam, o qual combina as ideias dos otimizadores *Momentum* e RMSProp e é um algoritmo de taxa de aprendizado adaptativa, o que requer menos ajuste em relação a esta taxa se comparado aos outros (GÉRON, 2019).

2.0.8 Métricas de Desempenho

Mesmo se um algoritmo funciona dentro de um requisito proposto, é necessário analisar outras métricas de avaliação. Uma única métrica não é suficiente para descrever a maior parte dos eventos do mundo real (DOSHI-VELEZ; KIM, 2017). Para avaliar se uma solução é a melhor encontrada para o problema podemos trabalhar com algumas métricas. Um algoritmo é considerado completo quando é garantido encontrar uma solução, quando ela existe. É considerado preciso de acordo com a razão de previsões e valores. Também é possível analisar as complexidades de tempo e espaço, na qual a complexidade de tempo é o quanto demora para o algoritmo encontrar uma solução e a complexidade de espaço é a quantidade de memória utilizada por este para sua execução (LUCCI; KOPEC; MUSA, 2022).

Uma medida de desempenho bastante utilizada para problemas de regressão é o Erro Quadrático Médio (MSE). Essa medida mostra quantos erros foram gerados pelo sistema em suas previsões, com um peso maior para erros maiores. A equação 1 mostra a fórmula matemática para calcular o MSE (GÉRON, 2019).

$$MSE(X, h) = \frac{1}{m} \sum_{i=1}^m \left(h(x^{(i)}) - y^{(i)} \right)^2, \quad (1)$$

na qual m é o número de instâncias do conjunto avaliado, $x^{(i)}$ é o vetor de características da instância, $y^{(i)}$ é o valor desejado para a saída do sistema, X é a matriz de entrada, ou seja, de características do sistema, e h é a função de previsão do sistema. Quando esta recebe o vetor $x^{(i)}$ retorna um valor previsto para a saída \hat{y} . Portanto, para cada previsão, esta métrica calcula o erro entre a saída prevista e a saída real e o eleva ao quadrado, para em seguida calcular a média de todos os erros encontrados.

Em redes neurais artificiais as funções de perda ajudam a otimizar a performance do modelo. Normalmente elas são utilizadas para medir alguma penalidade que pode ocorrer nas previsões, como o afastamento das previsões em relação às saídas esperadas. Elas são utilizadas para medir o custo médio de cada treinamento em relação a todo o conjunto de treinamento (GÉRON, 2019).

3 MÉTODO

Para a execução deste trabalho foram aplicados uma série de métodos de inteligência artificial e em seguida foi feita uma análise com base em sua eficiência. Para a verificação da sua eficiência foram levados em conta as métricas do erro médio quadrático e a perda. Classificação e regressão estão entre as funções mais comuns de métodos de inteligência artificial. A classificação é responsável por identificar um conjunto de dados dentre categorias especificadas. A regressão encontra um valor de saída de acordo com um conjunto de dados de entrada. Para a tarefa em questão os métodos devem ser capazes de realizar a tarefa de regressão. Os métodos de inteligência artificial escolhidos foram as redes neurais artificiais, a árvore de decisão, as máquinas de vetores e florestas aleatórias.

O sistema embarcado utilizado é capaz de coletar os seguintes dados: a temperatura interna da geladeira ou caixa térmica; a temperatura externa do local; o estado da porta, se está aberta ou fechada; e o tempo entre cada medida. Foram feitos testes em uma geladeira doméstica, respeitando a posição do termômetro definida pelo Manual de Rede de Frios. Isto é sem que haja contato do sensor com qualquer outra estrutura ou equipamento (MINISTÉRIO DA SAÚDE, 2017). Para cada teste foram alterados: o termostato da geladeira entre os valores mínimo, médio e máximo; o intervalo de abertura de porta; e o intervalo de tempo no qual a geladeira ficava aberta. A partir destas informações foram obtidos os conjuntos de dados de treinamento dos modelos.

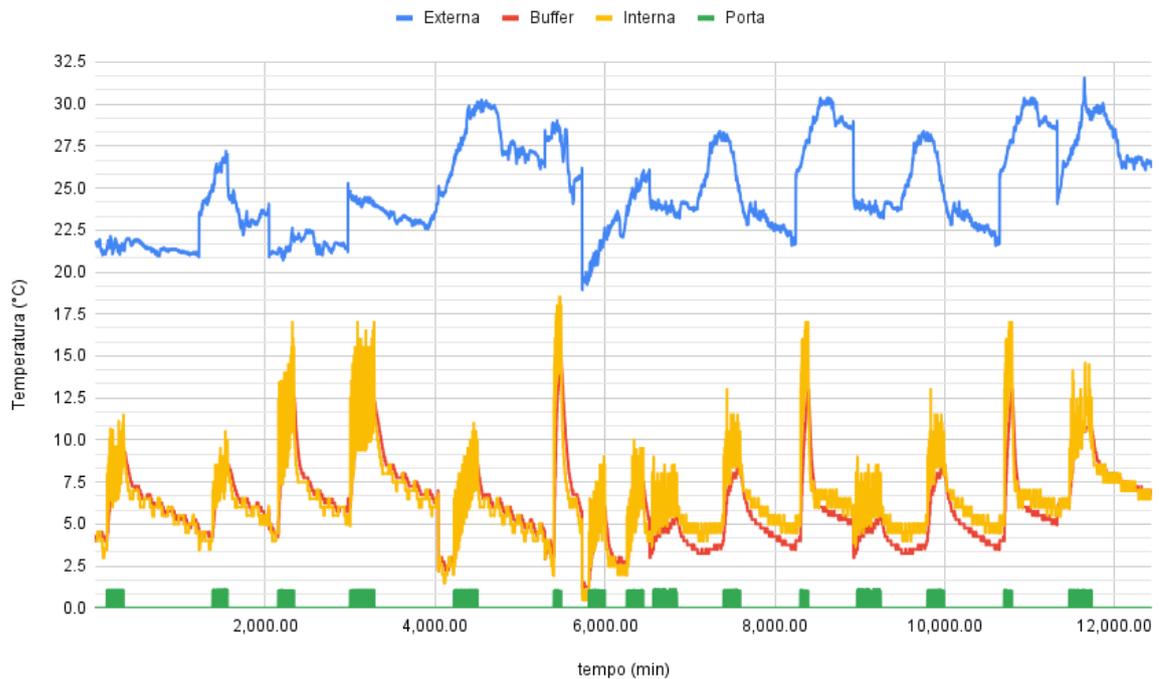
Analisando os dados de treinamento é possível treinar os modelos utilizando os dados coletados, sendo a principal informação a temperatura interna da geladeira.

3.0.1 Dados de Treinamento

Foram coletadas um total de 7573 amostras, sendo que cada amostra contém: os dados de temperatura interna do ar da geladeira; temperatura do ar no exterior da geladeira; temperatura do *buffer*, foi utilizado 5ml de glicerina diluída em 15 ml de água como líquido amortecedor; a informação de porta, aberta ou fechada; e o intervalo de tempo entre as coletas de dados. Foram realizados 2369 coletas com o termostato no médio, 2960 com o termostato no máximo e 2244 com o termostato no mínimo. Os testes variavam tanto no intervalo entre as aberturas de porta, que foram de 15, 10 ou 5 minutos, quanto no tempo que a porta ficava aberta, que foram de 60, 30 ou 15 segundos. A seguir podemos ver a Figura 7 a qual mostra todos estes dados concatenados.

Para que um dos dados não tenha prioridade em relação aos outros no

Figura 7 – Gráfico dos dados de treinamento.



Fonte: Elaborado pelo autor (2022).

treinamento é importante aplicar o escalonamento das características. Normalmente os algoritmos de aprendizado de máquina não funcionam de forma correta quando os dados possuem escalas muito diferentes (GÉRON, 2019).

Foram testados duas maneiras dos atributos obterem a mesma escala. O primeiro método utiliza a equação 2 para padronizar os dados de acordo com a sua variância (LU et al., 2017), centralizando seu resultado em zero. Já o segundo utiliza uma técnica mais comum, escala mínimo máximo, também conhecida como normalização, a qual pode ser vista na equação 3 e seu resultado termina entre 0 e 1 (GÉRON, 2019).

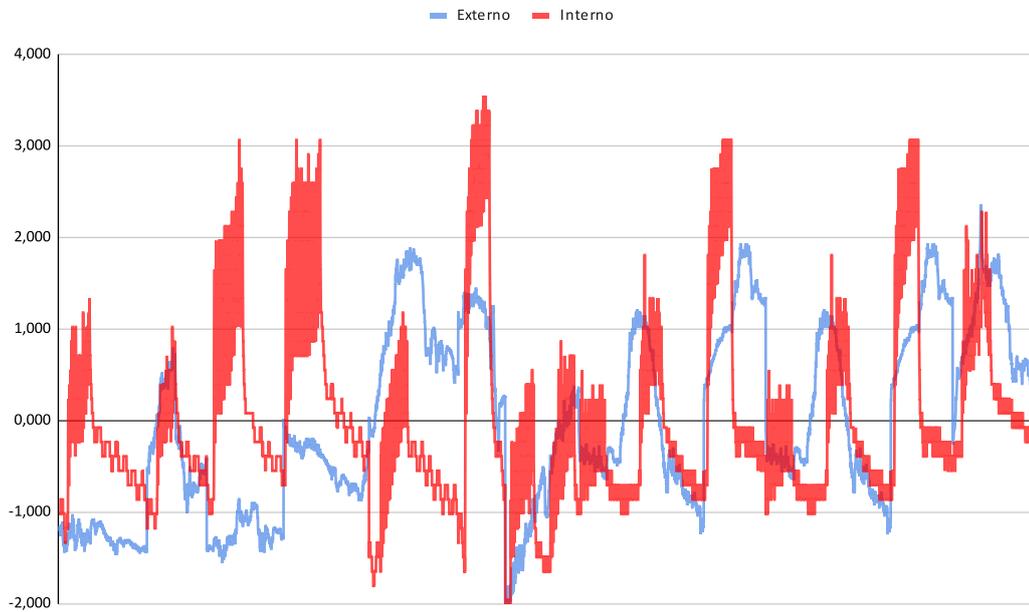
$$\tilde{x} = \frac{x - \langle x \rangle}{\sqrt{\langle (x - \langle x \rangle)^2 \rangle}} \quad (2)$$

$$\tilde{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3)$$

A característica da porta não foi alterada no método de normalização uma vez que a informação de porta aberta ou fechada já está entre os valores de 0 e 1. Na equação 2 os parenteses angulares significam o valor médio. Cada técnica foi aplicada em cada característica separadamente, assim o resultado dos métodos pode ser visualizado nas figuras 8 e 9. Foi escolhido não representar no gráfico as características da porta e do intervalo de tempo para facilitar a visualização do gráfico.

O intervalo de coleta de dados é de 120 segundos enquanto a porta está fechada e de 10 segundos enquanto a porta está aberta.

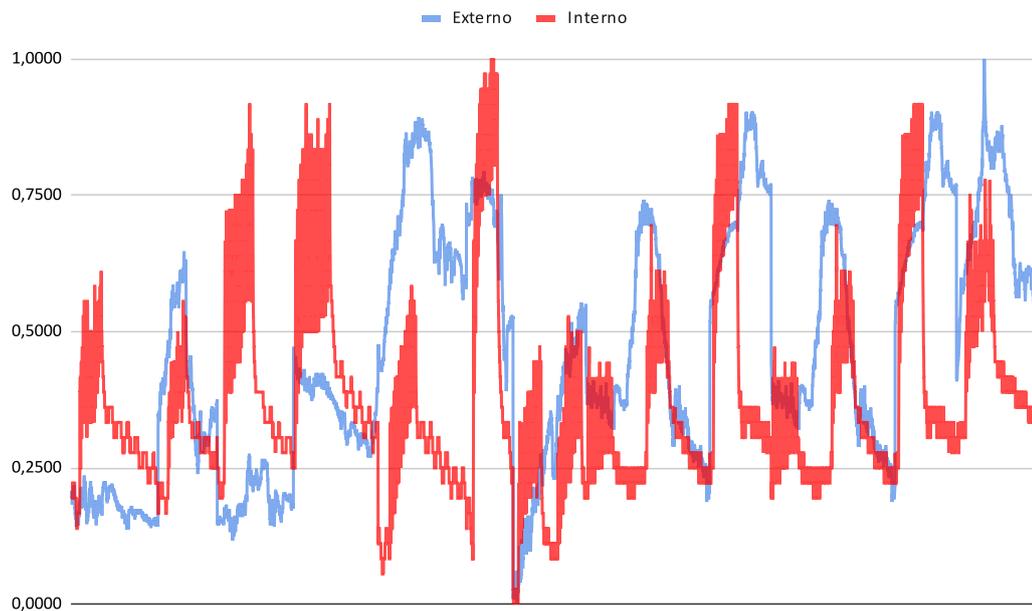
Figura 8 – Gráfico dos dados padronizados de acordo com a equação 2.



Fonte: Elaborado pelo autor (2022).

Para o treinamento de cada modelo as características do sistema foram separadas em nove conjuntos de dados. Os conjuntos possuem os seguintes dados: a temperatura interna T_{int} ; a temperatura externa T_{ext} ; o intervalo de tempo entre as coletas das amostras Δt ; o estado da porta P ; e a diferença entre a temperatura interna e externa ΔT . Quais dados cada conjunto possui podem ser vistos na tabela 1. É interessante testar a importância de cada característica. Uma vez que um número menor de características representa menos memória necessária para implementar no sistema embarcado.

Figura 9 – Gráfico dos dados normalizados de acordo com a equação 3.



Fonte: Elaborado pelo autor (2022).

Tabela 1 – Definição dos conjuntos de treinamento

Conjuntos de treinamento	Dados
Conjunto 1	T_{int} ,
Conjunto 2	T_{int} , T_{ext} ,
Conjunto 3	T_{int} , T_{ext} , Δt
Conjunto 4	T_{int} , T_{ext} , Δt , P
Conjunto 5	T_{int} , T_{ext} , Δt , P , ΔT
Conjunto 6	T_{int} , T_{ext} , Δt , ΔT
Conjunto 7	T_{int} , T_{ext} , ΔT
Conjunto 8	T_{int} , ΔT
Conjunto 9	ΔT

Fonte: Elaborado pelo autor (2022).

Portanto, além das características de cada técnica de inteligência artificial, todas as execuções variavam no método de obter a mesma escala dos dados e nos conjuntos de treinamento.

3.0.2 Árvores de Regressão

Para as Árvores de Regressão o principal parâmetro a ser modificado é sua profundidade, assim foram feitos 440 treinamentos. Nos treinamentos foram realizadas mudanças no conjunto de dados treinados e na profundidade da árvore.

3.0.3 Florestas Aleatórias

Para as florestas aleatórias seu principal parâmetro é a quantidade de árvores dentro do ensemble. Os treinamentos foram realizados alterando o número de estimadores, os conjuntos de treinamento e os dados, entre normalizados e padronizados.

3.0.4 Máquina de Vetores

As SVMs possuem dois parâmetros mais importantes para serem alterados nas tarefas de regressão. Eles são a função de ativação e o parâmetro "C", parâmetro de regularização, e para o caso da função de ativação ser a polinomial há um parâmetro a mais que é o grau da função. Foram treinados algoritmos com as funções de ativação polinomial, sigmoide e gaussiana. Para a função polinomial seu grau foi variado entre 3 e 6. Também foram variados os dados de entrada, mudando entre conjuntos de treinamento e a técnica de obter a mesma escala.

3.0.5 Redes Neurais Artificiais

Para o desenvolvimento, treinamento e avaliação das RNAs foram criados alguns modelos de redes neurais artificiais diferentes. Cada modelo foi testado e comparado variando os seus dados de entrada e o número de épocas alvo. Cada modelo é uma rede neural do tipo *feedforward*, e treinado com o algoritmo de retropropagação

3.0.5.1 Modelos

Foram criados seis modelos de redes neurais artificiais densas para treinamento e comparação de resultados. Os modelos variam entre os hiperparâmetros ajustáveis de uma RNA: na quantidade de camadas ocultas, em quantos perceptrons tem em cada camada e na função de ativação de cada camada (GÉRON, 2019). Todos os modelos podem ser observados nas tabelas 2 a 7.

Tabela 2 – Modelo 1

Camadas	Número de Neurônios	Função de ativação
Camada 1	80	elu
Camada 2	50	tanh
Camada 3	30	relu
Camada 4	10	relu
Camada 5	8	relu
Camada 6	4	relu
Camada 7	1	linear

Fonte: Elaborado pelo autor (2022).

Tabela 3 – Modelo 2

Camadas	Número de Neurônios	Função de ativação
Camada 1	80	ELU
Camada 2	90	tanh
Camada 3	100	tanh
Camada 4	90	tanh
Camada 5	100	tanh
Camada 6	90	tanh
Camada 7	100	tanh
Camada 8	100	ReLU
Camada 9	80	ReLU
Camada 10	100	ELU
Camada 11	80	ELU
Camada 12	1	linear

Fonte: Elaborado pelo autor (2022).

Tabela 4 – Modelo 3

Camadas	Número de Neurônios	Função de ativação
Camada 1	600	ELU
Camada 2	300	ELU
Camada 3	150	ELU
Camada 4	75	ELU
Camada 5	50	ELU
Camada 6	1	–

Fonte: Elaborado pelo autor (2022).

Tabela 5 – Modelo 4

Camadas	Número de Neurônios	Função de ativação
Camada 1	110	ELU
Camada 2	110	ReLU
Camada 3	110	tanh
Camada 4	110	tanh
Camada 5	110	ReLU
Camada 6	1	–

Fonte: Elaborado pelo autor (2022).

3.0.6 Outros Parâmetros

Além dos hiperparâmetros já citados podemos variar o otimizador e o inicializador para cada modelo. O inicializador escolhido é o Glorot. Conforme elucidado no capítulo anterior, esta técnica ajuda a aliviar o efeito da diluição do gradiente de

Tabela 6 – Modelo 5

Camadas	Número de Neurônios	Função de ativação
Camada 1	60	ELU
Camada 2	40	Relu
Camada 3	30	Relu
Camada 4	20	Relu
Camada 5	10	Relu
Camada 6	5	Relu
Camada 7	1	ELU

Fonte: Elaborado pelo autor (2022).

Tabela 7 – Modelo 6

Camadas	Número de Neurônios	Função de ativação
Camada 1	4	ELU
Camada 2	6	ELU
Camada 3	5	ELU
Camada 4	4	ELU
Camada 5	3	ELU
Camada 6	2	ELU
Camada 7	1	ELU

Fonte: Elaborado pelo autor (2022).

retropropagação. Para o otimizador, foi escolhido o otimizador Adam, pois ele combina as vantagens do otimizador Momentum e RMSProp.

Também foi utilizada a função *early_stop* a qual tem como objetivo minimizar o sobreajuste dos dados. Nela é definido o número de épocas alvo, a partir da época alvo é monitorada a função de custo. Se o algoritmo identificar que a função de custo não está mais diminuindo, a função *early_stop* termina o treinamento da rede (KERAS, 2022).

4 ANÁLISE DE DADOS

Serão apresentados e comentados neste capítulo os resultados de cada técnica de inteligência artificial. Estes foram obtidos respeitando a metodologia descrita no capítulo anterior. A análise dos resultados dos métodos de inteligência artificial é feita comparando a métrica do erro médio quadrático. É importante frisar que, para gerar todos os gráficos, foram utilizados dados de validação e não de treinamento. E os gráficos não possuem unidades uma vez que o erro quadrático médio, a profundidade da árvore, os conjuntos de treinamento, o número de estimadores, o parâmetro de regularização e o número de épocas são valores adimensionais.

4.0.1 Árvores de Decisão

Foram realizadas diversas execuções do algoritmo de árvores de decisão com a tarefa de regressão e gerados gráficos com suas saídas. A análise dos resultados é feita a partir do impacto causado pela variação da profundidade e dos conjuntos de treinamento.

4.0.1.1 Profundidade

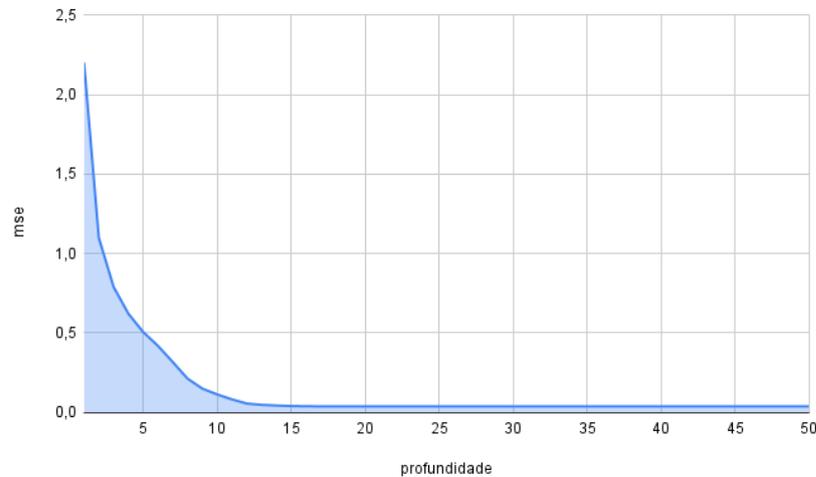
A profundidade foi alterada unitariamente entre os valores 1 e 50. Para cada valor de profundidade a rede foi treinada duas ou mais vezes, com os dados de entrada sem nenhum tipo de tratamento. A partir destas execuções foram encontrados 180 resultados.

Com base na figura 10, é notório o melhor ajuste da árvore conforme o aumento da profundidade. É interessante observar também que a partir de uma profundidade próxima de 15 a árvore não é mais capaz de diminuir o erro médio quadrático de forma significativa.

4.0.1.2 Conjuntos de treinamento

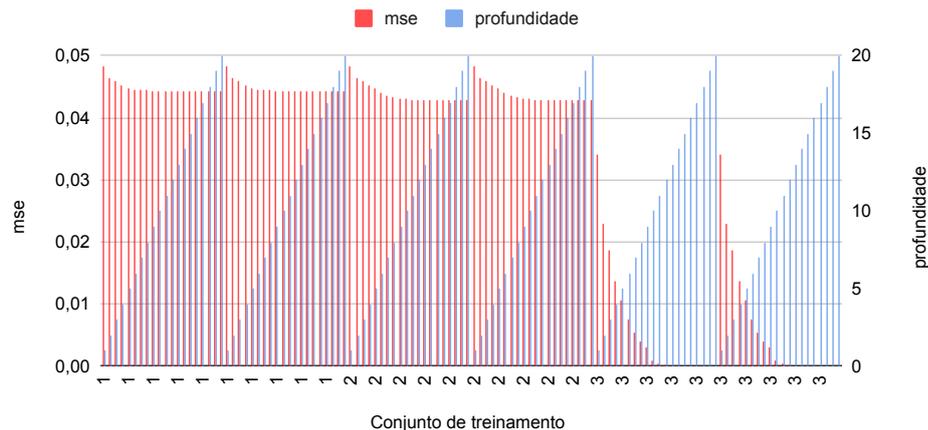
Para cada conjunto de treinamento o código foi executado variando sua profundidade de 2 a 20. A partir da análise do impacto da profundidade foi escolhido não treinar com a profundidade aumentando até 50. Foram realizadas 360 execuções, em metade destas foram utilizados os dados normalizados e na outra metade os dados padronizados. Visando uma melhor visualização do resultado dos treinamentos, estes foram separados nas figuras 11, 12 e 13. Não houve uma separação entre o tipo de técnica para obter a mesma escala, pois a diferença entre os resultados não foi significativa.

Figura 10 – Gráfico do MSE de acordo com a profundidade.



Fonte: Elaborado pelo autor (2022).

Figura 11 – Gráfico do MSE da árvore para os conjuntos 1 a 3.



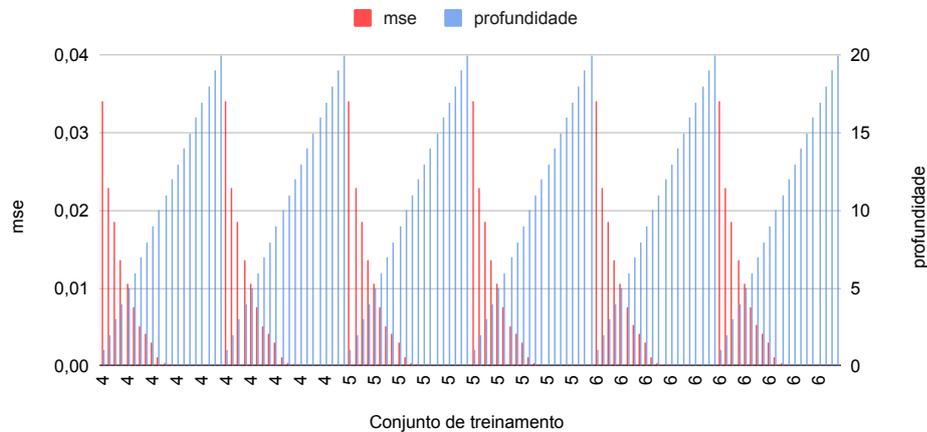
Fonte: Elaborado pelo autor (2022).

A partir deste método foram coletados resultados significativamente melhores para os conjuntos 3 a 6, os quais contém uma maior quantidade de entradas. Outra semelhança entre eles é que todos possuem entre seus dados, a temperatura interna da geladeira, a temperatura externa e o intervalo de tempo entre as coletas de dados. Para as árvores de decisão o melhor resultado se aproximou muito de zero, o MSE foi de $1,69 \mu$.

4.0.2 Florestas Aleatórias

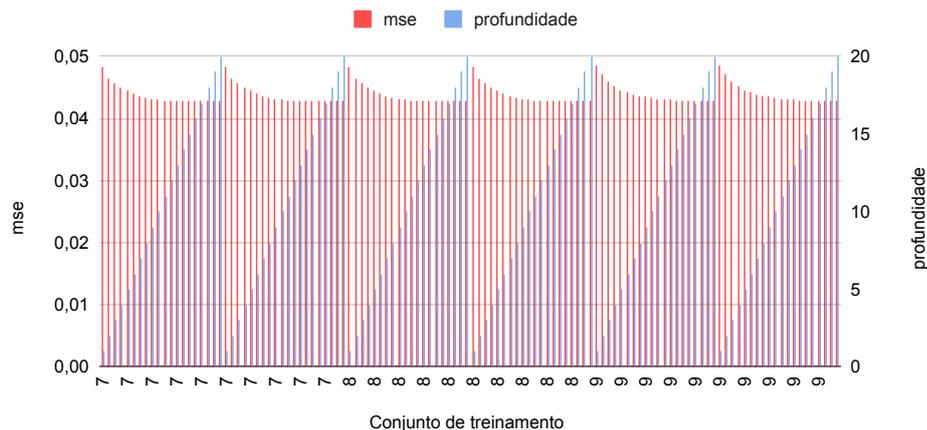
Os resultados obtidos pelo treinamento das Florestas Aleatórias foram analisados a partir da variação da quantidade de estimadores e dos conjuntos de treinamento. Utilizando a análise dos resultados das árvores de decisão, a profundidade

Figura 12 – Gráfico do MSE da árvore para os conjuntos 4 a 6.



Fonte: Elaborado pelo autor (2022).

Figura 13 – Gráfico do MSE da árvore para os conjuntos 7 a 9.

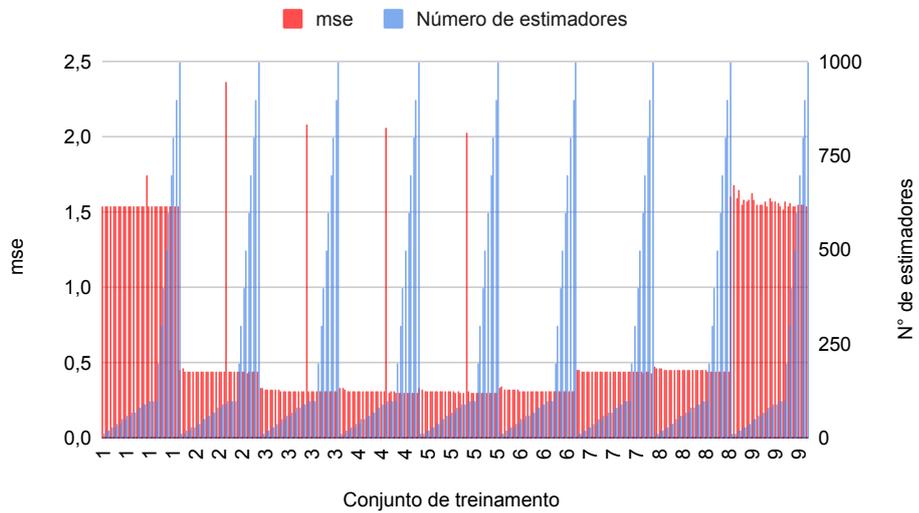


Fonte: Elaborado pelo autor (2022).

limite para as árvores utilizadas no ensemble foi definida em 15. O resultado dos treinamentos utilizando os dados normalizados pode ser visto na figura 14, o menor valor encontrado para o MSE foi 0,304. Utilizando os dados padronizados na figura 15, para este conjunto o menor valor encontrado para o MSE foi 0,306.

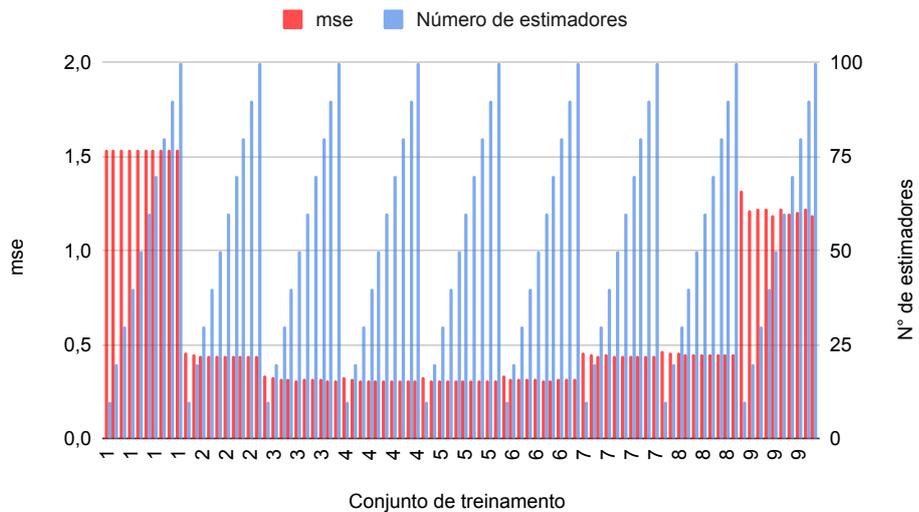
É possível notar, a partir da análise das figuras 14, que os conjuntos de treinamento 3, 4, 5 e 6 obtiveram desempenho melhor do que o restante. Outro ponto importante é que o efeito do aumento de número de estimadores acima de 100 estimadores não é notório ao observar a métrica. Observando a figura 15 e comparando com a anterior é notória a semelhança do seu comportamento. Indo ao encontro de informação dos resultados das árvores de decisão, as árvores não são sensíveis à características com escalas diferentes.

Figura 14 – Gráfico do MSE da floresta com dados normalizados para os conjuntos 1 a 9.



Fonte: Elaborado pelo autor (2022).

Figura 15 – Gráfico do MSE da floresta com dados padronizados para os conjuntos 1 a 9.



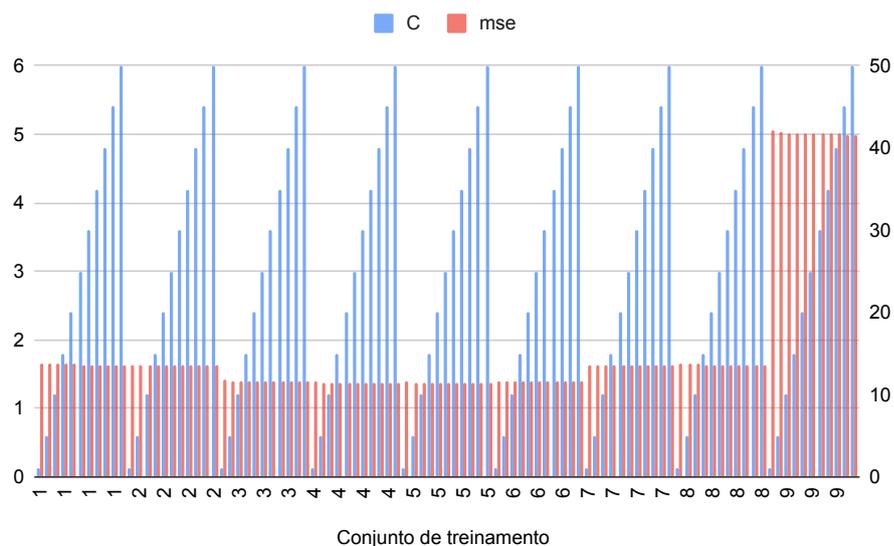
Fonte: Elaborado pelo autor (2022).

4.0.3 Máquinas de Vetores de Suporte

Para as máquinas de vetores de suporte os testes envolvendo a função sigmoide resultaram em erros muito expressivos. O algoritmo de treinamento foi executado 165 vezes e apenas em nove execuções o MSE ficou abaixo de 6, com seu menor valor sendo 1,68. Para esta execução foi utilizado o conjunto de treinamento 3 e o parâmetro de regularização C foi de 1.

Para as funções de ativação gaussiana e polinomial os resultados foram melhores. Eles estão representados nas figuras 16 a 19. A função polinomial obteve o melhor resultado na execução com os dados normalizados, com o conjunto de treinamento 5, com parâmetro de regularização 50 e grau 3. E para a função gaussiana o menor erro médio quadrático foi encontrado, com parâmetro de regularização 20, e conjunto de treinamento 6.

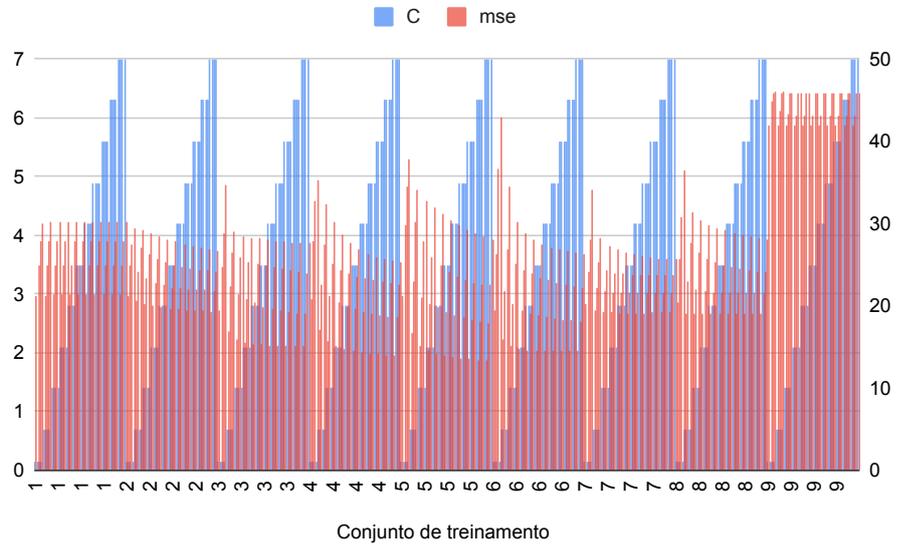
Figura 16 – Gráfico do MSE da SVM com dados normalizados para a função gaussiana.



Fonte: Elaborado pelo autor (2022).

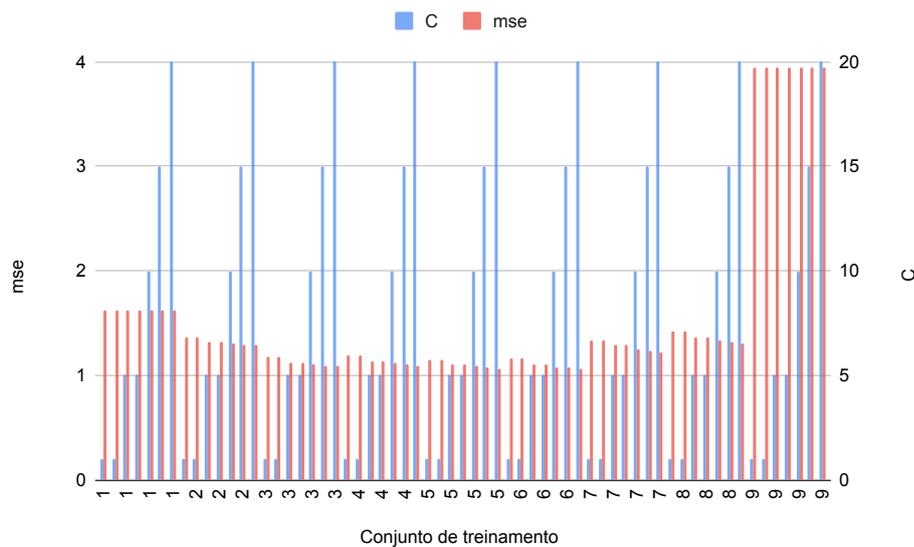
Com base na análise dos gráficos das figuras 16 a 19 é possível afirmar que os melhores desempenhos estão nos conjuntos com mais características. Pode-se concluir também que das três funções de ativação analisadas a gaussiana obteve melhor desempenho médio. Ao compararmos os dados de entrada, no que se refere às técnicas de normalização e padronização, a função gaussiana obteve resultados melhores com os dados padronizados e a função polinomial com os dados normalizados. Entre todos os testes envolvendo as SVMs o melhor resultado foi da função gaussiana com parâmetro de regularização 20, e conjunto de treinamento 6. Seu MSE foi de 1,060

Figura 17 – Gráfico do MSE da SVM com dados normalizados para a função polinomial.



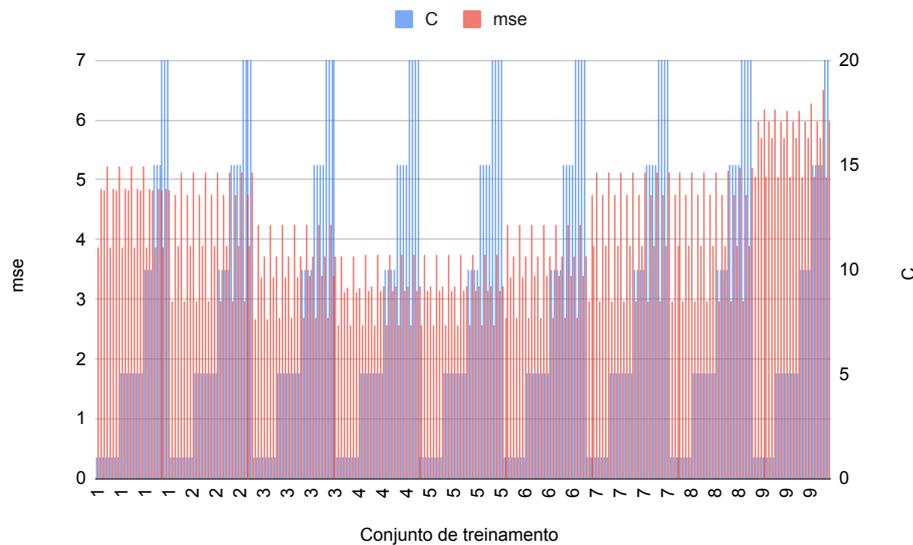
Fonte: Elaborado pelo autor (2022).

Figura 18 – Gráfico do MSE da SVM com dados padronizados para a função gaussiana.



Fonte: Elaborado pelo autor (2022).

Figura 19 – Gráfico do MSE da SVM com dados padronizados para a função polinomial.



Fonte: Elaborado pelo autor (2022).

4.0.4 Redes Neurais Artificiais

Os resultados obtidos através do treinamento dos modelos de redes neurais artificiais foram divididos no impacto do número de épocas e da variação dos conjuntos de treinamento. Foram realizadas ao todo 467 execuções. Também foi feita uma análise individual dos modelos.

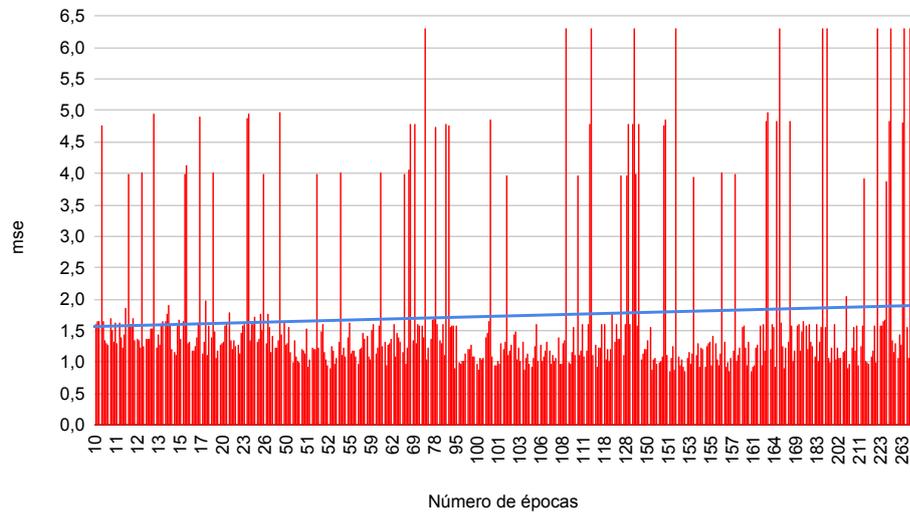
4.0.4.1 Número de Épocas

Os treinamentos realizados variando o número de épocas, sem considerar as variações nos dados de entrada, estão expressados através da figura 20. É difícil visualizar se o MSE aumenta ou diminui conforme aumenta o número de épocas no gráfico. Para solucionar este problema foi traçada uma curva de tendência. Esta curva é representada como uma linha azul na figura 20. Assim podemos notar uma tendência de aumento do erro conforme aumentamos o número de épocas, se considerarmos todos os dados treinados.

4.0.4.2 Conjuntos de Treinamento

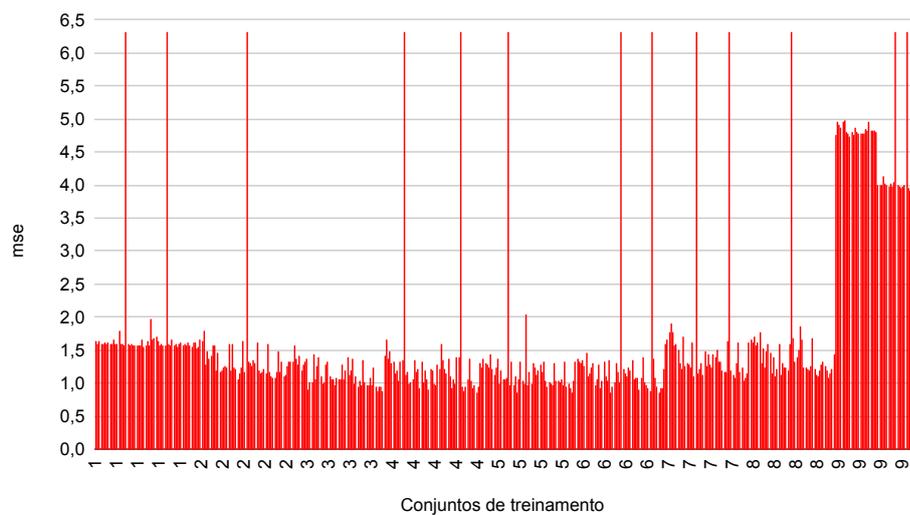
Ao separar os resultados para cada conjunto de treinamento diferente e representar na figura 21, é possível ter uma visualização mais clara do impacto de cada conjunto nos resultados finais. É possível notar o baixo desempenho do conjunto de treinamento 9. O conjunto de treinamento 1 também possui uma desvantagem em relação aos outros, porém não tão acentuada como o analisado anteriormente.

Figura 20 – Gráfico do MSE em relação as épocas da RNA.



Fonte: Elaborado pelo autor (2022).

Figura 21 – Gráfico do MSE em relação aos conjuntos de treinamento da RNA.



Fonte: Elaborado pelo autor (2022).

4.0.4.3 Modelos

Entre os modelos foi analisado tanto o menor valor encontrado para o MSE quando o valor médio. Se compararmos as tabelas 8 e 9 nota-se que a tabela com os dados padronizados obteve resultados melhores. É comum entre as tabelas que o Modelo 3 obteve o melhor desempenho tanto em relação ao MSE mínimo quanto ao valor médio do erro. Os 10 melhores desempenhos do Modelo 3 foram obtidos com os conjuntos de treinamento 4, 5 e 6.

Sabendo que o modelo 3 obteve melhor resultado em comparação aos outros modelos de RNAs foram realizados mais treinamentos. Ao todo foram mais 30 execuções do modelo 3 variando o número de épocas entre os conjuntos de treinamento 4, 5 e 6. Assim foi encontrado o melhor desempenho entre as redes neurais artificiais. Com o conjunto 5 e 507 épocas foi obtido um MSE de 0,800.

Tabela 8 – Tabela de MSE dos modelos de RNAs para os dados padronizados

Modelos	MSE mínimo	MSE médio
1	0,943	1,227
2	0,883	1,196
3	0,857	1,116
4	0,857	1,160
5	0,894	1,204
6	1,217	1,424

Fonte: Elaborado pelo autor (2022).

Tabela 9 – Tabela de MSE dos modelos de RNAs para os dados normalizados

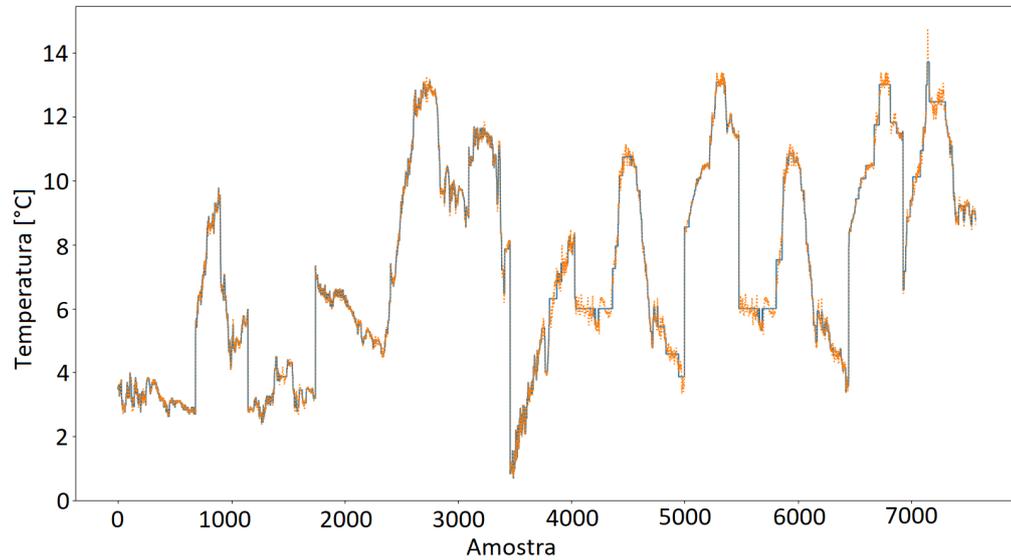
Modelos	MSE mínimo	MSE médio
1	1,003	1,370
2	0,980	1,256
3	0,863	1,229
4	0,908	1,238
5	0,964	1,313
6	1,279	1,578

Fonte: Elaborado pelo autor (2022).

4.0.5 Comparação Entre as Técnicas

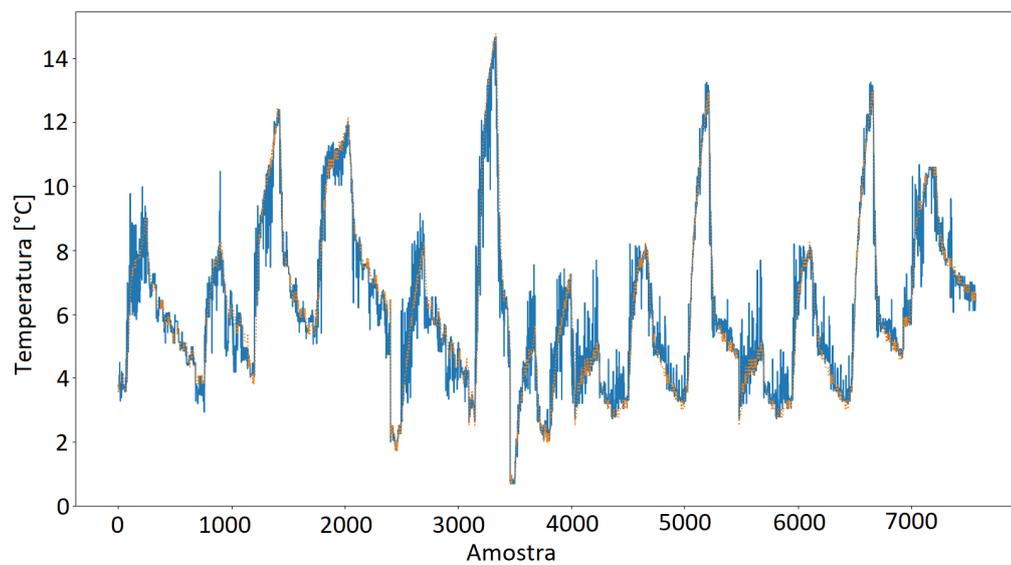
Após a análise de cada técnica separadamente foi encontrado um melhor resultado para cada técnica. Estes resultados estão representados nas figuras 22 a 25. Nestas figuras é apresentado o valor esperado, traço pontilhado laranja, e o valor previsto, traço azul.

Figura 22 – Gráfico de previsões e valores esperados da Árvore de Decisão.



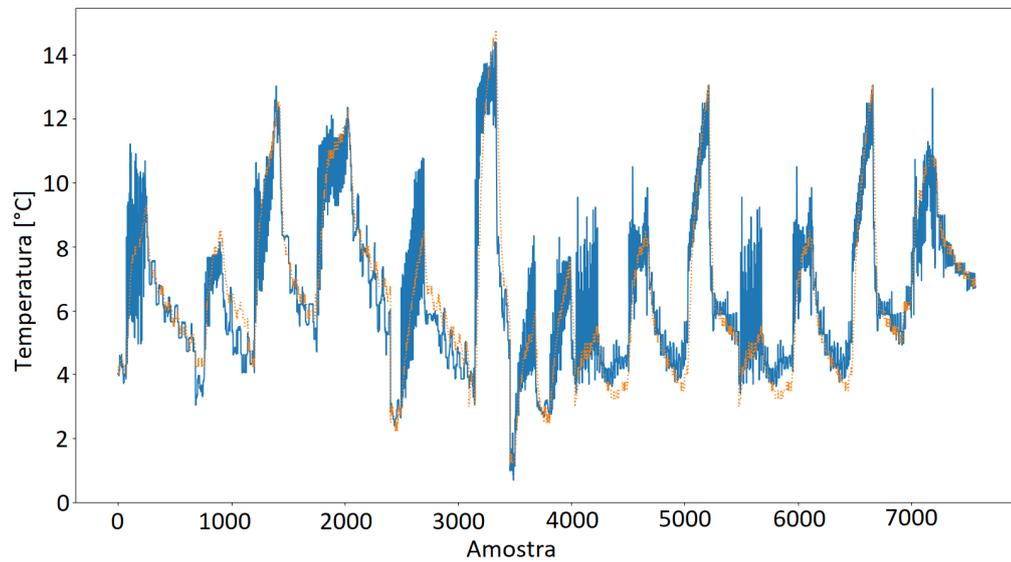
Fonte: Elaborado pelo autor (2022).

Figura 23 – Gráfico de previsões e valores esperados da Floresta Aleatória.



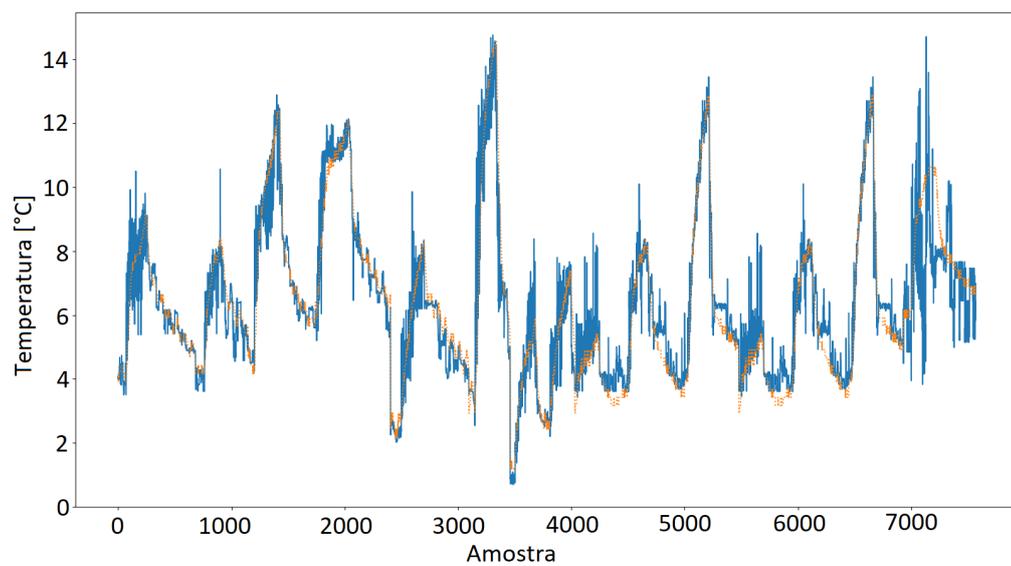
Fonte: Elaborado pelo autor (2022).

Figura 24 – Gráfico de previsões e valores esperados da Máquina de Vetores de Suporte.



Fonte: Elaborado pelo autor (2022).

Figura 25 – Gráfico de previsões e valores esperados das Redes Neurais Artificiais.



Fonte: Elaborado pelo autor (2022).

Ao analisar as saídas previstas e comparando com os valores reais de cada modelo, algumas observações podem ser feitas. Apesar da árvore possuir o menor erro entre todos os modelos, ela ficou sobreajustada, o que provavelmente implica na ineficiência de prever novos valores. Entre os outros métodos, a floresta aleatória se destacou com um MSE de 0,304, comparado ao 1,060 da SVM e 0,800 da RNA. A SVM foi a pior técnica de inteligência artificial.

Em relação aos conjuntos de treinamento, todas as técnicas desempenharam melhor com os conjuntos com um maior número de entradas, especialmente os conjuntos 5 e 6. Entre normalização e padronização as florestas aleatórias e as árvores de decisão não obtiveram desempenhos notoriamente distintos. Para as redes neurais artificiais os dados padronizados terminaram com um resultado melhor. Para as SVMs dependeu da função de ativação. A função polinomial e a função sigmoide se aproximaram mais do resultado esperado com os dados normalizados, enquanto a função gaussiana melhorou com os dados padronizados.

Tabela 10 – Tabela dos resultados finais

Técnica	MSE
Árvore de decisão	1,69 μ
Floresta aleatória	0,304
Máquina de Vetores de Suporte	1,060
Redes neurais artificiais	0,800

Fonte: Elaborado pelo autor (2022).

5 CONCLUSÕES

As vacinas foram descoberta essencial para combater diversas doenças. Foram centro da atenção do mundo durante a recente pandemia de COVID-19. Para mantermos a qualidade e capacidade de imunização das vacinas é de suma importância um armazenamento na faixa de temperatura ideal. Conseqüentemente, é necessário haver uma rede de frio bem estruturada.

A partir deste problema foi desenvolvido o sistema embarcado para monitoramento de temperatura. Para que o sistema fique de acordo com o Manual de Rede de Frio é necessária a retirada do *buffer* físico de temperatura (MINISTÉRIO DA SAÚDE, 2017). Assim, foram desenvolvidas e treinadas quatro técnicas de inteligência artificial.

A partir da análise dos resultados é possível concluir que dentre os modelos de inteligência artificial o que obteve a maior acurácia, considerando a métrica escolhida (erro médio quadrático), foram as florestas aleatórias. Os resultados das SVMs foram os piores entre os modelos testados. Isso corrobora com a afirmação de Géron, que as máquinas de vetores de suporte não são ideais para problemas não lineares em larga escala (GÉRON, 2019).

Todos os modelos propostos encontraram soluções para todos os dados em todos os treinamentos, apesar do seu erro, portanto é possível concluir que são algoritmos completos. Em relação à complexidade de tempo, as árvores de decisão foram o algoritmo que encontrava as respostas mais rapidamente, normalmente alguns segundos, seguido das florestas aleatórias. Os treinamentos das máquinas de vetores de suporte e as redes neurais artificiais, dependendo dos parâmetros escolhidos, chegaram a durar mais do que uma hora. Não foi medida a complexidade de espaço dos algoritmos.

Os objetivos do trabalho foram alcançados. Foram realizadas diversas execuções dos quatro modelos de inteligência artificial propostos, variando seus parâmetros e entradas. Todos os modelos foram capazes de simular o *buffer* de temperatura, variando sua acurácia. Foi feita uma análise individual de cada técnica e em seguida uma comparação entre os modelos. Finalmente foi definido que o modelo das florestas aleatórias obteve o melhor desempenho obedecendo as métricas analisadas, isto é um MSE de 0,304.

Os resultados deste trabalho foram analisados apenas em relação a dados já coletados. Em uma pesquisa futura seria possível aplicar os modelos desenvolvidos neste trabalho no sistema embarcado e realizar testes em tempo real. E a partir destes testes seria possível também analisar se as conclusões feitas vão ao encontro dos

resultados prácticos.

REFERÊNCIAS

- APS, L. R. d. M. M. et al. Eventos adversos de vacinas e as consequências da não vacinação: uma análise crítica. **Revista de saúde pública**, v. 52, n. 1, p. 40–53, 2018.
- BURRELL, J. How the machine ‘thinks’: understanding opacity in machine learning algorithms. **Big data & society**, SAGE Publications, London, England, v. 3, n. 1, p. 1–12, 2016.
- CHEN, K.; SHAW, Y. Applying back propagation network to cold chain temperature monitoring. **Advanced Engineering Informatics**, v. 25, n. 1, p. 11–22, 2011.
- DOSHI-VELEZ, F.; KIM, B. **Towards A Rigorous Science of Interpretable Machine Learning**. arXiv, 2017. Disponível em: <https://arxiv.org/abs/1702.08608>.
- EMOND, J. P. The cold chain. In: **RFID technology and applications**. Cambridge, Reino Unido: Cambridge University Press, 2008. p. 144–156.
- GÉRON, A. **Mãos à obra: aprendizado de máquina com scikit-learn & tensorflow - conceitos, ferramentas e técnicas para a construção de sistemas inteligentes**. Rio de Janeiro: Altas Books, 2019.
- GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: **Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics**. PMLR, 2010. (Proceedings of Machine Learning Research, v. 9), p. 249–256. Disponível em: <https://proceedings.mlr.press/v9/glorot10a.html>.
- HANSON, C. M. et al. Is freezing in the vaccine cold chain an ongoing issue? a literature review. **Vaccine**, v. 35, n. 17, p. 2127–2133, 2017.
- IBM CLOUD EDUCATION. **Random Forest**. 2020. Disponível em: <https://www.ibm.com/cloud/learn/random-forest>. Acesso em: 17 dez. 2022.
- KERAS. **EarlyStopping**. 2022. Disponível em: https://keras.io/api/callbacks/early_stopping/. Acesso em: 21 out. 2022.
- LU, Z. et al. Reservoir observers: model-free inference of unmeasured variables in chaotic systems. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, v. 27, n. 4, p. 041102, abr. 2017.
- LUCCI, S.; KOPEC, D.; MUSA, S. **Artificial Intelligence in the 21st Century**. Mercury Learning & Information, 2022. ISBN 9781683922230. Disponível em: <https://books.google.com.br/books?id=5MPxswEACAAJ>.
- MINISTÉRIO DA SAÚDE. **Manual de rede de frio do programa nacional de imunizações**. 2017. Disponível em: https://bvsmis.saude.gov.br/bvs/publicacoes/manual_rede_frio_programa_imunizacoes_5ed.pdf. Acesso em: 21 fev. 2022.
- MINISTÉRIO DA SAÚDE. **Apresentação**. 2022. Disponível em: <http://pni.datasus.gov.br/apresentacao.asp>. Acesso em: 19 fev. 2022.

MINISTÉRIO DA SAÚDE. **Rede de Frio**. 2022. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/calendario-nacional-de-vacinacao/rede-de-frio>. Acesso em: 17 dez. 2022.

PATINE, F. D. S. et al. Analysis of vaccine loss due to temperature change. **Revista brasileira de enfermagem**, Associação Brasileira de Enfermagem, Brazil, v. 74, n. 1, p. e20190762–8, 2021.

ROSSIGNATTI, C. et al. Sistema embarcado para o monitoramento de vacinas: Aprendizagem baseada em projetos no pet emb. **Congresso Brasileiro de Educação em Engenharia**, v. 49, n. 1, p. 1–12, abr. 2021.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986.

TAULLI, T. Ai foundations. In: **Artificial intelligence basics: a non-technical introduction**. Berkeley, CA: Apress, 2019. p. 1–17. ISBN 978-1-4842-5028-0.

THE NEW YORK TIMES. **New navy device learns by doing**. 1958. Disponível em: <http://www.nytimes.com/1958/07/08/archives/new-navy-device-learns-by-doing-psychologist-shows-embryo-of.html>. Acesso em: 15 fev. 2022.

THE WORLD HEALTH ORGANIZATION. **Smallpox**. 2022. Disponível em: https://www.who.int/health-topics/smallpox#tab=tab_1. Acesso em: 19 fev. 2022.

TSANG, Y. P. et al. An Internet of Things (IoT)-based risk monitoring system for managing cold supply chain risks. **Industrial management & data systems**, 118, n. 7, SI, p. 1432–1462, 2018.