

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
CÂMPUS MEDIANEIRA  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS  
COMPUTACIONAIS PARA O AGRONEGÓCIO

EDUARDA ARAUJO ANTONIOLLI

**PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA EM UM  
FRIGORÍFICO: UM ESTUDO DE CASO UTILIZANDO REGRESSÃO  
LINEAR, REDES NEURAIS E MÁQUINAS DE VETORES DE  
SUPORTE**

DISSERTAÇÃO DE MESTRADO

MEDIANEIRA

2021

EDUARDA ARAUJO ANTONIOLLI

**PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA EM UM  
FRIGORÍFICO: UM ESTUDO DE CASO UTILIZANDO REGRESSÃO  
LINEAR, REDES NEURAIIS E MÁQUINAS DE VETORES DE  
SUPORTE**

**Electrical energy consumption forecast in a poultry processing facility: a  
case study using linear regression, neural networks and support vector  
machine**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio da Universidade Tecnológica Federal do Paraná (UTFPR), como requisito parcial para obtenção do título de "Mestre em Tecnologias Computacionais para o Agronegócio"

Orientador: Prof. Dr. José Airton Azevedo dos Santos

Co-orientador: Prof. Me. Alex Lemes Guedes

**MEDIANEIRA**

**2021**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite *download* e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



EDUARDA ARAUJO ANTONIOLLI

**PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA EM UM FRIGORÍFICO: UM ESTUDO DE CASO  
UTILIZANDO REGRESSÃO LINEAR, REDES NEURAIS E MÁQUINAS DE VETORES DE SUPORTE**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Tecnologias Computacionais Para O Agronegócio da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Tecnologias Computacionais Aplicadas À Produção Agrícola E Agroindústria.

Data de aprovação: 29 de Outubro de 2021

Prof Jose Airton Azevedo Dos Santos, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Leandro Antonio Pasa, Doutorado - Universidade Tecnológica Federal do Paraná

Prof.a Luani Back Silvina, Doutorado - Faculdade Dinâmica das Cataratas (Udc)

Prof Paulo Cesar Tonin, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 04/11/2021.

## RESUMO

Esta pesquisa teve como objetivo realizar a previsão do consumo de energia elétrica de um frigorífico. Para isso, foram desenvolvidos modelos, de séries temporais, usando algoritmos de aprendizagem, como regressão linear, redes neurais e máquinas de vetores de suporte, através do *software* WEKA. A entrada foi o histórico de consumo da unidade, do período de novembro de 2016 a dezembro de 2020. Os hiperparâmetros dos modelos foram escolhidos a partir da avaliação do MAPE e do MAE do conjunto de teste. Para a validação, realizou-se a previsão de um semestre a frente, ou seja, para o período de janeiro de 2021 a junho de 2021. As estimativas foram comparadas com o consumo real observado na unidade. O modelo com melhor desempenho foi o SVM, com um MAPE de 3,38%. Os modelos de regressão linear e redes neurais, apesar de apresentarem um erro acima do MAPE estabelecido para este trabalho, de 5%, também são considerados aptos para estimar o consumo futuro de energia elétrica para a unidade em questão.

**Palavras-chave:** projeção, modelagem, algoritmo de aprendizagem, unidade industrial.

## ABSTRACT

This research aimed to forecast the electricity consumption of a poultry facility unit. For this, time series models were developed using learning algorithms, such as linear regression, neural networks and support vector machines, using *WEKA software*. The input was the unit's consumption history, from November 2016 to December 2020. The hyperparameters of the models were chosen from the evaluation of the MAPE and MAE of the test set. For validation, a forecast of one semester ahead was performed, that is, for the period from January 2021 to June 2021. The data were compared with the actual consumption observed in the unit. The model with the best performance was the SVM, with a MAPE of 3.38%. The linear regression models and neural networks, despite presenting an error above the MAPE established for this work, of 5%, are also considered capable of estimating the future consumption of electricity for the unit in question.

**Keywords:** projection, modeling, learning algorithm, industrial unit

## LISTA DE FIGURAS

FIGURA 1	– Ciclo de aprendizagem de máquina .....	15
FIGURA 2	– Diagrama de blocos da rede neural .....	17
FIGURA 3	– Camadas de uma rede neural .....	18
FIGURA 4	– SVM com $\epsilon=0,1, 0,2$ e $0,5$ .....	21
FIGURA 5	– Procedimento simplificado da SVR .....	22
FIGURA 6	– SVM com margens suaves .....	22
FIGURA 7	– Diagrama de blocos da SVM .....	23
FIGURA 8	– Interface inicial do <i>software</i> WEKA versão 3.8.4 .....	24
FIGURA 9	– Fluxograma da transformação dos dados de base horária para mensal ...	28
FIGURA 10	– Banco de dados no formato ARFF .....	30
FIGURA 11	– Interface <i>Basic configuration</i> do WEKA .....	31
FIGURA 12	– Interface de configuração da regressão linear no WEKA .....	34
FIGURA 13	– Interface de configuração das redes neurais no WEKA .....	35
FIGURA 14	– Fluxograma para configuração das redes neurais .....	35
FIGURA 15	– Interface de configuração das SVM no WEKA .....	37
FIGURA 16	– Consumo mensal de energia elétrica de 2016 a 2021 .....	40
FIGURA 17	– Consumo anual de 2016 a 2021 .....	42
FIGURA 18	– Linha de tendência polinomial para o consumo mensal .....	43
FIGURA 19	– Previsões de regressão linear dos dados de teste para o parâmetro de <i>Ridge</i>	45
FIGURA 20	– Previsões de regressão linear dos dados de teste para o método de seleção	46
FIGURA 21	– Previsões de regressão linear dos dados de teste para o ajuste de variância	47
FIGURA 22	– Previsões de redes neurais dos dados de teste para os neurônios .....	48
FIGURA 23	– Previsões de redes neurais dos dados de teste para a taxa de aprendizagem	49
FIGURA 24	– Previsões de redes neurais dos dados de teste para o <i>momentum</i> .....	50
FIGURA 25	– Previsões de redes neurais dos dados de teste para o tempo de treinamento	51
FIGURA 26	– Previsões de redes neurais dos dados de teste para o ajuste de variância .	52
FIGURA 27	– Previsões de SVM dos dados de teste para o <i>kernel</i> .....	53
FIGURA 28	– Previsões de SVM dos dados de teste para o otimizador .....	54
FIGURA 29	– Previsões de SVM dos dados de teste para o parâmetro C .....	55
FIGURA 30	– Previsões de SVM dos dados de teste para o ajuste da variância .....	56
FIGURA 31	– Comparação entre as previsões dos modelos de regressão linear, redes neurais e máquina de vetores de suporte com o consumo .....	58

## LISTA DE TABELAS

TABELA 1	– Ponto de medição, data, hora e tipo de energia disponibilizadas no SCDE	27
TABELA 2	– Medições ativa e reativa de consumo disponibilizadas no SCDE	28
TABELA 3	– Média do consumo mensal por ano	41
TABELA 4	– Teste de Tukey para o tratamento mês	41
TABELA 5	– Teste de Tukey para o tratamento ano	42
TABELA 6	– Equações de linha de tendência para o consumo mensal	43
TABELA 7	– MAE e MAPE dos modelos de regressão linear para o parâmetro de <i>Ridge</i>	44
TABELA 8	– MAE e MAPE dos modelos de regressão linear para o método de seleção	45
TABELA 9	– MAE e MAPE dos modelos de regressão linear para o ajuste de variância	46
TABELA 10	– Configuração do modelo de regressão linear	47
TABELA 11	– MAE e MAPE dos modelos de redes neurais para os neurônios	48
TABELA 12	– MAE e MAPE dos modelos de redes neurais para a taxa de aprendizagem	49
TABELA 13	– MAE e MAPE dos modelos de redes neurais para o <i>momentum</i>	50
TABELA 14	– MAE e MAPE dos modelos de redes neurais para o tempo de treinamento	51
TABELA 15	– MAE e MAPE dos modelos de redes neurais para o ajuste de variância	52
TABELA 16	– Configuração do modelo de redes neurais	52
TABELA 17	– MAE e MAPE dos modelos de SVM para o <i>kernel</i>	54
TABELA 18	– MAE e MAPE dos modelos de SVM para o otimizador	55
TABELA 19	– MAE e MAPE dos modelos de SVM para o parâmetro C	56
TABELA 20	– MAE e MAPE dos modelos de SVM para o ajuste da variância	56
TABELA 21	– Configuração do modelo de máquina de vetores de suporte	57
TABELA 22	– MAPE dos modelos de regressão linear, redes neurais e máquina de vetores de suporte	59
TABELA 23	– MAE dos modelos de regressão linear, redes neurais e máquina de vetores de suporte	59
TABELA 24	– RMSE dos modelos de regressão linear, redes neurais e máquina de vetores de suporte	59

## LISTA DE SIGLAS

MAE	Mercado Atacadista de Energia Elétrica
CCEE	Câmara de Comercialização de Energia Elétrica
ANEEL	Agência Nacional de Energia Elétrica
TUSD	Tarifa do Uso do Sistema de Distribuição
ONS	Operador Nacional do Sistema Elétrico
PLD	Preço da Liquidação das Diferenças
ACL	Ambiente de Contratação Livre
SVM	Máquina de vetor de suporte
SVR	Regressão de vetor de suporte
WEKA	<i>Waikato Environment for Knowledge Analysis</i>
SCDE	Sistema de Coleta de Dados de Energia
CVS	<i>Comma-Separated Values</i>
MAE	Erro absoluto médio
MAPE	Erro absoluto médio percentual
RMSE	Raiz Quadrada do Erro Médio

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>9</b>
1.1	JUSTIFICATIVA	10
1.2	OBJETIVOS	11
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>12</b>
2.1	MERCADO LIVRE DE ENERGIA ELÉTRICA	12
2.2	MODELAGEM DE SISTEMAS	14
2.3	REDES NEURAIS	16
2.4	MÁQUINA DE VETORES DE SUPORTE	20
2.5	WEKA	23
2.6	TRABALHOS RELACIONADOS	25
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>26</b>
3.1	METODOLOGIA	26
3.2	APRESENTAÇÃO DA UNIDADE CONSUMIDORA	26
3.3	OBTENÇÃO DOS DADOS DE ENTRADA	27
3.4	ANÁLISE ESTATÍSTICA	29
3.5	CONFIGURAÇÕES DO WEKA	29
3.5.1	Ambiente <i>Explorer - Basic configuration</i>	30
3.5.2	Ambiente <i>Explorer - Advanced configuration</i>	31
3.6	CONFIGURAÇÕES DOS MODELOS	32
3.6.1	Regressão	33
3.6.2	Rede neural	34
3.6.3	Máquina de vetores de suporte	37
3.7	AVALIAÇÃO E COMPARAÇÃO ENTRE OS MODELOS	38
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>40</b>
4.1	ANÁLISE ESTATÍSTICA	40
4.2	DESENVOLVIMENTO DOS MODELOS	44
4.2.1	Regressão linear	44
4.2.2	Redes neurais	47
4.2.3	Máquina de vetores de suporte	53
4.3	COMPARAÇÃO ENTRE OS MODELOS	57
<b>5</b>	<b>CONCLUSÕES</b>	<b>63</b>
	REFERÊNCIAS	<b>65</b>

## 1 INTRODUÇÃO

As previsões, ou seja, a capacidade de avaliar o histórico e projetar uma resposta, pode beneficiar o setor elétrico. As projeções são um meio de fornecer informações para tomadas de decisões criteriosa, que proporcionam economia e segurança no fornecimento de energia elétrica. Assim, evitam-se questões como apagões, cortes, deterioração dos sistemas e, até mesmo, paralização de indústrias e serviços. O planejamento de investimentos, redes e capacidade produtiva é diretamente dependente do consumo (CASTRO; FERRARI, 2016).

Ao observar a literatura da área nota-se que grande parte dos trabalhos são focados em previsões com um aspecto nacional, como Ogcü et al. (2012), Kaytez et al. (2015) e Vasconcelos et al. (2017), regional, como em Santos e Chaucoski (2020) e Nizami e Al-Garni (1995), ou para distribuidoras, como nos trabalhos de Campos (2008) e Fogliatto et al. (2005). Ou seja, consideram apenas volumes expressivos de energia. Alguns autores, como Amaral (2020) e Carvalho (2019), realizaram as previsões para consumidores residenciais. Porém, as projeções de consumo para unidades industriais, como a realizada por Castoril et al. (2018), são de grande importância.

Supor um consumo acima do real faz uso de recursos financeiros em investimentos supérfluos. Subestimar esse valor implica em custos maiores de operação e potenciais faltas de energia, que acarretam em penalidades e multas. Desse modo, além da contribuição para uma operação estável e confiável, as previsões representam uma vantagem operacional e financeira (KAYTEZ et al., 2015)

Essa questão é ainda mais importante para consumidores do mercado livre de energia. Com a reestruturação do setor elétrico, nos anos 90, abriu-se a possibilidade à alguns consumidores, que cumpram requisitos específicos, de escolher seu fornecedor de energia elétrica. Neste caso, a compra de energia é feita através de negociações bilaterais, onde são acertadas entre as partes questões como preço, prazo e o volume de energia a ser adquirido (BARROS et al., 2020; PINTO, 2018).

A natureza da energia elétrica torna sua previsão complexa, uma vez que é impactada por

muitos fatores, como condições climáticas e econômicas. A escolha do modelo adequado depende do acesso aos dados históricos, do horizonte da previsão e pode ainda ter outras considerações, dependendo de cada caso (ZENG et al., 2017; KAYTEZ et al., 2015; MORDJAOUI et al., 2017).

É cada vez mais comum a disponibilização, dos dados de consumo, pelas empresas que fornecem energia ou, até mesmo, a instalação de equipamentos específicos para esse fim. Com o avanço da tecnologia e o aumento da capacidade de coleta e armazenamento de dados, surge a oportunidade de aplicar técnicas e ferramentas, como algoritmos de aprendizagem, para extrair informações, auxiliando na compreensão de elementos e aumentando o conhecimento (WITTEN et al., 2016; CASTRO; FERRARI, 2016; DANTAS FILHO, 2019)

Diante do exposto, identifica-se a necessidade de projetar o consumo de energia elétrica em unidades industriais. O objetivo deste trabalho é desenvolver modelos, para previsão do consumo de energia elétrica, para um frigorífico de aves de grande porte, localizado no interior do Paraná. Para isso, propõe-se o uso da ferramenta WEKA, um *software* que, além de interface amigável e um ambiente simples de operar, permite utilizar, como base de aprendizagem, diferentes técnicas, entre elas, regressão linear, redes neurais e máquina de vetor de suporte. Propõe-se, neste trabalho, comparar os resultados das previsões, usando esses algoritmos, visando identificar benefícios e desvantagens de cada um deles.

Este documento será subdividido em cinco seções, para facilitar seu entendimento. Inicia-se com a hipótese, justificativa e objetivos. O capítulo 2 apresenta o referencial teórico, visando esclarecer os principais conceitos e definições, além de outros trabalhos nessa área que foram relevantes. Em seguida, no capítulo 3, propõe-se a metodologia. No capítulo 4 apresentam-se os resultados obtidos dos modelos implementados neste trabalho, e também uma comparação com os resultados encontrados na literatura. Por fim, as conclusões, limitações e ideias para continuidade desta pesquisa estão no capítulo 5.

## 1.1 JUSTIFICATIVA

Como contribuição teórica, este trabalho pode auxiliar na construção do conhecimento dos modelos de previsão para unidades industriais, em especial, para plantas frigoríficas. Para a empresa que forneceu os dados, espera-se que o desenvolvimento deste trabalho seja útil para o planejamento estratégico, auxiliando na tomada de decisões, identificação de pontos de melhoria e otimização de investimentos. Além disso, este projeto pode ser o fundamento de um estudo mais profundo das características dessa unidade em particular.

Os resultados obtidos neste trabalho são valiosos, principalmente considerando-se que

a unidade estudada faz parte do mercado livre de energia. Para a compra de energia, as projeções são uma fonte rica para ponderar as negociações de volumes, prazos e condições.

Com a expectativa de abertura do mercado livre a outros clientes, vem a tona a necessidade de previsões para esses consumidores. Assim, como contribuição social, espera-se que este trabalho possa ser usado como base para outras unidades, colaborando para uma definição sensata do volume de energia de compra.

## 1.2 OBJETIVOS

O objetivo deste trabalho é comparar modelos, de séries temporais, para previsão do consumo de energia elétrica de um frigorífico, situado no interior do estado do Paraná. Para isso, foram traçados os seguintes objetivos específicos:

- Obter e tratar os dados históricos de consumo de energia elétrica da unidade consumidora;
- Conhecer e trabalhar com o *software* WEKA, utilizando a ferramenta de previsão e os algoritmos de aprendizagem;
- Desenvolver modelos de previsão com base em regressão linear, redes neurais e máquina de vetor de suporte;
- Validar os modelos gerados através da avaliação do erro médio absoluto e erro absoluto médio percentual;
- Comparar os modelos desenvolvidos e identificar qual deles tem melhor desempenho.

## 2 FUNDAMENTAÇÃO TEÓRICA

Para um melhor entendimento das características da energia elétrica e das ferramentas usadas nesse trabalho, apresentam-se neste capítulo os principais itens e conceitos relacionados. Uma breve introdução ao mercado de energia elétrica e suas características são apresentadas na seção 2.1. Em seguida, no item 2.2, tem-se os algoritmos de aprendizagem, utilizados neste trabalho, para realizar as previsões.

### 2.1 MERCADO LIVRE DE ENERGIA ELÉTRICA

Os consumidores de energia elétrica são supridos, em sua maior parte, pelas distribuidoras e, em alguns casos, pelas transmissoras. Porém, atualmente, o mercado permite a desvinculação parcial das questões técnicas das comerciais. Isso significa que, apesar da conexão física da unidade consumidora com a rede elétrica ser sempre através da distribuidora, a contratação da energia pode ser realizada por outra empresa, geralmente uma comercializadora (BARROS et al., 2020).

O setor elétrico passou por diversas transformações até chegar ao modelo atual. Um marco importante foi a reestruturação do setor elétrico, na década de 90. As empresas, que eram verticalizadas (ou seja, que centralizam a cadeia de produção), foram desmembradas em distribuição, transmissão e geração. Foi também criada uma nova categoria, as comercializadoras. As distribuidoras e transmissoras atuam, de forma controlada, em um mercado totalmente regulado. As empresas de geração e comercialização passam a atuar em um mercado livre, onde imperam a competitividade e a concorrência (BARROS et al., 2020; PINTO, 2018).

Com a finalidade de controlar e viabilizar as transações de energia elétrica foi criado o Mercado Atacadista de Energia Elétrica (MAE). Em 2004, tornou-se a Câmara de Comercialização de Energia Elétrica (CCEE). Entre as suas atribuições, Barros et al. (2020), destaca:

- manter e registrar os contratos de fornecimento de energia elétrica entre as empresas do

setor elétrico;

- prover a medição e registro da energia elétrica gerada, transmitida e consumida;
- apurar eventuais descumprimentos dos limites de contratação de energia elétrica e, quando necessário, aplicar as respectivas penalidades.

O marco legal do mercado livre foi em 1995, com a promulgação da Lei nº 9.704. Ele estabelece que toda unidade consumidora, com demanda contratada superior a 500 kW, pode atuar no mercado livre de energia. Os potenciais cliente livres podem obter energia de qualquer fonte geradora ou ainda de fontes incentivadas, como eólica, biomassa, solar, entre outras. Com o objetivo de incentivar o consumo dessas fontes, a Agência Nacional de Energia Elétrica (ANEEL) determina que seja concedido um desconto de, no mínimo, 50 % na Tarifa do Uso do Sistema de Distribuição/Transmissão (TUSD) para os contratos com essa energia. Caso a demanda seja superior a 2.000 kW, o consumidor é classificado como livre e deve comprar de fontes convencionais, como grandes hidrelétricas e térmicas, sem direito a desconto na parcela da TUSD (BARROS et al., 2020).

Ao optar pelo mercado livre, o consumidor deve adquirir a energia que pretende consumir de uma comercializadora ou diretamente de um gerador. Essa contratação é feita em um mercado de livre concorrência, com negociações bilaterais. Apesar da compra de energia de um fornecedor de escolha própria, como a distribuidora (ou a transmissora, eventualmente) possui a responsabilidade de manter a rede elétrica em condições de suprir o fornecimento de energia elétrica para a unidade, ela continua a ser remunerada pela TUSD (BARROS et al., 2020).

O sistema de medição é definido pela concessionária local, porém, deve seguir um padrão do Operador Nacional do Sistema Elétrico (ONS) e da CCEE. O principal requisito é um acesso de telecomunicação ao medidor exclusivo para a CCEE acessar os seus dados. A cada mês, a energia efetivamente consumida é medida e comparada com o valor contratado. Caso o consumo seja maior que o montante contratado, a unidade consumidora deve comprar a diferença no chamado mercado de curto prazo. Caso contrário, deve vender o excedente. Seja na compra ou na venda, o preço dessa energia é definido pela CCEE, sendo chamado de Preço de Liquidação das Diferenças PLD (BARROS et al., 2020).

O preço da energia elétrica é definido em função da relação entre oferta e procura. Seu valor pode sofrer grandes variações. Uma das opções para o consumidor é contratar à longo prazo, firmando o valor futuro da energia elétrica a ser pago. Durante a vigência do contrato, o consumidor está protegido de oscilações de preço. As diferenças entre a energia medida e a

contratada, contudo, serão pagas ou recebidas de acordo com o PLD, que é passível de variações (BARROS et al., 2020; PINTO, 2018).

A partir do momento em que um consumidor decide migrar para o mercado livre, ele deve ter conhecimento de que há uma série de obrigações a serem cumpridas. Migrar para o Ambiente de Contratação Livre (ACL) deve ser uma ação muito bem planejada, visto que carrega um alto grau de risco. É comum contratar uma consultoria especializada para dar suporte nessa análise, uma vez que existem muitas regras nesse mercado. Como as decisões relacionadas a esse setor envolvem grandes pagamentos, o responsável deve ter pleno conhecimento de todas as vantagens e desvantagens da associação ao ACL (BARROS et al., 2020).

A compreensão e análise do faturamento de energia permite estabelecer relações importantes entre hábitos e consumo. Os indicadores de uso de energia refletem o perfil de consumo da unidade, permitindo um amplo estudo das características de consumo da instalação (BUENO, 2008). Na indústria, é importante a determinação da demanda futura o mais a frente possível. Estimativas acuradas significam economia pela otimização da operação, configuração e, até mesmo, manutenções (WITTEN et al., 2016).

## 2.2 MODELAGEM DE SISTEMAS

A regressão modela uma ou mais variáveis de saída e as variáveis independentes, ou de controle, utilizando métodos estatísticos. Ou seja, estima uma função a partir de pares entrada-saída. No caso da regressão linear, a adequação da função é feita através do erro quadrático. Quando há uma única variável independente, é dita regressão linear simples. Caso haja mais de uma saída, é denominada multivariada (CASTRO; FERRARI, 2016).

Quando uma equação de regressão aproxima suficientemente bem os dados de treinamento, ela pode ser usada para estimar o valor da saída a partir do valor da variável dependente. A regressão linear é um método simples e ótimo para previsões numéricas. Porém, se os dados apresentam dependência não-linear, é difícil identificar a otimização. Apesar disso, modelos lineares podem ser construídos em blocos, para métodos de aprendizagem mais complexos (WITTEN et al., 2016).

Modelos lineares expressam uma saída como uma combinação de atributos, com pesos pré-determinados. Na Equação 1 o  $x$  representa a classe;  $a_1, a_2, \dots, a_k$  são os valores dos atributos; e  $w_0, w_1, w_2, \dots, w_k$  são os pesos. Os pesos são calculados a partir dos dados de

treinamento (WITTEN et al., 2016).

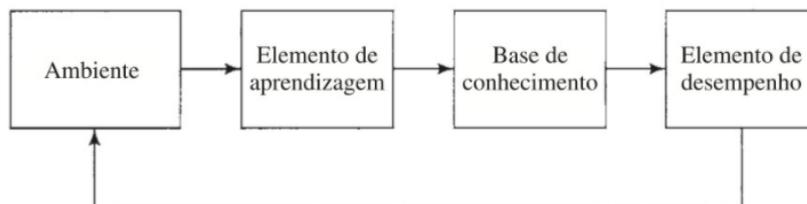
$$x = w_0 + w_1 a_1 + w_2 a_2 + \dots + w_k a_k \quad (1)$$

O ajuste dos pesos pode ser feito por meio de técnicas de aprendizagem de máquinas. Através dela, é possível para acessar informações e extrair conhecimento de um conjunto de dados, auxiliando na tomada de decisões (WITTEN et al., 2016).

A aprendizagem de máquina visa desenvolver programas computacionais capazes de melhorar seu desempenho automaticamente por meio da experiência. É importante ressaltar que não envolve consciência e nem inteligência, mas a extração de informações a partir dos dados de maneira automática. Os sistemas são capazes de se adaptar ou mudar seu comportamento, ou seja, otimizar seu critério de desempenho, com base em experiências passadas, também chamadas de exemplos ou dados de entrada (CASTRO; FERRARI, 2016).

A Figura 1 mostra esse ciclo. O ambiente fornece alguma informação para o elemento de aprendizagem. A base de conhecimento é aperfeiçoada por essa informação. O elemento de desempenho usa a base de conhecimento para executar a tarefa. Normalmente, a informação fornecida pelo ambiente é imperfeita, de modo que o elemento de desempenho não sabe previamente como preencher dados ausentes ou ignorar elementos que não são importantes. Assim, inicialmente, a máquina opera por suposição. Depois, recebe realimentação do elemento de desempenho, que permite a avaliação de suas hipóteses e a revisão, se necessário (HAYKIN, 2001).

**Figura 1: Ciclo de aprendizagem de máquina**



**Fonte: Haykin (2001)**

Nas seções 2.3 e 2.4 serão descritos, com maiores detalhes, as redes neurais e as máquinas de suporte de vetores, dois algoritmos de aprendizagem de máquina.

### 2.3 REDES NEURAIIS

As redes neurais são uma alternativa que permite que os modelos sejam construídos sem o conhecimento explícito da relação entre as variáveis. São capazes de interpolar pontos de dados discretos e aprender tendências dos conjuntos de informações (NIZAMI; AL-GARNI, 1995; HAYKIN, 2001).

O poder das redes neurais vem da sua capacidade de aprendizado. A generalização permite que a rede neural produza saídas adequadas para entradas que não estavam presente durante o treinamento. O procedimento usado é chamado de algoritmo de aprendizagem (HAYKIN, 2001).

Um neurônio é a unidade de processamento de informação fundamental para a operação da rede neural. O conhecimento adquirido é armazenado na conexão entre os neurônios, conhecidos como pesos sinápticos. Os pesos podem assumir valores positivos ou negativos, se o comportamento for excitatório ou inibitório, respectivamente (HAYKIN, 2001; FACELI et al., 2016).

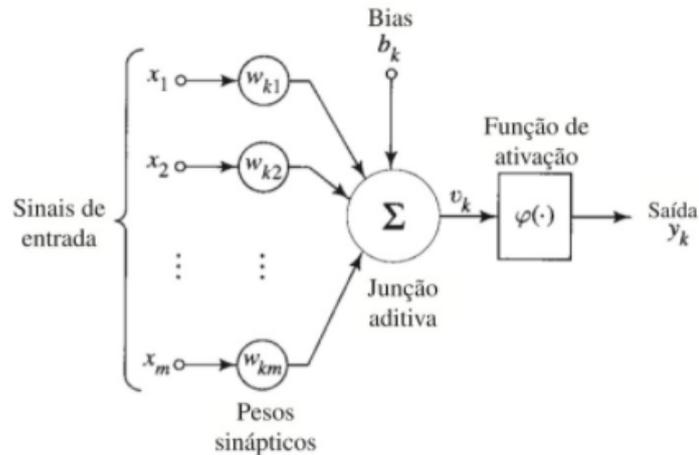
O diagrama de blocos da Figura 2 mostra o modelo de um neurônio. Nele, há um conjunto de sinapses, caracterizados por um peso. O sinal de entrada  $x_j$  é multiplicado pelo peso sináptico  $w_{kj}$ . O somador junta os sinais de entrada, devidamente ponderados. Por fim, a função de ativação restringe a amplitude da saída do neurônio, limitando o intervalo permissível do sinal de saída a um valor finito. Tipicamente, este intervalo é normalizado entre  $[0, 1]$  ou  $[-1, 1]$ . Neste trabalho, utilizou-se a função de ativação sigmóide (Equação 2), que preenche os requisitos de ser não-linear e diferenciável, conforme recomendação de Haykin (2001). O bias, representado por  $b_k$ , tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação (HAYKIN, 2001; SILVA et al., 2010).

$$\varphi(v) = \frac{1}{1 + \exp^{-\alpha v}} \quad (2)$$

Em termos matemáticos, o neurônio  $k$  pode ser descrito pelas Equações 3 e 4. Os sinais de entrada são  $x_1, x_2, x_3, \dots, x_m$ . A variável  $w$  representa os pesos sinápticos do neurônio. O  $\varphi$  é a função de ativação. A variável  $u_k$  representa o sinal de saída do neurônio. O bias ( $b_k$ ) tem o efeito de aplicar uma transformação à saída  $u_k$ .

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3)$$

**Figura 2: Diagrama de blocos da rede neural**



**Fonte: Haykin (2001)**

$$y_k = \varphi(u_k + b_k) \quad (4)$$

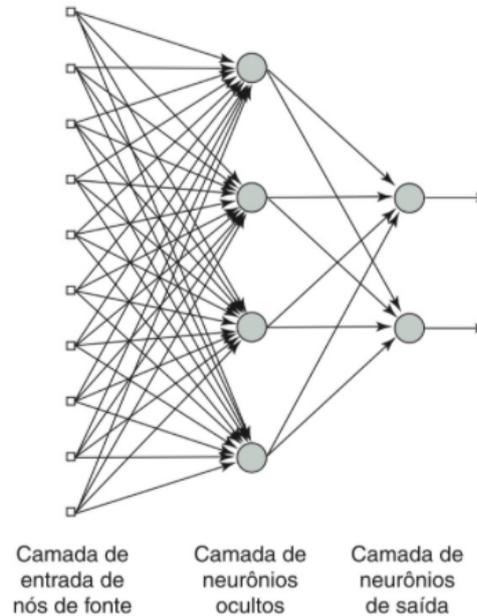
A aprendizagem supervisionada envolve a modificação dos pesos sinápticos de uma rede neural pela aplicação de um conjunto de amostras de treinamento ou exemplos de cada tarefa. Cada exemplo consiste em um sinal de entrada único e uma resposta desejada correspondente (HAYKIN, 2001).

Apresenta-se para a rede um conjunto de dados e os pesos sinápticos da rede são modificados para minimizar a diferença entre a resposta desejada e a resposta real da rede, produzida pelo sinal de entrada, de acordo com um critério estatístico apropriado. O treinamento é repetido para diversos exemplos, até que a rede alcance um estado estável, onde não há mais modificações significativas nos pesos sinápticos. Assim, a rede aprende dos exemplos ao construir um mapeamento de entrada e saída para o problema considerado (HAYKIN, 2001). Uma das propriedades mais interessantes, das redes neurais, é a capacidade de trabalhar e prever mesmo com uma base de dados incompleta, com ruído ou imprecisões (OGCU et al., 2012).

A rede neural artificial pode ser dividida em três partes, denominadas camadas, como mostra a figura 3. A camada de entrada é responsável pelo recebimento de informações advindas do meio externo. As camadas escondidas, intermediárias ou ocultas, são compostas por neurônios que extraem as características associadas ao processo ou sistema. A última camada, de saída, produz e apresenta os resultados finais da rede, advindos dos processamentos feitos pelos neurônios das camadas anteriores (HAYKIN, 2001; SILVA et al., 2010).

O sucesso dos modelos de redes neurais dependem das suas características, como

**Figura 3: Camadas de uma rede neural**



**Fonte: Haykin (2001)**

número de neurônios, camadas, funções utilizadas, algoritmo de aprendizado, pesos iniciais e número de iterações (OGCU et al., 2012). A arquitetura de rede neural define a disposição dos neurônios, através do direcionamento das conexões sinápticas. A topologia, considerando determinada arquitetura, pode ser definida pelas diferentes formas de composições estruturais que esta poderá assumir (SILVA et al., 2010). Haykin (2001) identifica três classes de arquitetura fundamentalmente diferentes.

- Redes alimentadas adiante (do inglês, *feedforward*) com camada única;
- Arquitetura alimentadas adiante com camadas múltiplas;
- Redes recorrentes ou realimentadas.

As redes alimentadas adiante, de camadas múltiplas, são constituídas de uma ou mais camadas escondidas de neurônios. O número de camadas escondidas e neurônios constituintes dependem, sobretudo, do tipo de complexidade do problema a ser mapeado e da quantidade e qualidade dos dados disponíveis (HAYKIN, 2001; SILVA et al., 2010).

Ao adicionar camadas, a rede se torna capaz de extrair estatísticas de ordem elevada. Nessa categoria inclui-se as redes *Perceptron* multicamadas (*multilayer perceptron* - MLP), muito usadas, principalmente, em problemas relacionados a classificação de funções, aproximação

funcional e problemas envolvendo sistemas variantes no tempo (HAYKIN, 2001; SILVA et al., 2010).

As redes *perceptron* multicamadas são treinadas através do algoritmo de retropropagação de erro. A aprendizagem consiste em dois passos. Na propagação, um vetor de entrada é aplicado e seu efeito se propaga através da rede, camada por camada. Nessa etapa, todos os pesos sinápticos da rede são fixos. Um conjunto de saídas é produzido como resposta da rede. A resposta real é subtraída de uma resposta alvo, produzindo um sinal de erro. A partir dessa informação, ocorre a retropropagação, que é um passo para trás, ou seja, os pesos sinápticos são ajustados de acordo com a regra de correção do erro (HAYKIN, 2001).

Uma apresentação completa do conjunto de treinamento é denominado uma época. Devido aos neurônios ocultos, o processo de aprendizagem decide quais as características do padrão de entrada devem ser representadas, de modo que a visualização desse processo é mais complexa. Assim, os neurônios ocultos exercem um papel crucial na operação das redes *perceptron* porque agem como detectores de características (HAYKIN, 2001).

Como propriedades das redes neurais Haykin (2001) e Silva et al. (2010) destacam:

- Não-linearidade: um neurônio artificial pode ser linear ou não-linear. É uma propriedade importante, particularmente se o mecanismo físico responsável pela geração do sinal de entrada for não-linear;
- Adaptabilidade: as redes neurais tem uma capacidade inata de adaptar seus pesos sinápticos a modificações do ambiente. Além da possibilidade de ser retreinada para pequenas modificações, nas condições operativas do ambiente, consegue operar em um ambiente não estacionário, ou seja, a rede neural pode ser projetada para modificar seus pesos sinápticos em tempo real;
- Tolerância a falhas: seu desempenho se degrada suavemente sob condições de operações adversas. Devido a natureza distribuída da informação armazenada na rede, o dano deve ser extenso para que a resposta global de rede seja prejudicada;
- Habilidade de generalização: após o processo de treinamento da rede, essa é capaz de generalizar o conhecimento adquirido, possibilitando estimar soluções até então desconhecidas;
- Organização de dados: a rede é baseada em características intrínsecas do conjunto de informações, de modo que é capaz de realizar sua organização interna, visando possibilitar o agrupamento de padrões que apresentam particularidades em comum.

## 2.4 MÁQUINA DE VETORES DE SUPORTE

As máquinas de vetores do suporte (*support vector machines* - SVM) são embasadas na teoria de aprendizado estatístico e podem ser usadas para implementar processos de aprendizagem. As propriedades dessa máquina de aprendizado permitem que tenham capacidade de generalizar para dados não vistos (HAYKIN, 2001; SMOLA; SCHÖLKOPF, 2004).

Os vetores de suporte consistem em um subconjunto dos dados de treinamento que são extraídos pelo algoritmo. O hiperplano ótimo, formado na superfície, é dado pela equação 5, onde  $x$  é o vetor de entrada,  $w$  é um vetor de peso ajustável e  $b$  é um bias (HAYKIN, 2001).

$$w^T x + b = 0 \quad (5)$$

Diferentes máquinas de aprendizagem podem ser construídas de acordo com a geração do núcleo do produto interno. Enquanto o algoritmo de retropropagação é planejado especificamente para treinar redes MLP, o algoritmo de aprendizagem por vetor de suporte tem uma aplicabilidade mais ampla, incluindo máquinas de aprendizagem polinomial, redes de função de base radial e também *perceptrons* de duas camadas (HAYKIN, 2001; FACELI et al., 2016).

O uso de funções *kernel* para SVMs não lineares torna o algoritmo eficiente, porque permite a construção de hiperplanos simples em um espaço de alta dimensão, tornando-o um problema solucionável computacionalmente (USHA; BALAMURUGAN, 2016).

Da mesma forma que as redes neurais, o SVM pode ser aplicado para solução de problemas de classificação, agrupamento de dados e regressão. Para este último, quando é denominado SVR, do inglês *support vector regression*, ou regressão de vetor de suporte, deve haver algumas reformulações no problema de otimização para lidar com as características e objetivos desses problemas (FACELI et al., 2016).

A regressão de vetor de suporte tem como objetivo obter uma função que produza saídas contínuas para os dados de treinamento (FACELI et al., 2016). Para problemas de regressão não-lineares, um estimador de quadrados mínimos tem um desempenho fraco, principalmente com a presença de ruídos. Para superar essas limitações, é necessário um estimador robusto, que seja insensível a pequenas variações no modelo. Para um procedimento ótimo de estimação robusta usa-se o erro absoluto como a quantidade a ser minimizada. Assim, a função de perda é dada pela equação 6, onde  $d$  é a resposta desejada e  $y$  é a saída do estimador.

$$L(d, y) = |d - y| \quad (6)$$

A resposta desejada  $d$  é mostrada na equação 7, onde passa a ser descrito em função de uma função  $f(x)$ , descrita por um vetor, e um ruído,  $v$ , conforme Haykin (2001). Considera-se  $x$  o conjunto de dados de treinamento. Como  $d$  corresponde a saída do modelo, o problema consiste em fornecer uma estimativa da dependência de  $d$  em relação a  $x$ .

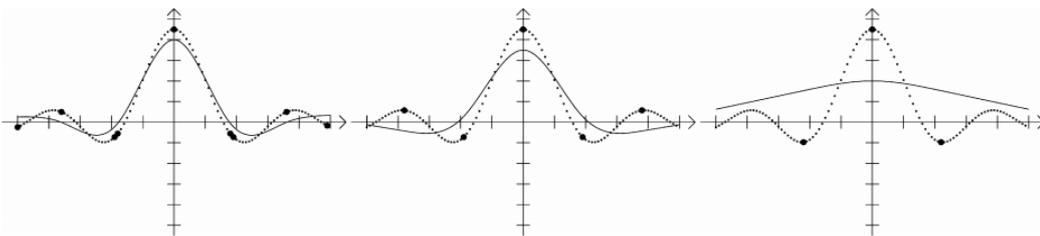
$$d = f(x) + v \quad (7)$$

Para construir uma máquina de vetor de suporte pode-se usar uma extensão da função de perda, dada pela equação 8, denominada função de perda insensível a  $\varepsilon$ , onde  $\varepsilon$  é um parâmetro predeterminado. A função é zero se o valor absoluto do desvio da saída do estimador  $y$  em relação à resposta desejada  $d$  for menor ou igual a zero. Caso contrário, ela é igual ao valor absoluto do desvio menos  $\varepsilon$ .

$$L_{\varepsilon}(d, y) = \begin{cases} |d - y| - \varepsilon, & \text{para } |d - y| \geq \varepsilon \\ 0, & \text{caso contrário} \end{cases} \quad (8)$$

Utilizando a figura 4, Smola e Schölkopf (2004) compara três SVMs com diferentes parâmetros  $\varepsilon$ , demonstrando que o algoritmo seleciona a função mais plana que aproxima os dados originais com dada precisão. Quanto menor a precisão necessária, menos vetores de suporte podem ser utilizados.

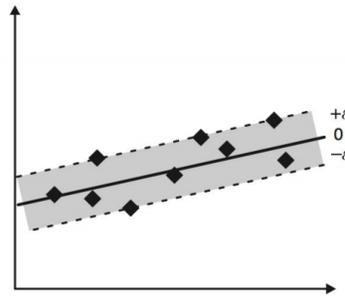
**Figura 4: Da esquerda para direita: SVM com  $\varepsilon=0,1, 0,2$  e  $0,5$**



**Fonte: Smola e Schölkopf (2004)**

Para Smola e Schölkopf (2004), o conjunto de equações 8 representa que serão aceitas diferenças entre a função encontrada e os dados de treinamento, desde que esse desvio se mantenha dentro do parâmetro  $\varepsilon$ . Essa situação é mostrada, de forma ilustrada, na figura 5.

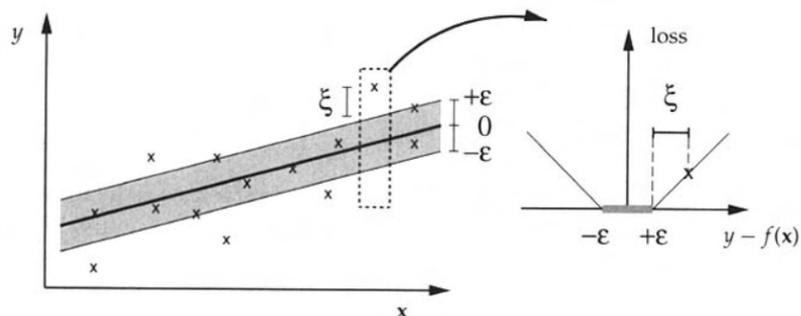
**Figura 5: Procedimento simplificado da SVR**



Fonte: Faceli et al. (2016)

Porém, essa análise e funções funcionam apenas se considerarmos que a função é viável. Para ajustar essa restrição, pode-se adicionar as variáveis soltas  $\xi$  e  $\xi'$ . Neste caso, serão permitidos alguns erros, como mostrado na figura 6, tornando a margem da função mais suave (SMOLA; SCHÖLKOPF, 2004; SCHÖLKOPF; SMOLA, 2002; HAYKIN, 2001).

**Figura 6: SVM com margens suaves**



Fonte: Schölkopf e Smola (2002)

Monta-se um problema dual, com o uso de uma lagrangiana, representado por  $\alpha_i$  e  $\bar{\alpha}_i$ , que é descrito em termos de produtos internos entre objetos e recorre-se ao uso de *kernels*, dado por  $K$ . Na prática, são usadas funções de base radial ou sigmoidais, para realizar as regressões não lineares. Esse *kernel* implica o mapeamento dos objetos para um espaço de características, onde é encontrada a função linear mais regular e com baixo erro de treinamento. Assim, de acordo com Faceli et al. (2016) o problema de otimização final solucionado é dado por:

$$-0,5 \sum_{i=1}^n (\alpha_i - \bar{\alpha}_i)(\alpha_j - \bar{\alpha}_j)K(x_i - x_j) - \varepsilon \sum_{i=1}^n (\alpha_i + \bar{\alpha}_i) + \sum_{i=1}^n y_i(\alpha_i - \bar{\alpha}_i)$$

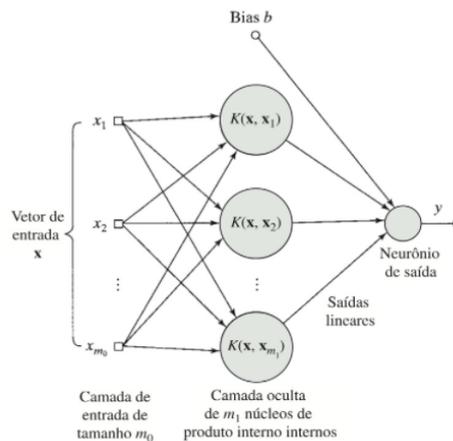
Com as restrições:

$$(9) \quad \begin{cases} \sum_{i=1}^n (\alpha_i - \bar{\alpha}_i) = 0 \\ \alpha_i, \bar{\alpha}_i \in [0, C] \end{cases}$$

Apresenta-se uma equação de otimização, em 9, porém, ainda sem solução. Diversas técnicas e ferramentas podem ser aplicadas para resolver esse problema com suas restrições, entre elas, pacotes matemáticos de *softwares* (FACELI et al., 2016).

Nota-se uma grande semelhança entre as figuras 7, que apresenta a arquitetura para uma máquina de vetores de suporte, e a 2, que mostra a arquitetura de uma rede neural. Entretanto, ao contrário das redes neurais, nas máquinas de vetores de suporte, o número de neurônios e os pesos são determinados automaticamente, em função do número de vetores de suporte e seus valores, respectivamente, sem a necessidade de intervenção do projetista (HAYKIN, 2001).

**Figura 7: Diagrama de blocos da SVM**



**Fonte: Haykin (2001)**

## 2.5 WEKA

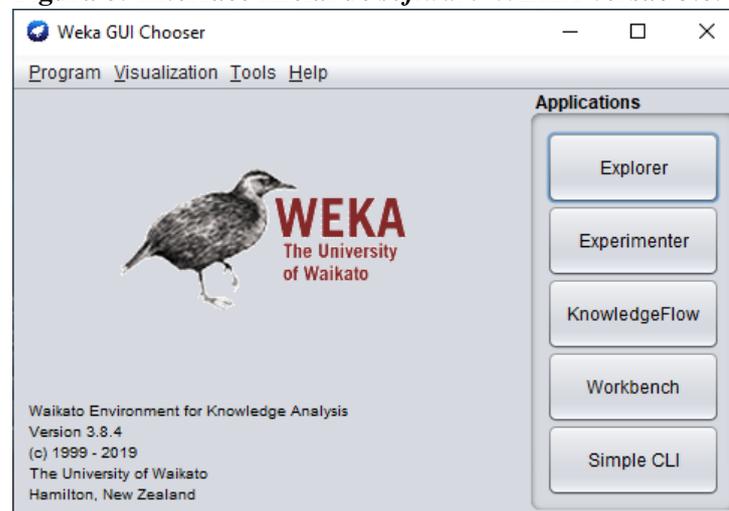
O *software* WEKA (*Waikato Environment for Knowledge Analysis* - em tradução livre, o Ambiente para Análise de Conhecimento Waikato), é um sistema desenvolvido pela Universidade de Waikato, na Nova Zelândia. Escrito em Java, consiste em uma coleção de algoritmos de aprendizagem de máquina e ferramentas de processamento de dados. Oferece suporte para todo o processo de mineração de dados, que inclui a preparação das entradas, a avaliação estatística

dos esquemas de aprendizado e a visualização dos dados de entrada e de saída. Conforme Witten et al. (2016), o WEKA pode ser utilizado de três formas principais:

- Aplicação de métodos de aprendizagem para análise das saídas e coleta de informações;
- Utilização dos modelos de aprendizagem para predições de novos casos;
- Comparação da performance de diferentes métodos de aprendizagem.

Ao iniciar o Weka, na versão 3.8.4, mostrada na figura 8, pode-se optar entre quatro interfaces: *Explorer*, *Knowledge Flow*, *Experimenter* e *Workbench*.

**Figura 8: Interface inicial do software WEKA versão 3.8.4.**



**Fonte: Autoria própria (2021)**

A interface *Experimenter* foi desenvolvida para auxiliar na avaliação dos métodos e parâmetros utilizados, através da comparação entre diferentes técnicas. Permite automatizar o processo de testes de classificações, filtros e parâmetros de configuração, performando testes estatísticos e testes de significância. Essa função é útil principalmente em um conjunto grande de dados (WITTEN et al., 2016).

O WEKA utiliza as abordagens de aprendizagem de máquina e mineração de dados para modelar séries temporais, pela transformação dos dados, em uma forma padrão que os algoritmos podem processar. O *software* remove a ordem temporal de cada entrada e codifica a dependência temporal, com o uso de um campo de entrada adicional. Os campos passam a serem referenciados através de uma variável de atraso (FRANK et al., 2016).

## 2.6 TRABALHOS RELACIONADOS

Na literatura, nota-se uma preferência para realizar previsões para grandes consumidores, como nações, nos trabalhos de Ogcü et al. (2012), Kaytez et al. (2015) e Vasconcelos et al. (2017), ou distribuidoras, como apresentado por Campos (2008) e Fogliatto et al. (2005).

O artigo de Castoril et al. (2018) propôs a validação de um sistema de previsão de demanda, em uma indústria de produção de laticínios, comprovando que há campo para aplicação de projeções e identificando a importância desta em plantas fabris.

Zanin et al. (2002), Cardoso (2004), Bueno (2008) e Frozza (2013), apesar de não trabalharem com previsões, trouxeram considerações importantes sobre o uso da energia e aplicação de medidas de eficiência energética em frigoríficos.

Em 1995, Nizami e Al-Garni (1995) propuseram a construção de uma rede neural para prever o consumo de energia elétrica em uma província da Arábia Saudita. Em comparação com o modelo de regressão, as redes neurais apresentaram melhores resultados. O trabalho de Azadeh et al. (2008) também encontrou resultados melhores com as redes neurais em comparação com séries temporais e simulação computacional para estimar o consumo de energia elétrica no Irã.

Outros trabalhos reforçam que a rede neural não é sempre o método mais adequado. Hernandez Neto e Fiorelli (2008), apesar de encontrarem uma melhor performance da rede neural quando comparada ao software EnergyPlus, não tiveram resultados satisfatórios. Kaytez et al. (2015) avaliou o modelo de LS-SVM, uma regressão por vetor de suporte, para previsão de consumo de energia elétrica na Turquia. Através da comparação com redes neurais, regressão e regressão por vetor de suporte, observou-se que o LS-SVM foi considerado melhor que os outros devido a maior sensibilidade nesse estudo. No trabalho de Ogcü et al. (2012) a regressão por vetor de suporte também teve uma performance superior as redes neurais. Apesar da rede neural ser capaz de prever o consumo de energia elétrica na Turquia, com base no histórico de consumo mensal, o método SVR fez uma previsão mais próxima da realidade.

Witten et al. (2016) e Usha e Balamurugan (2016) utilizaram o *software* WEKA para realizar suas previsões. Usha e Balamurugan (2016) utilizaram a ferramenta de séries temporais do WEKA para identificar padrões nos dados de consumo de energia elétrica. Witten et al. (2016) propôs um modelo de predição de cargas diárias e comparou os dados com regressão linear, utilizando essas informações para fins didáticos do WEKA.

### 3 MATERIAIS E MÉTODOS

Nesta seção, aborda-se a metodologia utilizada nesta pesquisa, detalhando os métodos propostos para realização dos experimentos, bem como as técnicas usadas.

#### 3.1 METODOLOGIA

Para o desenvolvimento do trabalho, no que diz respeito à abordagem, é uma pesquisa quantitativa, que enfatiza o raciocínio dedutivo, lógico e argumentos mensuráveis (GERHARDT; SILVEIRA, 2009). Os procedimentos estruturados para a coleta de dados e a análise numérica, através de procedimentos estatísticos, adequa-se a este conceito. Com relação à natureza, é uma pesquisa aplicada, pois busca soluções a problemas específicos, gerando conhecimentos para aplicação prática (GERHARDT; SILVEIRA, 2009). Considera-se também um estudo de caso, uma vez que conhece com profundidade as características de uma unidade consumidora específica, sem interferência do pesquisador no objeto a ser estudado (MARCONI; LAKATOS, 2003). Busca familiaridade com um problema, através de levantamento bibliográfico, experiências práticas e análise de exemplos que estimulem a compreensão (GERHARDT; SILVEIRA, 2009).

#### 3.2 APRESENTAÇÃO DA UNIDADE CONSUMIDORA

A empresa, com sede no oeste do Paraná, atua em diversos ramos, entre eles, redes de postos de gasolina e supermercados, unidades de recebimento de grãos e indústrias. Junto à instituição, identificou-se que a área com custos mais representativos, relacionados à energia elétrica, é uma das unidades industriais de aves.

O processo, na planta, inclui o recebimento das aves, abate, congelamento e armazenamento. Além do frango inteiro, a unidade produz cortes individuais, e também a transformação da carne de frango em outros produtos, como empanados e linguiças. Nessa indústria, em conformidade com o trabalho de Frozza (2013), identificou-se a casa de máquinas como o

principal consumo energético, onde estão localizados os compressores para geração de frio.

Devido a seu grande porte, a empresa optou pelo atendimento em 138 kV, com subestação própria. Apesar dos custos de implementação da subestação, as tarifas da distribuidora são mais baixas para essa tensão, tanto para o consumo quanto para a demanda (BARROS et al., 2020).

Além das considerações de tamanho e custos com energia, a escolha desta unidade consumidora está vinculada a operacionalidade da mesma. A unidade em questão deixou o ambiente cativo e migrou para o mercado livre em 2016, visando economia ao poder optar por diferentes fornecedores de energia elétrica.

### 3.3 OBTENÇÃO DOS DADOS DE ENTRADA

Como a unidade consumidora, em estudo, faz parte do mercado livre de energia, os dados de consumo são disponibilizados, pela CCEE, na plataforma de Sistema de Coleta de Dados de Energia (SCDE). O relatório do sistema fornece informação do ponto de medição, data, hora, tipo de medição (bruta ou líquida) e energia ativa gerada e consumida, assim como energia reativa gerada e consumida. A origem da medição, que pode ser do tipo coleta diária, inspeção lógica, hora ajustada ou hora faltante, também faz parte dos dados fornecidos. As tabelas 1 e 2 mostram um trecho da planilha. Como a unidade em questão não é produtora de energia, as colunas referentes a medição ativa e reativa de geração foram omitidas, neste trabalho, para facilitar a visualização.

**Tabela 1: Informações do ponto de medição, data, hora e tipo de energia disponibilizadas no SCDE**

Ponto de medição	Data	Hora	Tipo de energia
PRLAMTENTR101	01/11/2016	1	Líquida
PRLAMAENTR101	01/11/2016	2	Líquida
⋮	⋮	⋮	⋮
PRLAMTENTR101	30/11/2016	23	Líquida
PRLAMAENTR101	30/11/2016	24	Líquida

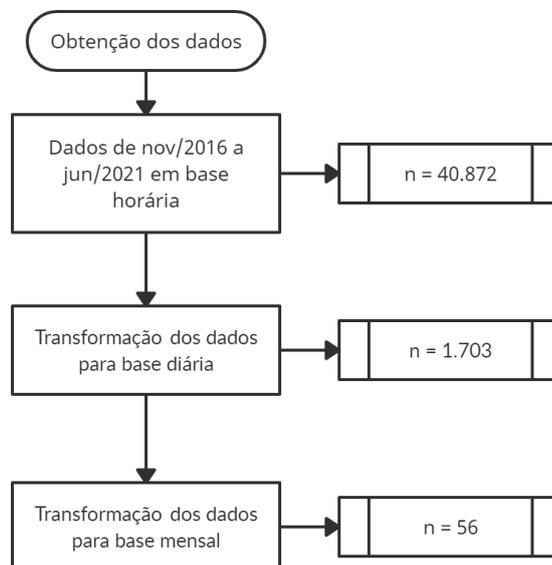
**Fonte: Autoria própria (2021)**

**Tabela 2: Medições ativa e reativa de consumo disponibilizadas no SCDE**

Ativa Consumo (kWh)	Reativa Consumo (kVARh)	Origem da Coleta
1036,80	281,52	Inspeção Lógica
1054,80	287,28	Inspeção Lógica
⋮	⋮	⋮
210,00	0	Hora Ajustada
209,00	0	Hora Ajustada

**Fonte: Autoria própria (2021)**

Os dados de consumo da unidade, em MWh, foram obtidos no período de novembro de 2016 a junho de 2021. Os arquivos são disponibilizados no formato *Comma-Separated Values* (CSV). Como as análises são feitas em base mensal, optou-se por transformar os dados para o consumo acumulado no mês. Na figura 9, o fluxograma mostra o processo aplicado e o efeito no número de amostras. Ao final da transformação, o banco de dados continha 56 amostras.

**Figura 9: Fluxograma da transformação dos dados de base horária para mensal**

**Fonte: Autoria própria (2021)**

Nos dados brutos, é possível encontrar uma série de problemas, como incompletude, discrepâncias ou ruídos, que podem atrapalhar o aprendizado dos algoritmos (CASTRO; FERRARI, 2016). O processo de limpeza e transformação dos dados é feito automaticamente pelo WEKA (WITTEN et al., 2016). Os dados, depois da limpeza e transformação, ficaram prontos para serem usados como entradas dos modelos.

### 3.4 ANÁLISE ESTATÍSTICA

Visando a exploração e um entendimento inicial da base de dados, foram usadas ferramentas de estatística descritiva, conforme recomendado por Castro e Ferrari (2016). Através destas, são obtidas informações sobre a planta em questão, como padrões de consumo mensais, anuais e expectativa de crescimento.

Observando os dados em cada período, foram identificados padrões de consumo. Além da avaliação dos gráficos, optou-se pelo teste de Tukey, através do *software* SISVAR, com nível de significância de 5% para classificar as variações de consumo.

O levantamento e diagnóstico do consumo de uma planta são importantes, não somente para controle, mas também para identificação de oportunidades de conservação. O conhecimento de áreas com uso significativo de energia é fundamental para priorizar ações com maiores resultados (FROZZA, 2013). Além disso, a quantificação expõe quais etapas e setores são mais representativas no custo total de produção, oportunizando redução de custos e aumento de produtividade. Conseqüentemente, torna o produto mais competitivo no mercado (BUENO, 2008).

Finalmente, usando o *software* Excel, foram geradas equações com base em funções logarítmica, linear, potência, exponencial e polinomial de 2º grau, na tentativa de descrever o consumo da unidade. Essa regressão foi avaliada através do coeficiente de correlação de Pearson.

### 3.5 CONFIGURAÇÕES DO WEKA

O desenvolvimento dos modelos foram implementados através do *software* WEKA, também usado por Usha e Balamurugan (2016) e Vasconcelos et al. (2017). Para inserir os dados, foi usado o formato ARFF, um tipo de arquivo de atributo e relação. Uma amostra, de como o arquivo foi construído, é apresentada na figura 10. Os requisitos são escritos por linhas, com os valores de cada atributo separados por uma vírgula. Para a inclusão dos dados, o formato de data, usado pelo WEKA, é o proposto pela ISO-8601, conforme instruções de Witten et al. (2016).

**Figura 10: Banco de dados no formato ARFF**

```

1      % Arquivo no formato ARFF
2      %
3      @relation 'Banco de dados - 2016 a 2020'
4
5      @attribute Data date "MM-yyyy"
6      @attribute Consumo numeric
7
8      @data
9      11-2016,5851,370
10     12-2016,6474,573
11
12     .
13     .
14     12-2020,8811,488

```

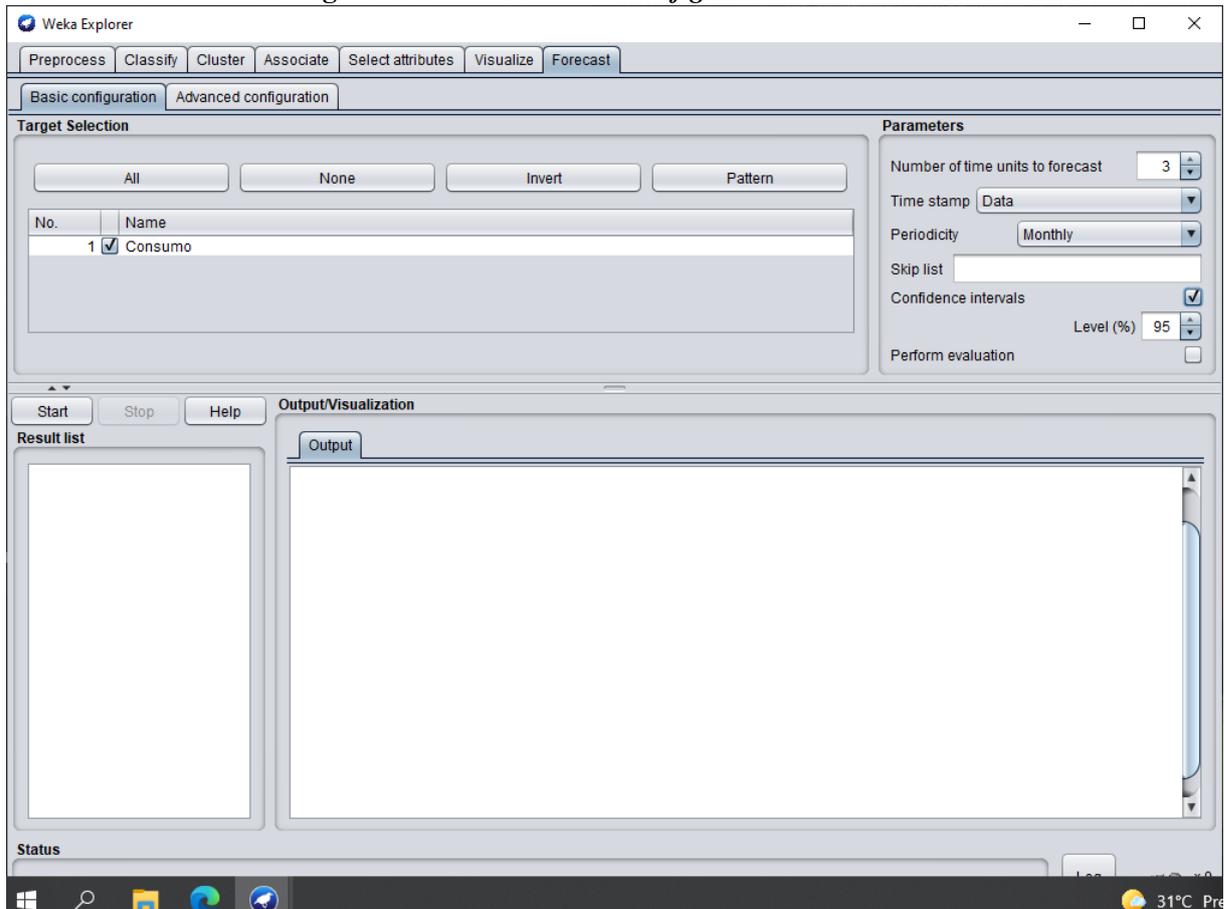
**Fonte: Adaptado de Witten et al. (2016).**

Ao fazer o *upload* da base de dados, na interface *Explorer*, o próprio programa já fornece algumas medidas de dispersão, como valores mínimos, máximos, média e desvio padrão. O WEKA fornece diversas possibilidades, entre elas classificação, agrupamento (*cluster*), visualização e, o de maior interesse para essa pesquisa, a previsão, utilizando a aba *Forecast*. Ao entrar na aba *Forecast* são disponibilizadas as áreas de *Basic configuration* e *Advanced configuration*, que serão detalhadas na sequência.

### 3.5.1 AMBIENTE *EXPLORER* - *BASIC CONFIGURATION*

A janela da *Basic configuration* é mostrada na figura 11. A seção *Target selection* permite que o usuário selecione as variáveis para as quais deseja fazer a previsão, que no caso foi o consumo.

**Figura 11: Interface *Basic configuration* do WEKA**



**Fonte: A autoria própria (2021)**

Na seção *Parameters*, configuram-se as variáveis relacionadas as unidades de tempo e a avaliação da previsão. O *Number of time units* controla o horizonte da previsão. Como foram feitas as previsões para seis meses a frente, ou seja, para os meses de janeiro, fevereiro, março, abril, maio e junho de 2021, foi atribuído o valor 6. O *Time stamp* foi configurado para que os dados tenham um rótulo do tipo *datas*. O campo *Periodicity* foi cadastrado como mensal. As caixas *Confidence intervals* e *Perform evaluation*, para avaliação, foram acionadas.

### 3.5.2 AMBIENTE EXPLORER - ADVANCED CONFIGURATION

As opções de configurações avançadas são apresentadas na aba *Advanced configuration*, onde seis outras telas de configuração podem ser acessadas através de abas.

O WEKA fornece diferentes bases de aprendizagem que podem ser utilizadas, sendo essas acessadas na aba *Basic learner*. Dentre elas, a regressão linear, o processo gaussiano, a *perceptron* multicamadas (MLP) e *SMOReg*, uma máquina de vetores de suporte. Cada algoritmo

de aprendizagem permite a seleção de mais opções de configuração específicas.

A aba *Lag creation* permite que o usuário controle e manipule como são criadas as variáveis de atraso (do inglês, *lagged*). Nesta tela também pode ser implementada o ajuste de variância dos dados de entrada, através do campo *Ajust for variance*.

As abas *Periodic attributes* e *Overlay data* não foram usadas. A primeira customiza a periodicidade dos trabalhos. A segunda restringe os dados que serão usados como entradas.

As métricas de avaliação são selecionadas na aba *Evaluation*, que podem ser usadas de modo único ou combinadas em qualquer sequência que o usuário desejar. São fornecidas sete opções. Dentre elas, para esse trabalho foram selecionadas as seguintes:

- MAE - *Mean Absolute Error* (Erro absoluto médio)
- MAPE - *Mean Absolute Percentage Error* (Erro absoluto médio percentual)

Nesta aba também é configurada a porção dos dados que será retirada do treinamento para uma validação independente. Uma vez que o processo de aprendizagem pode ser comparado a uma parametrização, é necessário dividir o conjunto de dados disponíveis em treinamento e teste. Durante o treinamento, o modelo se ajusta de acordo com os dados. Depois de treinado, os dados de teste avaliam o desempenho. Essa separação também evita o ajuste excessivo do modelo (WITTEN et al., 2016; HAYKIN, 2001).

Witten et al. (2016) sugere o uso de dois terços dos dados para treinamento e o restante para validação. Uma recomendação semelhante é dada por Haykin (2001), que propõe 60 % para treinamento e 40 % para teste. Neste trabalho optou-se por usar 70 % dos dados para treinamento e 30 % para teste. Para treinamento e teste foram usados apenas os dados de novembro de 2016 a dezembro de 2020. Desse modo, foram apresentados 35 dados para treinamento e 15 para teste. Os dados restantes, de 2021, foram usados para uma validação posterior.

A aba *Output* controla os elementos gráficos e textuais que serão produzidos e exibidos.

### 3.6 CONFIGURAÇÕES DOS MODELOS

Para a base de aprendizagem podem ser selecionados diferentes algoritmos, cada qual com suas características. Neste trabalho, devido aos bons resultados encontrados na literatura, optou-se por utilizar modelos de regressão, redes neurais e máquina de vetor de suporte.

A escolha da arquitetura adequada, para um dado conjunto, envolve uma série de fatores, como número de dados, a quantidade de ruído, a complexidade da função e a distribuição

estatística dos dados de treinamento. Seguindo a recomendação de Faceli et al. (2016), foi realizado um processo de tentativa e erro para avaliar as diferentes configurações.

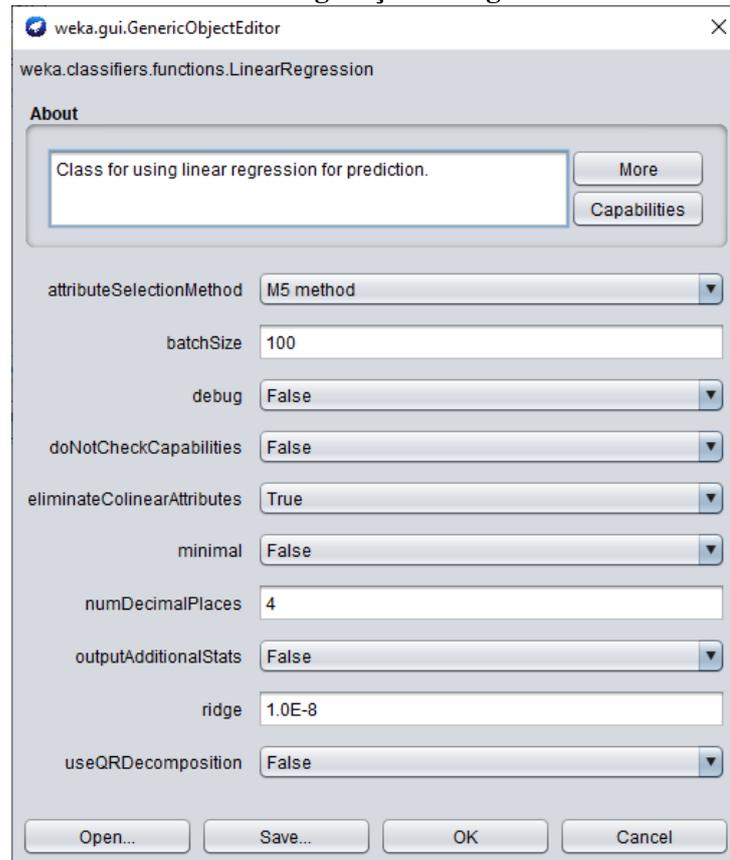
São realizadas alterações nos atributos, de acordo com as características de cada método de aprendizagem. A cada alteração, analisa-se o MAPE e o MAE do conjunto de teste. Pelo MAPE médio, julga-se melhoras e pioras nos modelos a cada modificação. Esse processo foi feito visando identificar as configurações ótimas, para cada algoritmo, para a previsão do consumo de energia elétrica da unidade em estudo.

Como cada algoritmo de aprendizagem possui suas particularidades, a metodologia usada para regressão linear, redes neurais e máquina de vetores de suporte são mostradas nos itens 3.6.1, 3.6.2 e 3.6.3, respectivamente.

### 3.6.1 REGRESSÃO

Para a regressão linear, configura-se os itens mostrados pela figura 12. O próprio *software* fornece informações sobre cada um dos elementos disponíveis para ajuste. O desempenho da regressão linear, quando os dados de treinamento estão correlacionados, pode ser reduzido. O WEKA, quando o campo *eliminateColinearAttributes* estiver ativado, detecta e remove os atributos de entrada altamente correlacionados.

**Figura 12: Interface de configuração da regressão linear no WEKA**



**Fonte: Autoria própria (2021)**

Segundo Vasconcelos (2017), o parâmetro de *Ridge* também ajuda a ajustar essa situação, evitando uma inflação dos coeficientes. A caixa *ridge* permite a administração do valor do parâmetro de Ridge. Visando identificar o melhor desempenho da regressão, foram testados os valores de  $1 \times 10^{-12}$ ,  $1 \times 10^{-8}$ ,  $1 \times 10^{-5}$ ,  $1 \times 10^{-3}$  e 1.

Para o método de seleção de atributos, foram testados o método M5, *Greedy* e sem método. Também pode ser utilizada a decomposição QR, ativada pelo campo *useQRDecomposition*. Porém, optou-se por não utilizar esse artifício.

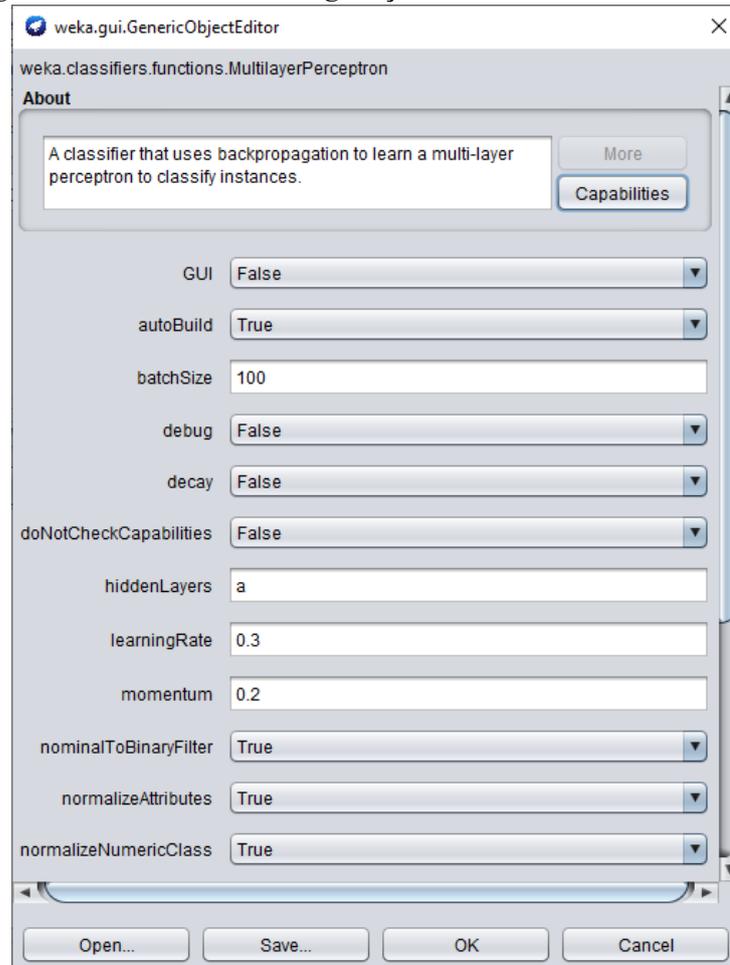
Após a definição dessas características, foi testado como o modelo responde a um ajuste de variância no conjunto de dados. São comparados os resultados com e sem esse ajuste, visando identificar qual o mais adequado.

### 3.6.2 REDE NEURAL

Para as redes neurais, conforme a tela exibida na figura 13, o WEKA fornece diversas opções de configurações. Como a rede neural pode ser personalizada em muitos pontos, foi

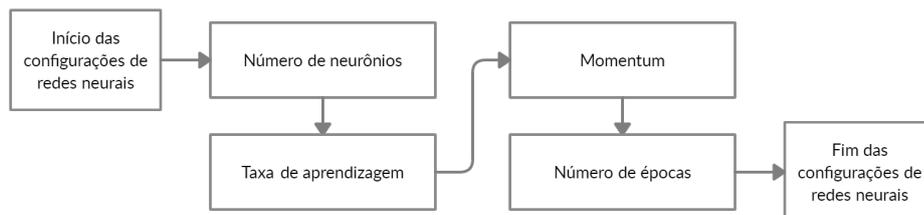
estabelecida uma ordem na configuração dos hiperparâmetros, conforme a figura 14. Verifica-se, através do MAPE, qual ajuste é o mais adequado e o hiperparâmetro é configurado. Considera-se, então, que o item em questão foi ajustado com o seu melhor valor e passa-se a avaliar o próximo hiperparâmetro.

**Figura 13: Interface de configuração das redes neurais no WEKA**



Fonte: Autoria própria (2021)

**Figura 14: Fluxograma para configuração das redes neurais**



Fonte: Autoria própria (2021)

O primeiro item a ser avaliado é o número de neurônios na camada oculta. No campo *HiddenLayers* o usuário configura e personaliza o número neurônios na camada oculta. Para Haykin (2001), uma única camada oculta é suficiente para que uma *perceptron* multicamadas consiga aproximar uma função. É menos provável que uma rede de tamanho mínimo aprenda idiossincrasias ou ruídos dos dados de treinamento. Entretanto, não há garantias que a única camada oculta é ótima no sentido de tempo de aprendizagem, facilidade de implementação e capacidade de generalização.

Para grande parte das aplicações, uma rede com uma única camada oculta é suficiente. De acordo com (HAYKIN, 2001) é preciso analisar os problemas causados pela interação dos neurônios em uma rede de camada única. Tendo em vista essa situação, foram testadas redes com 1, 2, 5, 10 e 50 neurônios na camada oculta.

Em seguida, foi feito o ajuste da taxa de aprendizagem  $\eta$ , que tem grande influência no tempo de convergência da rede. Caso seja pequena, podem ser necessárias muitas iterações para alcançar um bom modelo. Por outro lado, a escolha de uma taxa elevada provoca oscilações e dificulta a convergência (FACELI et al., 2016). Configura-se a taxa de aprendizagem da rede para 0, 1, 0,3, 0,5, 0,7 e 1. Os resultados são analisados pelo MAPE e pelo MAE.

A adição do *momentum* pode amenizar os problemas de convergência de uma taxa de aprendizado alta, uma vez que quantifica o grau da importância da variação de peso de um ciclo para outro (FACELI et al., 2016). Foram testados os valores de 0, 1, 0,3, 0,5, 0,7 e 0,9.

Faceli et al. (2016) destaca a natureza estocástica das redes neurais, que tem como consequência a dependência da ordem de apresentação dos objetos e dos pesos iniciais dos neurônios. O ajuste dos pesos é realizado em função do número de épocas configurados. Para este trabalho foram feitos testes com 1, 100, 500, 1000 e 5000 épocas. Caso tenham poucas épocas, os pesos não são suficientemente ajustados, gerando um erro grande. Porém, um valor grande pode causar *overfitting*, quando a rede deixa de aprender e se torna superajustada aos dados de treinamento (HAYKIN, 2001; FACELI et al., 2016).

Finalizadas as configurações da rede, o WEKA oferece algumas opções de tratamento de dados, como transformação e normalização, através dos filtros *NominalToBinary*, *NominalAttributes* e *NormalizeNumericClass*. Todas essas funções foram acionadas, visando acelerar a resposta. Cabe ressaltar que a normalização é apenas interna, ou seja, a saída será dada na escala original.

Por fim, da mesma forma que para regressão, foi testado se o ajuste na variância do conjunto de dados causa uma melhora ou não na performance.

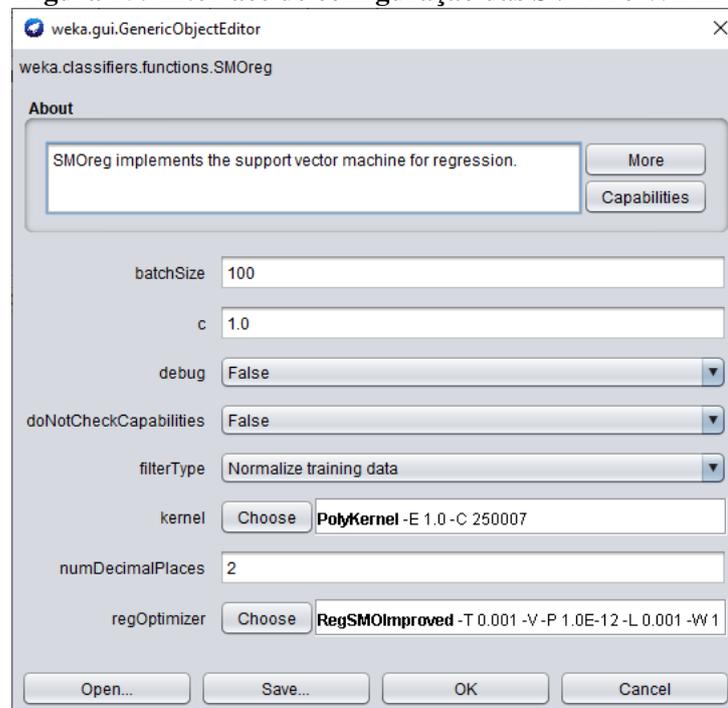
### 3.6.3 MÁQUINA DE VETORES DE SUPORTE

A tela para a máquina de vetores de suporte é mostrada pela figura 15. Ao contrário das redes neurais, a máquina de vetores de suporte possui poucos itens configuráveis.

Para o tratamento dos dados, o campo *filterType* determina se os dados serão normalizados, padronizados ou se não serão transformados. Os dados foram normalizados, mas não foram padronizados ou transformados.

Os itens essenciais a serem definidos são o função *kernel*, o otimizador e o parâmetro *C*. Os mesmos foram avaliados em função do MAPE e do MAE para os dados de teste.

**Figura 15: Interface de configuração das SVM no WEKA**



**Fonte: Autoria própria (2021)**

Para que a SVM seja usada com sucesso, é necessário utilizar uma função de *kernel* adequada. Ela pode ser selecionada através do item *kernel*, que oferece as funções NPolyKernel, PolyKernel, Puk e RBF. De acordo com Bonesso (2013), na literatura, o mais comum é utilizar o linear e o RBF, usada por Ruas et al. (2008) e Kaytez et al. (2015). Contrariando esse indicativo, Vasconcelos et al. (2017) usou a função polinomial. Conforme Abreu (2016), a definição da função de *kernel* deve ser feita através de testes, uma vez que não se pode assumir que uma função é melhor que a outra, pois o desempenho depende do conjunto de dados utilizado.

Em seguida, no campo *regOptimizer*, foi selecionado o algoritmo de aprendizagem,

*regSMO* ou *regSMOImproved*, que possui uma adaptação do critério de parada.

Concluindo as configurações, o parâmetro  $C$  é ajustado no campo  $c$ . Ele controla a compensação dos erros cometidos na fase de treinamento e a margem de aceitação do algoritmo (BONESSO, 2013). Foram testados valores de 0, 1, 2, 5 e 10.

A avaliação do ajuste da variância também foi feita para a máquina de vetores de suporte.

### 3.7 AVALIAÇÃO E COMPARAÇÃO ENTRE OS MODELOS

A validação de uma análise pode ser através de diferentes formas, como a comparação com outras técnicas, verificação da capacidade do método gerado e, até mesmo, por um especialista. A avaliação de qual das medidas é a mais adequada depende da aplicação (CASTRO; FERRARI, 2016).

Em termos práticos, um método de previsão é bom independente da medida de erro usada. Porém, alguns pontos podem ser analisados. Witten et al. (2016) reforça que, para analisar a performance do algoritmo, a avaliação deve ser feita somente com dados do conjunto de teste.

Castro e Ferrari (2016) recomenda o uso do erro quadrático médio e da raiz do erro quadrático médio. São medidas de fácil manipulação matemática, porém podem amplificar os efeitos dos *outliers*. Esse aumento não ocorre para erros relativos, que são normalizados, de modo que valores próximos da média implicam em melhores previsões. Os erros absolutos também são uma solução, uma vez que todos os tamanhos dos dados são tratados de acordo com a sua magnitude (CASTRO; FERRARI, 2016; WITTEN et al., 2016).

Tendo em vista as questões expostas acima, optou-se por fazer a avaliação das previsões através do erro absoluto médio, o MAE, e o erro médio absoluto percentual, o MAPE. Apesar de ambos serem usados na literatura, como nos casos de Amaral (2020), e Campos (2008), o mais comum é avaliar o MAPE, usado por Castoril et al. (2018), Kaytez et al. (2015), Mordjaoui et al. (2017), Ogcü et al. (2012), Ruas et al. (2008), Vasconcelos et al. (2017), entre outros.

Para Vasconcelos et al. (2017), uma previsão é bem sucedida quando tem-se um MAPE inferior a 10 %. Já de acordo com Campos (2008), os erros não indicam claramente o desempenho de determinado previsor, uma vez que dependendo do nível da parcela estocástica, um MAPE de 10 %, por exemplo, pode ser um ótimo desempenho. De forma semelhante, para uma série puramente determinística e estacionária, um MAPE de 1 % indica uma performance péssima.

Considerando a natureza da variável energia elétrica, sensível a diversos fatores que

não são controláveis, como condições climáticas e econômicas, assim como os valores de MAPE observados nos trabalhos de Ruas et al. (2008), Campos (2008) e Ogcü et al. (2012), foi estabelecido que, neste trabalho, os modelos deveriam apresentar um erro inferior a 5 % para serem considerados adequados para a previsão.

Após a validação dos modelos de regressão linear, redes neurais e máquinas de vetores de suporte, passa-se para a comparação entre eles. Para isso, verifica-se a previsão de um semestre a frente, ou seja, de janeiro de 2021 a junho de 2021. As estimativas são confrontadas com o consumo observado. A partir desses dados são obtidos o MAPE, MAE e a raiz quadrada do erro médio (RMSE) para cada um dos algoritmos. Com essas informações, foram feitas considerações a respeito do desempenho, vantagens e desvantagens de cada.

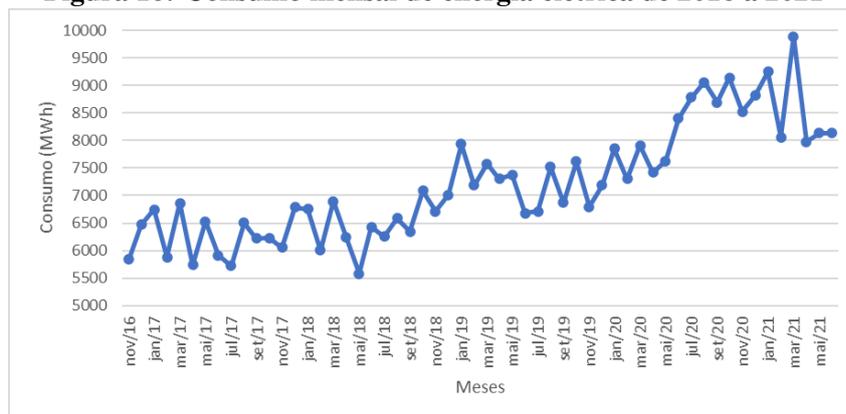
## 4 RESULTADOS E DISCUSSÕES

Nesta seção, apresentam-se os resultados obtidos neste trabalho. Inicia-se com uma análise estatística dos dados, visando compreender as características do consumo de energia elétrica na planta em questão. Na sequência, na seção 4.2, são apresentados os resultados obtidos dos modelos de regressão, redes neurais e SVM. Por fim, no item 4.3, apresentam-se as comparações das vantagens e desvantagens entre os algoritmos de aprendizado utilizados.

### 4.1 ANÁLISE ESTATÍSTICA

O consumo de energia elétrica mensal, em MWh, do período de novembro de 2016 a junho de 2021 é mostrado na figura 16. Os dados apresentam um coeficiente de variação de 13,93%, o que indica que são homogêneos. As oscilações, verificadas na figura 16, abrem margem para uma especulação sobre como os dias de trabalho e as condições climáticas influenciam no consumo de energia elétrica.

**Figura 16: Consumo mensal de energia elétrica de 2016 a 2021**



**Fonte: Autoria própria (2021)**

Conforme a tabela 3, que apresenta a média de consumo mensal, por ano, há um indicativo do aumento do uso da energia elétrica nesta indústria. Enquanto o consumo médio mensal de 2016 foi de 6.162,971 MWh, essa média subiu para 8.207,971 MWh em 2020.

**Tabela 3: Média do consumo mensal por ano**

Ano	Média de consumo mensal (MWh)
2016	6.162,971
2017	5.712,311
2018	6.491,943
2019	7.233,717
2020	8.207,971

**Fonte: Aatoria própria (2021)**

Para identificar se há uma grande variação entre os consumos mensais, aplicou-se o teste de Tukey com um nível de significância de 5% para a variável mês. Os resultados são mostrados na tabela 4. O teste de Tukey pode ser usado para localizar médias que são significativamente diferente das outras. Conforme a tabela 4, não há diferenças significativas, uma vez que as médias formam um grupo único.

**Tabela 4: Teste de Tukey para o tratamento mês**

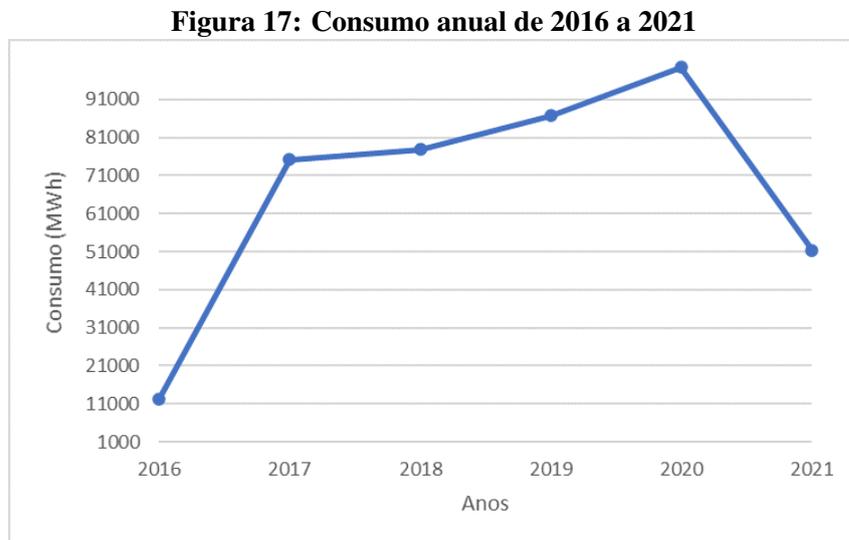
Tratamento (meses)	Média (MWh)	Resultado do teste
Agosto	5.954,587	a1
Maio	6.582,230	a1
Fevereiro	6.593,397	a1
Abril	6.674,038	a1
Novembro	6.787,717	a1
Junho	6.855,402	a1
Julho	6.867,857	a1
Setembro	7.035,462	a1
Dezembro	7.252,560	a1
Março	7.312,429	a1
Janeiro	7.320,247	a1
Outubro	7.516,884	a1

**Fonte: Aatoria própria (2021)**

Analisando as médias de consumo mensal, fornecidas pela tabela 4, observa-se que, os meses próximo ao inverno, que tem uma tendência ao clima mais ameno, apresentam uma média de consumo menor. A exceção é fevereiro, que tem um consumo baixo, apesar da alta temperatura. Entretanto, esse mês tem 28 ou 29 dias apenas.

A influência das condições climáticas também foi observada nos frigoríficos analisados por Zanin et al. (2002) e Frozza (2013). De acordo com eles, fatores externos, como temperatura e umidade, afetam o sistema de refrigeração, que é responsável por cerca de 80% do consumo de energia elétrica.

Contudo, a expansão da planta também deve ser considerada como uma variável de grande influência. A figura 17 mostra o consumo anual, de 2016 a 2021, com claros indicativos de crescimento do consumo a cada ano. Ressalta-se que os dados do primeiro e do último ano estão incompletos. Os dados de 2016 são somente a partir novembro, enquanto que os de 2021 são até junho.



**Fonte: Autoria própria (2021)**

Na sequência, com o consumo médio mensal, para cada ano, aplicou-se o teste de Tukey, com as mesmas condições anteriores. Os resultados são apresentados na tabela 5.

**Tabela 5: Teste de Tukey para o tratamento ano**

Tratamento (anos)	Média (MWh)	Resultado do teste
2017	5.712,311	a1
2016	6.162,971	a1
2018	6.491,943	a1
2019	7.233,717	a1/a2
2020	8.207,971	a2

**Fonte: Autoria própria (2021)**

O teste de Tukey, na tabela 5, separou a variável ano em dois grupos distintos. Observa-

se uma transição em 2019, quando os níveis de consumo tiveram um aumento significativo. Neste período houve expansão da planta, com inclusão de novos equipamentos e obras. A partir de 17/05/2020 houve aumento dos turnos de trabalho, com abates inclusive aos domingos. Antes disso, nestes dias eram mantidos apenas o frio e equipamentos de higienização.

Foi constada uma progressão do consumo ao longo dos anos, com exceção do ano de 2017, quando houve um decréscimo de 7% em relação ao consumo médio mensal do ano anterior.

A tabela 6 explora as possíveis linhas de tendência para explicar o crescimento do consumo mensal e os coeficientes de correlação de Pearson para cada uma delas. O coeficiente de correlação serve como um indicador relativo do grau de ajustamento. Para modelos mais lineares, quanto mais próximo de 1, melhor.

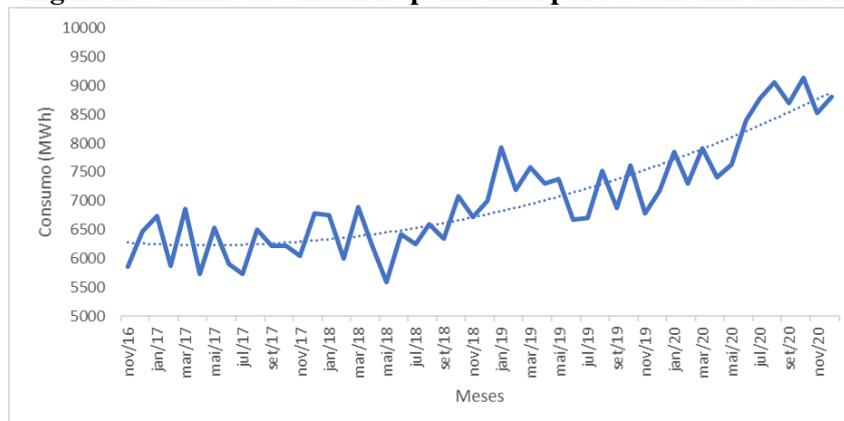
**Tabela 6: Equações de linha de tendência para o consumo mensal**

Tipo	Equação	Coeficiente de correlação
Logarítmica	$y = 76029\ln(x) - 804849$	0,7020
Linear	$y = 1,7538x - 69117$	0,7042
Potência	$y = 7x10^{-46}x^{10,561}$	0,7237
Exponencial	$y = 0,1781e^{0,0002x}$	0,7256
Polinomial 2º grau	$y = 0,0015x^2 - 129,55x + 2x10^6$	0,7849

**Fonte: Autoria própria (2021)**

O coeficiente de Pearson mais alto é obtido para a expressão polinomial. O gráfico com a linha de tendência para a função polinomial é exibido na figura 18. A equação em questão, apesar de auxiliar no entendimento do uso da energia, não consegue acompanhar com precisão as variações de consumo.

**Figura 18: Linha de tendência polinomial para o consumo mensal**



**Fonte: Autoria própria (2021)**

Fica evidente que as técnicas estatísticas, aplicadas neste trabalho, auxiliaram na análise dos dados. Apesar disso, a previsão do consumo, para essa planta, ainda é precária somente com essas ferramentas. Na sequência serão apresentados os resultados obtidos neste trabalho, usando o *software* WEKA e algoritmos de aprendizagem de máquina.

## 4.2 DESENVOLVIMENTO DOS MODELOS

Os itens 4.2.1, 4.2.2 e 4.2.3 apresentam os modelos de regressão linear, redes neurais e máquinas de vetor de suporte, respectivamente. Na sequência, na seção 4.3, procedeu-se com uma comparação entre os modelos.

### 4.2.1 REGRESSÃO LINEAR

Utilizando a interface do WEKA para regressão linear, foram configurados, para o parâmetro de *Ridge*, os valores de  $1 \times 10^{-12}$ ,  $1 \times 10^{-8}$ ,  $1 \times 10^{-5}$ ,  $1 \times 10^{-3}$  e 1. O WEKA, ao executar o algoritmo de regressão linear, forneceu, para o conjunto de testes, os valores médios dos MAEs e MAPEs, apresentados na tabela 7.

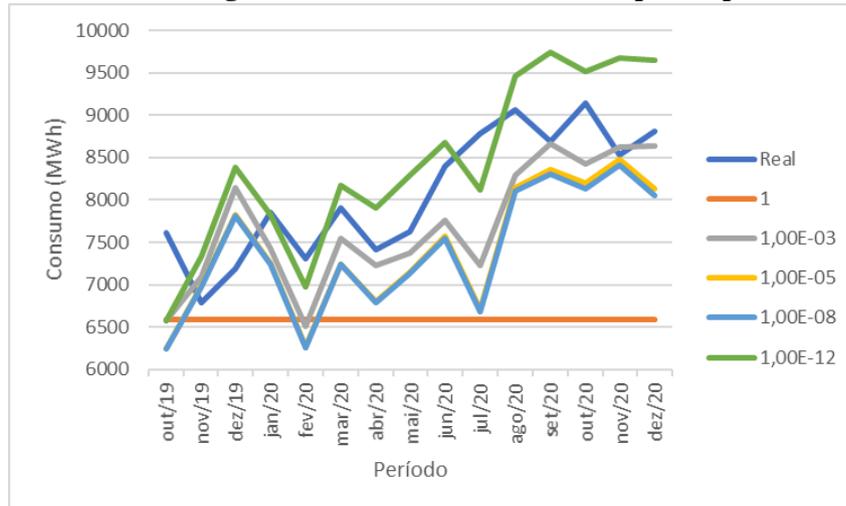
**Tabela 7: MAE e MAPE dos modelos de regressão linear para o parâmetro de *Ridge***

Ridge	MAE (MWh)	MAPE ( %)
$1e^{-12}$	592,689	7,31 %
$1e^{-8}$	787,031	9,56 %
$1e^{-5}$	763,268	9,28 %
$1e^{-3}$	553,033	6,78 %
1	1.541,459	18,33 %

**Fonte: Autoria própria (2021)**

Pode-se observar, da tabela 7, que a melhor performance, do modelo de regressão linear, foi obtida com o parâmetro de *Ridge*  $1 \times 10^{-3}$ , com um MAPE de 6,78%. As previsões para o período de teste, de outubro de 2019 a dezembro de 2020, são mostradas na figura 19.

**Figura 19: Previsões de regressão linear dos dados de teste para o parâmetro de Ridge**



**Fonte: Autoria própria (2021)**

Fica nítido, na figura 19, que o desempenho do modelo é pobre quando o parâmetro de Ridge é 1. O consumo é projetado como uma reta, que não se aproxima dos valores reais.

Para os outros valores de Ridge, notam-se também algumas discrepâncias entre os consumos previstos e reais, como nos meses de dezembro de 2019, julho de 2020 e novembro de 2020. Porém, de forma geral, as curvas apresentam um comportamento semelhante ao consumo efetivo.

Na sequência, depois da definição do parâmetro de Ridge, testa-se o algoritmo com e sem métodos de seleção de atributos. Utiliza-se, neste trabalho, os métodos de seleção de atributos M5 e Greedy. Os erros, para o conjunto de teste, são apresentados na tabela 8.

**Tabela 8: MAE e MAPE dos modelos de regressão linear para o método de seleção**

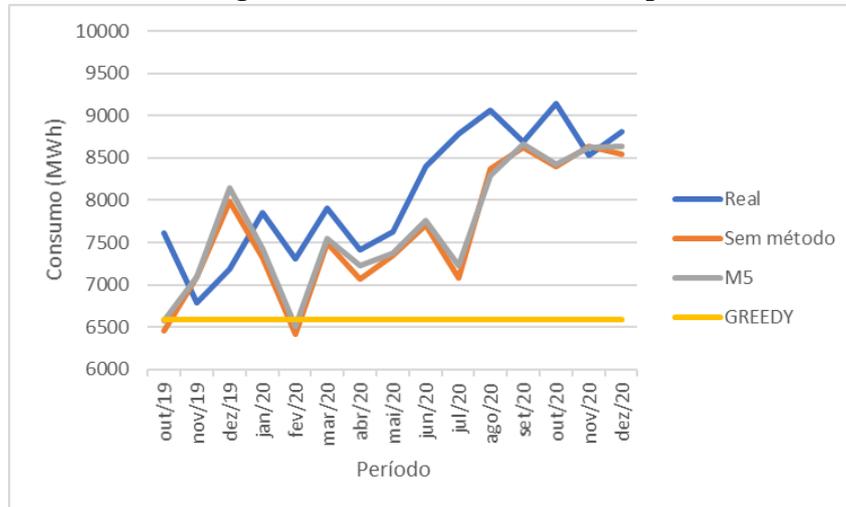
Método de seleção	MAE (MWh)	MAPE (%)
M5	553,033	6,78 %
Greedy	1.541,459	18,33 %
Sem método	603,687	7,40 %

**Fonte: Autoria própria (2021)**

Percebe-se, dos erros apresentados na tabela 8, que o método de seleção M5 obteve os melhores resultados, sendo, portanto, mais adequado para esse conjunto de dados. O método M5 trabalha com atributos e classes contínuas e gera uma *Regression Tree*, do inglês, uma árvore de regressão (DOSUALDO; REZENDE, 2003). A figura 20 apresenta os dados projetados, para os

dados de teste, para cada um dos métodos de seleção, em comparação com os do consumo real de outubro de 2019 a dezembro de 2020.

**Figura 20: Previsões de regressão linear dos dados de teste para o método de seleção**



**Fonte: Autoria própria (2021)**

Em seguida, realizou-se o teste do ajuste de variância, no item *Adjust for variance*, que compensa as variações nos dados de entrada. No caso da regressão, o ajuste afeta negativamente as previsões, conforme mostra a tabela 9

**Tabela 9: MAE e MAPE dos modelos de regressão linear para o ajuste de variância**

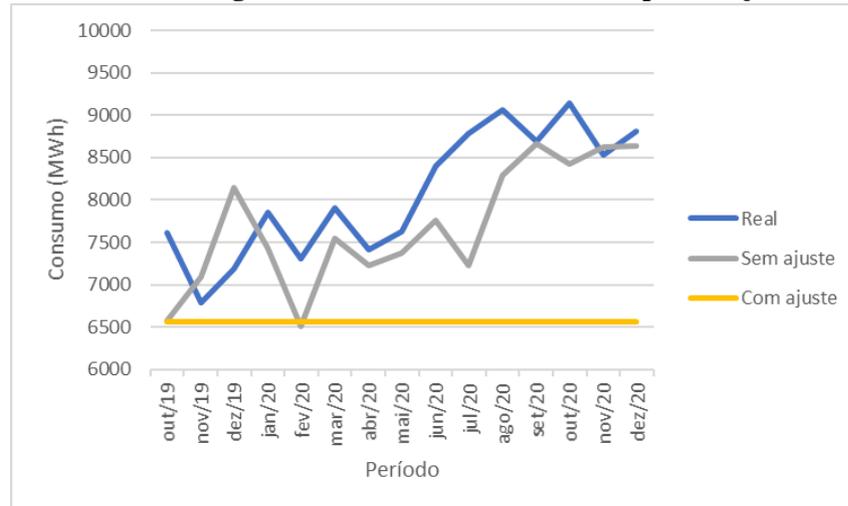
Ajuste de variância	MAE (MWh)	MAPE (%)
Sem ajuste	553,033	6,78 %
Com ajuste	1.565,432	18,62 %

**Fonte: Autoria própria (2021)**

Dessa forma, como o MAPE e o MAE aumentaram com o ajuste da variância, considerou-se que o mesmo não foi adequado, para esse conjunto de dados, e foi descartado. A figura 21 compara as projeções, com e sem o ajuste de variância.

Observa-se, da figura 21, que o ajuste tornou as previsões lineares. Tendo em vista que o consumo tem um comportamento variável, confirma-se que o ajuste não é adequado.

Em conclusão, o modelo escolhido tem as características apresentadas na tabela 10, que resulta em um MAPE, nos dados de treinamento, de 6,78%.

**Figura 21: Previsões de regressão linear dos dados de teste para o ajuste de variância**

**Fonte: Autoria própria (2021)**

**Tabela 10: Configuração do modelo de regressão linear**

Parâmetro	Configuração
<i>Ridge</i>	$1e^{-3}$
Método de seleção	M5
Ajuste da variância	sem ajuste

**Fonte: Autoria própria (2021)**

#### 4.2.2 REDES NEURAIAS

Visando obter um modelo, com resultados dentro dos níveis estabelecidos, configurou-se o número de neurônios na camada oculta, a taxa de aprendizagem, o *momentum* e o tempo de treinamento. Além disso, também foi avaliada uma adaptação nos dados de entrada, o ajuste de variância. Os hiperparâmetros foram alterados nessa sequência, partindo das configurações iniciais do WEKA e seguindo, com o modelo com o menor erro, para os testes da próxima variável.

A primeira configuração foi o número de neurônios. Foram testadas redes com 1, 2, 5, 10 e 50 neurônios. Para o conjunto de teste, o WEKA apresentou os erros médios mostrados na tabela 11.

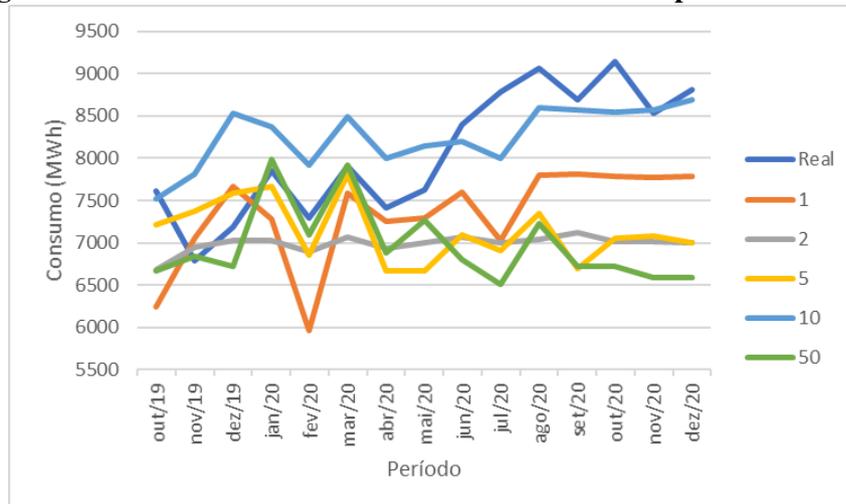
**Tabela 11: MAE e MAPE dos modelos de redes neurais para os neurônios**

Número de neurônios	MAE (MWh)	MAPE (%)
1	842,078	10,12 %
2	1.133,186	13,34 %
5	1.040,657	12,31 %
10	508,528	6,52 %
50	1.112,499	12,97 %

**Fonte: Autoria própria (2021)**

Confirma-se, da tabela 11, que, apesar de uma rede mais complexa ter mais neurônios, não necessariamente tem melhor desempenho. De fato, para esse conjunto de dados, o MAPE e o MAE são maiores com 50 neurônios do que com um.

Comparando o consumo real com os previstos, na figura 22, nota-se que o aumento verificado a partir de maio de 2020 só é acompanhado pela rede de 10 neurônios. Porém, nos meses anteriores, essa mesma rede superestimou o consumo de energia da unidade. Apesar disso, analisando o MAPE e o MAE, na tabela 11, comprova-se que a rede neural com 10 neurônios é a mais adequada para o conjunto de dados em estudo.

**Figura 22: Previsões de redes neurais dos dados de teste para os neurônios**

**Fonte: Autoria própria (2021)**

Definido o número de neurônios, segue-se para o ajuste da taxa de aprendizagem. Os erros de previsão, para o conjunto de teste, para as taxas de aprendizagem de 0, 1, 0, 3, 0, 5, 0, 7 e 1, são mostrados na tabela 12.

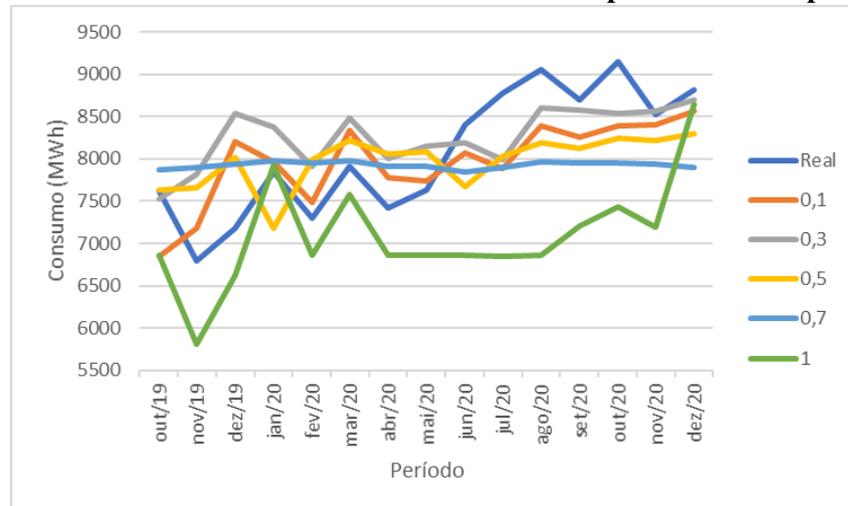
**Tabela 12: MAE e MAPE dos modelos de redes neurais para a taxa de aprendizagem**

Taxa de aprendizagem	MAE (MWh)	MAPE ( %)
0,1	438,538	5,40 %
0,3	508,538	6,52 %
0,5	612,978	7,59 %
0,7	651,514	7,95 %
1	1.140,685	13,58 %

**Fonte: Autoria própria (2021)**

Observa-se, na tabela 12, que o erro cresce a medida que aumenta a taxa de aprendizagem. Desse modo, a taxa de aprendizagem de 0,1 foi a mais adequada para o conjunto de dados, foi selecionado, neste trabalho,  $\eta = 0,1$ .

Analisando as projeções mensais, que são dadas na figura 23, observa-se que as previsões foram mais certas nos meses iniciais, de outubro de 2019 a maio de 2020. A partir de junho de 2020, o consumo teve um crescimento, que foi maior que as estimativas das redes neurais.

**Figura 23: Previsões de redes neurais dos dados de teste para a taxa de aprendizagem**

**Fonte: Autoria própria (2021)**

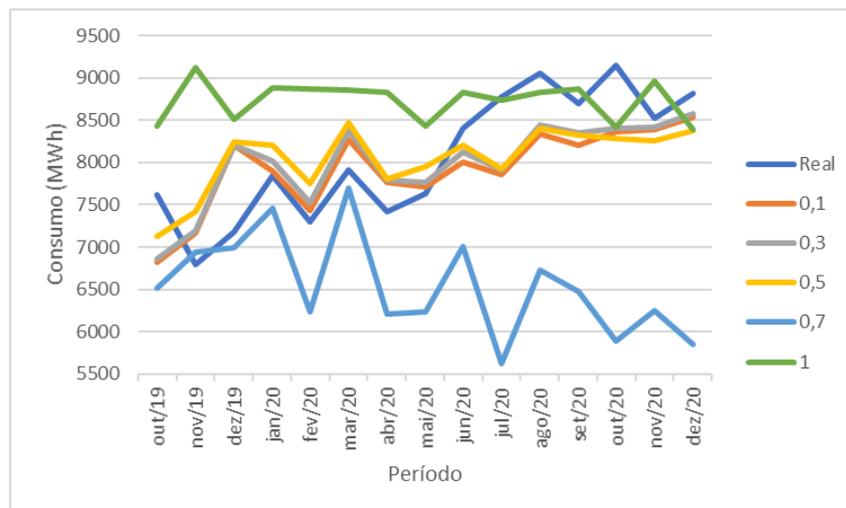
Passa-se a configurar o próximo hiperparâmetro, o *momentum*, em 0,1, 0,3, 0,5, 0,7 e 1. Os erros médios resultantes, para cada um desses valores, são apresentados na tabela 13.

**Tabela 13: MAE e MAPE dos modelos de redes neurais para o *momentum***

<i>Momentum</i>	MAE (MWh)	MAPE (%)
0,1	443,656	5,43 %
0,3	432,599	5,32 %
0,5	521,045	6,47 %
0,7	1.429,353	16,85 %
0,9	804,071	10,53 %

**Fonte: Autoria própria (2021)**

O *momentum* em 0,1 ou 0,3 geram erros semelhantes. A partir deste valor, observa-se, então, um crescimento do erro em função do *momentum*. As maiores diferenças entre o consumo e as estimativas são para o *momentum* de 0,7. A figura 24 apresenta as previsões, do período de outubro de 2019 a dezembro de 2020, para cada um dos *momentums* testados.

**Figura 24: Previsões de redes neurais dos dados de teste para o *momentum***

**Fonte: Autoria própria (2021)**

Observa-se, na figura 24, que, enquanto as previsões com o *momentum* 1 superestimou o consumo, para o *momentum* 0,7 subestimou. De todo modo, os dois são inadequados para o conjunto de dados.

As projeções, com o *momentum* em 0,1, 0,3 e 0,5, apresentaram comportamentos próximos e tiveram um bom desempenho, principalmente no período de janeiro de 2020 a abril de 2020. Definiu-se o *momentum* 0,3 como o valor mais adequado, pelo menor MAPE.

Em seguida, foram analisadas as previsões para 1, 100, 500, 1000 e 5000 épocas. O

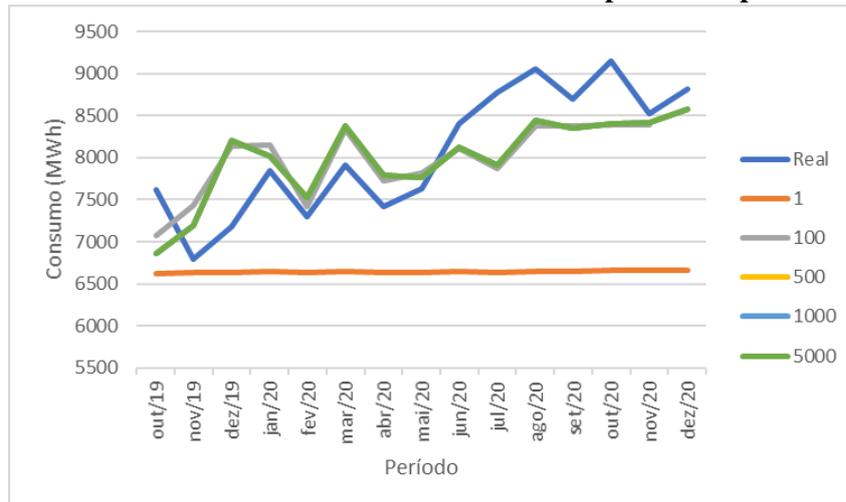
erros MAE e MAPE são mostrados na tabela 14. As previsões de outubro de 2019 a dezembro de 2020, de acordo com a variação do número de épocas, é apresentada na figura 25.

**Tabela 14: MAE e MAPE dos modelos de redes neurais para o tempo de treinamento**

Número de épocas	MAE (MWh)	MAPE ( %)
1	1.486,940	17,66 %
100	444,419	5,49 %
500	432,599	5,35 %
1000	432,258	5,35 %
5000	432,257	5,35 %

**Fonte: Autoria própria (2021)**

**Figura 25: Previsões de redes neurais dos dados de teste para o tempo de treinamento**



**Fonte: Autoria própria (2021)**

Na tabela 14, nota-se que não há grande melhora no desempenho quando o tempo de treinamento é maior que 500 épocas. De fato, a diferença entre 500, 1000 e 5000 épocas só é perceptível no MAE, e não no MAPE.

Observa-se, na figura 25, com exceção de uma única época de treinamento, que tem um péssimo desempenho, as outras previsões apresentam resultados próximos. De toda forma, como o aumento do esforço computacional não foi representativo, optou-se por utilizar 5000 épocas.

Por fim, testou-se o ajuste de variância no conjunto de dados. As previsões mensais são apresentadas na figura 26 e os erros médios na tabela 15. Apesar de ambas as curvas de previsão, da figura 26, apresentarem comportamentos próximos, a tabela 15 indica que o ajuste

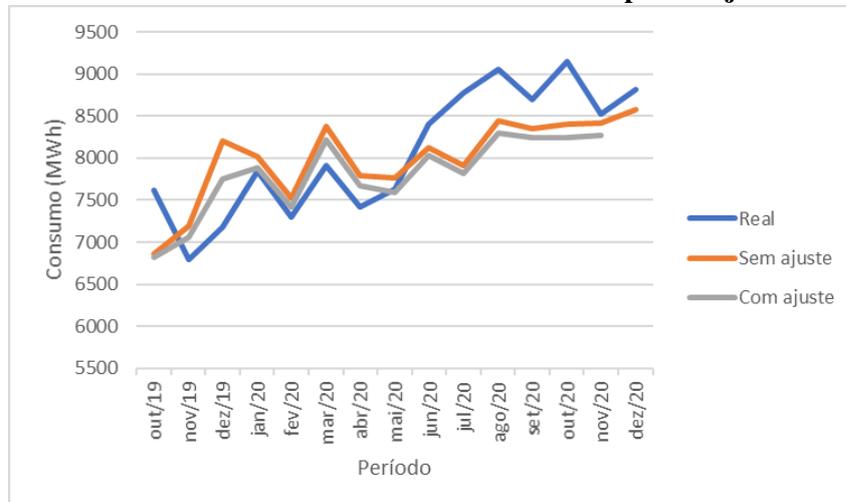
na variância implicou em uma melhora no desempenho da rede. De fato, esse ajuste faz com que o MAPE seja de 4,87%.

**Tabela 15: MAE e MAPE dos modelos de redes neurais para o ajuste de variância**

Ajuste da variância	MAE (MWh)	MAPE (%)
Sem ajuste	432,257	5,35 %
Com ajuste	406,454	4,87 %

**Fonte: Autoria própria (2021)**

**Figura 26: Previsões de redes neurais dos dados de teste para o ajuste de variância**



**Fonte: Autoria própria (2021)**

Assim, concluem-se as configurações da rede neural. A melhor performance foi com o formato mostrado na tabela 16. Destaca-se que o MAPE foi inferior a 5%, de modo que esta abordagem pode ser considerada como adequada, para a aplicação em previsão do consumo de energia elétrica da unidade industrial em estudo.

**Tabela 16: Configuração do modelo de redes neurais**

Parâmetro	Configuração
número de neurônios	10
taxa de aprendizagem	0,1
<i>momentum</i>	0,3
tempo de treinamento	5000
ajuste da variância	com ajuste

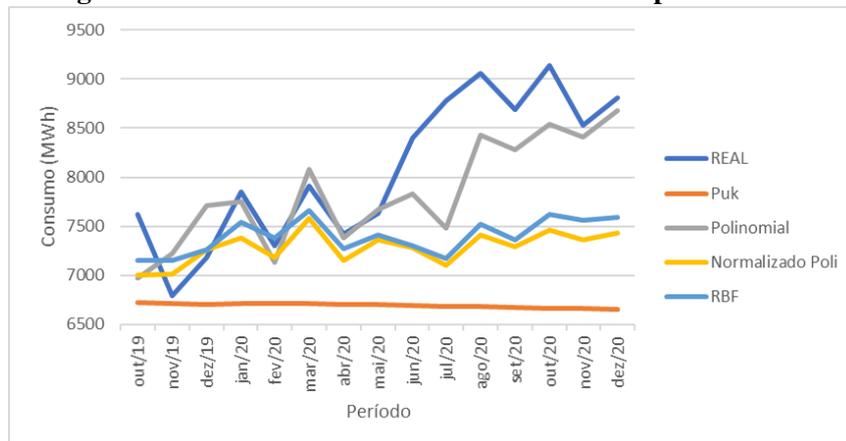
**Fonte: Autoria própria (2021)**

### 4.2.3 MÁQUINA DE VETORES DE SUPORTE

Para o modelo de máquina de vetores de suporte foram configurados a função de kernel, o algoritmo otimizador e o parâmetro  $C$ . As alterações foram feitas sequencialmente, sempre selecionando o resultado com menor erro para configurar a próxima variável.

Iniciou-se o processo selecionando a função de kernel. A figura 27 mostra as previsões, para os dados de teste, usando a função polinomial normalizada, polinomial, PUK e RBF.

**Figura 27: Previsões de SVM dos dados de teste para o kernel**



**Fonte: Autoria própria (2021)**

Observa-se, da figura 27, que a função Puk gera um resultado quase linear, de modo que não é adequada para este trabalho. As funções de *kernel* polinomial normalizada e RBF apresentam resultados semelhantes. Porém, nenhuma das duas consegue acompanhar o crescimento do consumo real observado a partir de maio de 2020. A função polinomial, apesar de também apresentar uma discrepância a partir de maio, volta a sugerir um aumento de consumo a partir do mês de agosto de 2020. Em dezembro de 2020, a previsão de consumo da função polinomial e o consumo real já são próximos novamente.

O MAPE e o MAE médios para cada função, mostrados na tabela 17, confirmam que a função mais adequada para o conjunto de dados, em estudo, é a função polinomial. Ela apresenta um MAE de  $381,366MWh$  e MAPE de  $4,58\%$ .

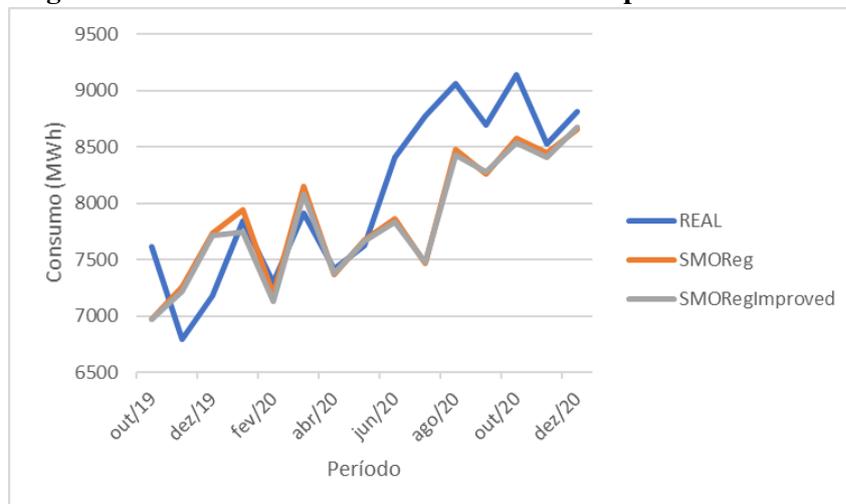
**Tabela 17: MAE e MAPE dos modelos de SVM para o kernel**

Função Kernel	MAE (MWh)	MAPE (%)
Polinomial	381,366	4,58 %
Polinomial normalizada	858,689	10,00 %
Puk	1.437,275	17,02 %
RBF	778,544	9,05 %

**Fonte: Autoria própria (2021)**

Ressalta-se, desde já, a boa atuação da aprendizagem por máquina de vetor de suporte. O MAPE encontrando de 4,58% é inferior ao mínimo estabelecido de 5%, mesmo sem a implementação de outros ajustes.

Seguindo com as configurações, escolhe-se o otimizador de aprendizagem. São disponibilizados o SMOReg e o SMORegImproved. Os dois algoritmos são testados. As previsões mensais, para o período de outubro de 2019 a dezembro de 2020 são mostradas na figura 28.

**Figura 28: Previsões de SVM dos dados de teste para o otimizador**

**Fonte: Autoria própria (2021)**

As previsões, da figura 28, mostram resultados próximos para os dois otimizadores. Para fevereiro de 2020, o algoritmo *SMOReg* teve um desempenho melhor. Já em abril de 2020, o *SMORegImproved* preveu um consumo mais próximo do real. Contudo, nenhum deles foi capaz de acompanhar o crescimento observado em junho de 2020.

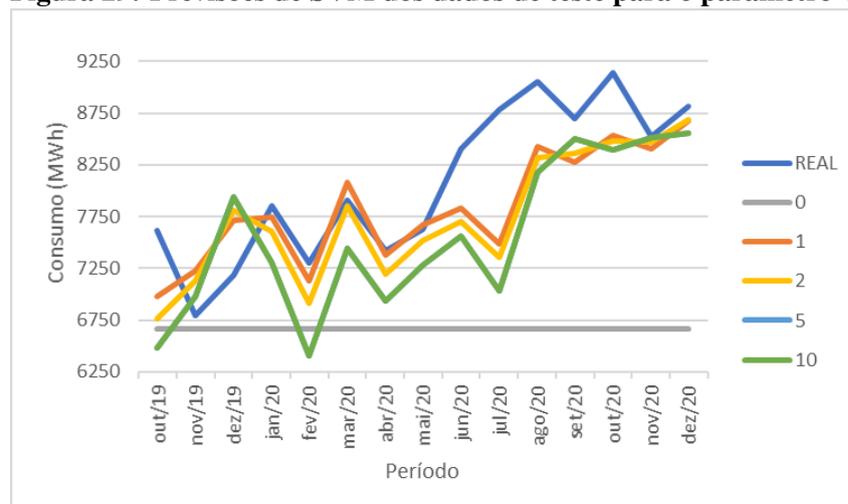
**Tabela 18: MAE e MAPE dos modelos de SVM para o otimizador**

Otimizador	MAE (MWh)	MAPE (%)
SMORegImproved	381,366	4,58 %
SMOReg	350,5883	4,62 %

**Fonte: Autoria própria (2021)**

De acordo com os erros, apresentados na tabela 18, o *SMOReg* apresentou um MAE menor. Entretanto, o MAPE do *SMORegImproved* foi inferior. Optou-se pelo otimizador *SMORegImproved*.

Por fim, o parâmetro C é configurado. São testados cinco valores: 0, 1, 2, 5 e 10. As previsões, para o período de outubro de 2019 a dezembro de 2020, para cada um dos casos, são mostradas na figura 29.

**Figura 29: Previsões de SVM dos dados de teste para o parâmetro C**

**Fonte: Autoria própria (2021)**

Pela figura 29, vê-se que, quando o parâmetro C foi zero, as previsões foram lineares. Assim, esse valor foi considerado inadequado.

Com exceção de dezembro de 2019, o parâmetro C, com valor 10, faz com que as previsões subestimem o valor de consumo. Uma resposta próxima a essa é encontrada quando o parâmetro C é configurado em 5.

As previsões, entre os meses de janeiro de 2020 e maio de 2020, acompanham bem a curva real, quando o parâmetro C foi configurado em 1 ou 2. A partir de maio de 2020, o consumo foi maior que o predito. Analisando o MAE e o MAPE, para os valores dos parâmetros,

na tabela 19, observa-se que os melhores resultados são obtidos para o parâmetro C igual a 1.

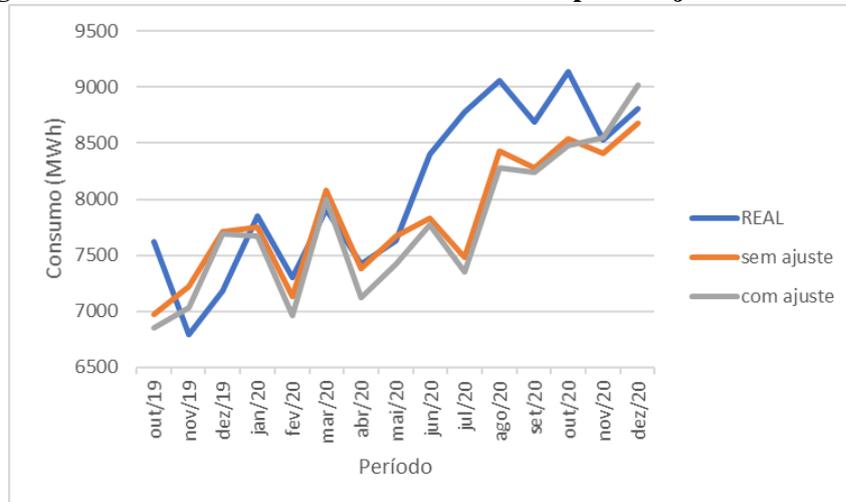
**Tabela 19: MAE e MAPE dos modelos de SVM para o parâmetro C**

Parâmetro C	MAE (MWh)	MAPE ( %)
0	1.370,068	16,20 %
1	381,366	4,58 %
2	440,469	5,33 %
5	632,771	7,74 %
10	632,771	7,74 %

**Fonte: Autoria própria (2021)**

Testou-se também a reação do modelo em relação ao ajuste da variância dos dados. As previsões, de outubro de 2019 a dezembro de 2020, com e sem esse ajuste, são mostradas na figura 30. O MAE e o MAPE para os dados de teste são apresentados na tabela 20. Desses dados, constata-se que o ajuste da variância impacta em uma piora no desempenho. Desse modo, preferiu-se seguir sem ele.

**Figura 30: Previsões de SVM dos dados de teste para o ajuste da variância**



**Fonte: Autoria própria (2021)**

**Tabela 20: MAE e MAPE dos modelos de SVM para o ajuste da variância**

Ajuste da variância	MAE (MWh)	MAPE ( %)
com ajuste	428,387	5,18 %
sem ajuste	381,366	4,58 %

**Fonte: Autoria própria (2021)**

Com isso, finalizam-se os itens a serem configurados. A tabela 21 mostra os parâmetros que foram encontrados para o melhor desempenho, nos dados de teste, do modelo de máquinas de vetores de suporte.

**Tabela 21: Configuração do modelo de máquina de vetores de suporte**

Parâmetro	Configuração
Função <i>kernel</i>	Polinomial
Otimizador	SMORegImproved
Parâmetro C	1
Ajuste da variância	sem ajuste

**Fonte: Autoria própria (2021)**

Nota-se, entretanto, que as mudanças nas configurações não trouxeram uma melhora de impacto na performance do modelo. De toda forma, o que se destaca, para as máquinas de suporte de vetores são a simplicidade, rapidez e eficiência com a qual os modelos são desenvolvidos. Através da configuração de somente três parâmetros, foi alcançado um MAPE de 4,58%.

De acordo com Ruas et al. (2008), os poucos ajustes facilitam a automatização do processo de previsão. Além disso, para Dosualdo e Rezende (2003), a principal vantagem desse método é o fato de trabalharem bem com um conjunto de dados de alta dimensão, uma vez que não existe o risco de encontrarem mínimos locais, um problema frequente quando se trabalha com redes neurais. Como desvantagem, novamente, há a incompreensão dos modelos ao ser humano, uma vez que não fica claro como as decisões são tomadas (DOSUALDO; REZENDE, 2003).

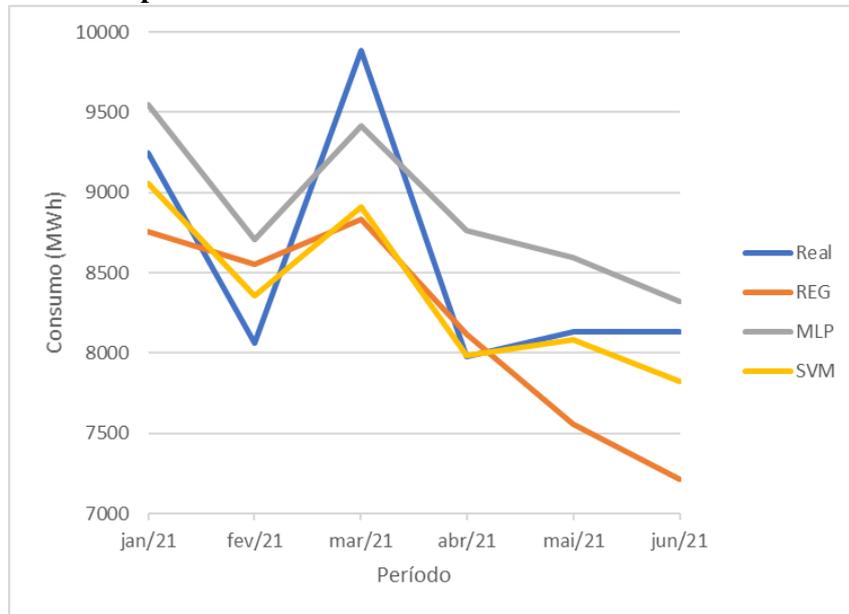
#### 4.3 COMPARAÇÃO ENTRE OS MODELOS

Os achados desta pesquisa avançam o conhecimento prévio na área. Este trabalho desenvolveu e validou modelos de regressão linear, redes neurais e máquinas de vetores de suporte, para a previsão do consumo de energia elétrica em uma planta frigorífica do oeste do Paraná. Contudo, para os dados de teste, a regressão linear apresentou um MAPE superior aos 5% propostos nesta pesquisa.

Com o intuito de verificar o desempenho dos modelos, foi previsto o consumo de energia elétrica de um período que não estava contido no banco de dados. Foram feitas as predições, de janeiro de 2021 a junho de 2021, usando a regressão linear, redes neurais e máquinas de vetores de suporte. Os resultados são apresentados na figura 31, onde também foram incluídas as

medições reais da unidade, de modo a avaliar a eficiência dos modelos.

**Figura 31: Comparação entre as previsões dos modelos de regressão linear, redes neurais e máquina de vetores de suporte com o consumo**



**Fonte: Autoria própria (2021)**

Observa-se, na figura 31 que, no primeiro trimestre, a variação do consumo foi mais intensa do que os modelos estimaram. Para o segundo trimestre, a regressão linear preveu uma queda, que não ocorreu. Já a rede neural, com exceção de março, superestimou o consumo. A máquina de vetores de suporte, apesar de ser mais suave que a curva real, apresentou resultados acurados, em especial em janeiro, abril e maio de 2021.

A tabela 22 mostra o MAPE das previsões, para os modelos, para os seis meses previstos, assim como a média do semestre. Para os dados não vistos, observa-se, da tabela 22, que o melhor modelo foi o de máquina de vetores de suporte, que teve um erro médio de 3,38%.

**Tabela 22: MAPE dos modelos de regressão linear, redes neurais e máquina de vetores de suporte**

Período	Regressão	Redes neurais	SVM
Janeiro/2021	5,27 %	7,85 %	2,03 %
Fevereiro/2021	6,10 %	2,10 %	3,71 %
Março/2021	10,64 %	14,24 %	9,90 %
Abril/2021	1,81 %	9,87 %	0,16 %
Mai/2021	7,12 %	5,62 %	0,64 %
Junho/2021	11,29 %	2,33 %	3,82 %
Média	7,04 %	5,65 %	3,38 %

**Fonte: Autoria própria (2021)**

Para os dados de validação, assim como para os de teste, a rede neural, com um MAPE de 5,65%, se saiu melhor que a regressão linear, que apresentou um MAPE de 7,04%. Apesar disso, pelos limites estabelecidos, nenhum dos dois foi considerado como adequado para a previsão do consumo de energia elétrica da unidade estudada.

Para avaliação, foram calculados também o MAE e RMSE para a regressão, redes neurais e máquina de vetores de suporte, que são mostrados nas tabelas 23 e 24. Nelas, os mesmos resultados encontrados pelo MAPE são confirmados. Para o conjunto de validação, as melhores estimativas obtidas foram usando máquinas de vetores de suporte, seguida das redes neurais e, por fim, da regressão linear.

**Tabela 23: MAE dos modelos de regressão linear, redes neurais e máquina de vetores de suporte**

Período	Regressão (MWh)	Redes neurais (MWh)	SVM (MWh)
Média	611,988	475,601	306,769

**Fonte: Autoria própria (2021)**

**Tabela 24: RMSE dos modelos de regressão linear, redes neurais e máquina de vetores de suporte**

Período	Regressão (MWh)	Redes neurais (MWh)	SVM (MWh)
Média	681,203	515,913	443,713

**Fonte: Autoria própria (2021)**

A superioridade do SVM, para a previsão do consumo de energia elétrica, é corroborada por Kaytez et al. (2015), Ogcü et al. (2012), Vasconcelos et al. (2017) e Usha e Balamurugan (2016), confirmando a eficácia da máquina de vetores de suporte para esse tipo de previsão.

Tanto SVM quanto as redes neurais são tolerantes a ruídos e podem ser usadas em problemas que requerem alta precisão. Apesar de, em termos técnicos, não se ter notado diferença entre os tempos de execução dos algoritmos, Haykin (2001) releva que as máquinas de vetor de suporte são mais lentas.

A comparação entre o desempenho das redes neurais e regressão linear, não é unânime na literatura. Enquanto Kaytez et al. (2015) e Usha e Balamurugan (2016) encontraram resultados melhores com as redes neurais, Nizami e Al-Garni (1995) e Azadeh et al. (2008), em conformidade com os achados dessa pesquisa, julgaram a rede neural mais adequada que a regressão linear para suas aplicações.

De acordo com Dosualdo e Rezende (2003), a regressão em mineração de dados preditiva é uma questão pouco explorada. A maioria das pesquisas são voltadas para os problemas de classificação e não de regressão.

De fato, no desenvolvimento de todos os algoritmos desta pesquisa, foram usados somente os dados de consumo mensais como entrada. Andrade (2009) expõe a fragilidade de modelos pautados na própria série histórica de consumo de energia elétrica. Apesar de apresentarem resultados interessantes, são mais susceptíveis a erros quando ocorrem variações bruscas na economia.

Kaytez et al. (2015) e Vasconcelos et al. (2017), desenvolveram modelos com entradas múltiplas. Enquanto Kaytez et al. (2015) usaram informações da capacidade de geração e consumo de energia elétrica, Vasconcelos et al. (2017) acrescentou além do histórico de consumo, o Produto Interno Bruto e população. Contudo, para Vasconcelos et al. (2017), a adição de variáveis não impactou em melhoras no desempenho de seus modelos.

O banco de dados deste trabalho, devido a disponibilidade, é limitado. Enquanto usou-se aqui os dados de 2016 e 2021, os modelos de Vasconcelos et al. (2017) e Kaytez et al. (2015) usaram históricos maiores que 30 anos.

Conforme Kaytez et al. (2015), os parâmetros que afetam o consumo de energia elétrica devem ser bem definidos para que seja possível identificar e adicionar os indicadores que tem influência no consumo. Cabe ressaltar que a unidade consumidora em questão, além de não ter um histórico de dados extenso, ainda não possui um controle das variáveis de influência no processo. Desse modo, é complexo estimar possíveis variáveis de impacto.

Assim, julgou-se que não cabe a comparação dos erros obtidos com relação a outros trabalhos. Tem-se em conta que cada banco de dados é único, seja por seu tamanho ou variáveis incluídas.

Apesar de todas os algoritmos usados terem parâmetros a serem configurados pelo usuário, destaca-se a complexidade das redes neurais. O número de variáveis implica em quase infinitas possibilidades. De fato, Faceli et al. (2016) recomenda que seja executada várias vezes para diferentes configurações de dados e valores iniciais, obtendo uma média de desempenho. Enquanto as máquinas de vetores de suporte, um método determinístico, não tem resultados dependentes da ordem de apresentação, as redes neurais, estocásticas, a hierarquia e pesos iniciais tem grande influência (FACELI et al., 2016).

A modelagem das redes neurais foi complexa. A alteração das variáveis não tem um impacto claro. Essa também é uma queixa de Dosualdo e Rezende (2003) e Faceli et al. (2016). Para eles, as soluções não são facilmente interpretadas pelos usuários, pois o conhecimento está embutido nos pesos e conexões da rede. Ruas et al. (2008) expõe como desvantagem, da rede neural, a necessidade de um grande número de testes e experiência por parte do projetista.

Campos (2008) menciona que a divergência da rede neural aumenta com a ampliação da previsão. O autor conclui que as redes neurais são eficientes para previsões de médio prazo, que implica em um prazo de até um ano à frente. Esse foi o horizonte usado por Vasconcelos et al. (2017) em seu trabalho, alcançando bons resultados. Todavia, essa não é uma regra estabelecida. Usha e Balamurugan (2016), por exemplo, fez previsões 3 anos a frente e indicou que esses dados são úteis para tomada de decisões para plantas de geração de energia.

Ainda nas questões metodológicas, cabe citar novamente a escolha do *software* usado para o desenvolvimento. Kaytez et al. (2015), por exemplo, usou o *software* matemático Matlab. Apesar de ser mais complexo para operar, fornece uma série de outros parâmetros e possibilidades de configurações, que podem trazer benefícios aos projetista.

Neste trabalho, optou-se pelo WEKA, também usado por Vasconcelos et al. (2017) e Usha e Balamurugan (2016). Sua interface gráfica é simples de operar e auxilia na execução de testes com diferentes algoritmos e configurações. Além disso, esse é um *software* de livre acesso e gratuito. Existem, de acordo com Witten et al. (2016), limitações relacionadas ao tamanho do banco de dados que pode ser usados. Também cita-se a impossibilidade da aplicação de algoritmos concomitantes. Porém, essas restrições não tiveram impacto nessa pesquisa, devido ao conjunto de dados reduzido e os algoritmos utilizados.

Da mesma forma que para Castoril et al. (2018), os achados dessa pesquisa trazem luz sobre a possibilidade da criação de *softwares* que sejam acessíveis a um grande número de unidades, que tem um consumo significativo e precisam estimar seus valores futuros, para otimizar investimentos e ações.

Entre os modelos desenvolvidos e usados neste trabalho, a SVM foi a única a ter previsões com MAPE inferior a 5%, para o conjunto de validação. De fato, a máquina de vetores de suporte obteve o melhor resultado. Contudo, visto que os erros obtidos, para a regressão linear e as redes neurais, ainda são pequenos, considerou-se que ambos também estão aptos a serem utilizados na previsão semestral, do consumo de energia elétrica, do frigorífico em estudo.

## 5 CONCLUSÕES

Esta pesquisa teve como objetivo o desenvolvimento e validação de modelos de previsão utilizando algoritmos de aprendizagem de máquina, em especial, a regressão linear, as redes neurais e as máquinas de vetores de suporte, para consumidores industriais de médio porte.

A estatística permite a organização, resumo, análise e interpretação dos dados, auxiliando na crítica e em argumentos para tomadas de decisões. Apesar da identificação do comportamento do conjunto de dados de forma mais clara e objetiva, essa análise não se mostrou suficiente para atender as necessidades de estimativa do consumo futuro. Assim, propõe-se o uso de aprendizagem de máquinas para fazer as previsões.

Para a regressão linear, foram configurados o parâmetro de *Ridge*, em  $1 \times 10^{-3}$  e o método de seleção M5. Nos dados de teste, obteve-se um MAPE de 6,78%.

As redes neurais são mais complexas para projetar, uma vez que uma série de hiperparâmetros pode ser alterados. Neste trabalho, foram usadas 10 neurônios na camada oculta, taxa de aprendizagem 0,1, *momentum* 0,3 e tempo de treinamento de 5000 épocas. Dentro os modelos desenvolvidos, somente a rede neural obteve um resultado melhor ao ajustar os dados de entrada de acordo com a variância. Após todos os ajustes, o MAPE, para o teste, foi de 4,87%

Em oposição as redes neurais, as máquinas de vetores de suporte são simples de configurar e apresentam boa performance. Para os dados de teste, obteve-se um MAPE de 4,58%. Esse resultado foi atingido utilizando a função *kernel* polinomial, o otimizador *SMORegImproved* e o parâmetro C igual a 1.

Observa-se que, somente a regressão linear não atingiu o critério mínimo, de MAPE inferior a 5%, para os dados de teste. Desse modo, considerando o nível estabelecido, este modelo não é considerado adequado para as previsões de consumo de energia para a unidade em estudo.

Após o ajuste dos modelos, utilizando os parâmetros encontrados para cada algoritmo de aprendizagem de máquina, foi realizada a validação. As previsão do consumo de energia elétrica,

de janeiro de 2021 a junho de 2021, foram comparadas com o consumo real. O modelo com melhor desempenho foi o SVM, com um MAPE de 3,38%. Nos dados de validação, novamente, a rede neural atingiu resultados melhores que a regressão. Para a previsão do semestre, a MLP teve um erro de 5,65%, frente a 7,04% da regressão linear.

Como contribuição teórica, os modelos desenvolvidos auxiliaram na construção do conhecimento. As previsões obtidas facilitam ao usuário estimar seu consumo futuro, ajudando-o no planejamento de suas atividades e demandas.

Podem ser implementadas melhorias nos modelos, como a adição de variáveis. Porém, as estimativas de consumo foram obtidas com base em informações que a unidade já possui atualmente, o histórico de dados de 2016 a 2020, e com um baixo custo, uma vez que *software* é gratuito.

Os trabalhos existentes, de forma geral, analisam o consumo do ponto de vista residencial, da distribuidora ou para grandes regiões. Identifica-se a necessidade de modelos para a previsão do consumidor, em especial o industrial. Tendo em vista o crescimento do mercado livre de energia, torna-se fundamental o acesso a ferramentas que permitam que o consumidor tenha informações sobre o seu consumo futuro. Os resultados, obtidos neste trabalho, também podem ser úteis para outros consumidores, com cargas com características semelhantes.

Os algoritmos de aprendizagem são ferramentas poderosas que podem e devem ser aplicadas visando facilitar atividades e funções. Vislumbra-se que, através dos algoritmos de aprendizagem, podem ser criados *softwares* acessíveis que auxiliem nas atividades e funções relacionadas a energia elétrica.

Para ampliar os conhecimentos obtidos com essa pesquisa, sugere-se, como trabalhos futuros:

- Utilizar os parâmetros encontrados para cada algoritmo de aprendizagem de máquina em outros *softwares*, como o *Matlab*, e comparar os resultados;
- Adicionar variáveis externas ao histórico do consumo de energia elétrica, em especial, relacionados a meteorologia ou economia;
- Melhorar as performances dos modelos a partir do uso de algoritmos híbridos. Sobretudo, com o uso de técnicas de inteligência artificial, que podem otimizar as previsões;

## REFERÊNCIAS

- ABREU, A. L. M. de. **Bootstrap e modelos de support vector machine - SVM**. Monografia (Tese de doutorado) — Universidade Federal do Paraná, 2016.
- AMARAL, H. L. M. **Desenvolvimento de uma nova metodologia para previsão de consumo de energia elétrica de curto prazo utilizando redes neurais artificiais e decomposição de séries temporais**. Monografia (Tese de doutorado) — Universidade de São Paulo, 2020.
- ANDRADE, F. F. **Um modelo para descrição e previsão da demanda dos consumidores industriais de energia elétrica**. Monografia (Tese de doutorado) — Universidade Federal de Santa Catarina, 2009.
- AZADEH, A.; GHADERI, S. F.; SOHRABKHANI, S. A simulated-based neural network algorithm for forecasting electrical energy consumption in iran. **Energy Policy**, v. 36, p. 2637–2644, 2008.
- BARROS, B. F.; BORELLI, R.; GEDRA, R. L. **Gerenciamento de energia: ações administrativas e técnicas de uso adequado da energia elétrica**. 3. ed. São Paulo: Editora Érica, 2020. ISBN 978-85-365-3306-3.
- BONESSO, D. **Estimação dos parâmetros de Kernel em um classificador SVM na classificação de imagens hiperespectrais em uma abordagem multiclasse**. Monografia (Dissertação de mestrado) — Universidade Federal do Rio Grande do Sul, 2013.
- BUENO, L. G. **Diagnóstico do uso de energia elétrica de um frigorífico de frangos de corte enfatizando medidas de eficiência energética**. Monografia (Tese de doutorado) — Universidade Estadual de Campinas, 2008.
- CAMPOS, R. J. **Previsão de séries temporais com aplicações a séries de consumo de energia elétrica**. Monografia (Dissertação de mestrado) — Universidade Federal de Minas Gerais, 2008.
- CARDOSO, G. B. **Potencial de redução de consumo de energia elétrica em entreposto frigorífico: um estudo de caso**. Monografia (Dissertação de mestrado) — Universidade Estadual Paulista, 2004.
- CARVALHO, D. J. **Métodos de previsão de consumo de energia elétrica residencial em grande volume de dados**. Monografia (Trabalho de conclusão de curso) — Universidade Federal Rural de Pernambuco, 2019.
- CASTORIL, B. R. et al. Previsão de demanda de energia elétrica no setor de produção do laticínio do ifmg - campus bambuí utilizando regressão de vetores de suport. **Revista de Informática Aplicada**, v. 14, n. 2, p. 62–77, 2018.
- CASTRO, L. N.; FERRARI, D. G. **Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações**. São Paulo: Editora Saraiva, 2016.

- DANTAS FILHO, L. L. N. **Análise de modelos para previsões de consumo de energia elétrica e seu uso na redução de custos**. Monografia (Trabalho de conclusão de curso) — Universidade Federal da Paraíba, 2019.
- DOSUALDO, D. G.; REZENDE, S. O. **Análise da precisão de métodos de regressão**. Monografia (Relatório técnico) — Universidade de São Paulo, 2003.
- FACELI, K. et al. **Inteligência artificial: uma abordagem de aprendizado de máquina**. 1. ed. São Paulo: Editora LTC, 2016. ISBN 978-85-216-1880-5.
- FOGLIATTO, F. S. et al. Previsão de demanda por energia elétrica - método e aplicação. **XXV Encontro Nacional de Engenharia de Produção**, 2005.
- FRANK, E.; HALL, M. A.; WITTEN, I. H. **The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques**. [S.l.]: Morgan Kaufmann, 2016.
- FROZZA, J. F. **Eficiência energética em uma indústria frigorífica: desafios da implantação**. Monografia (Dissertação de mestrado) — Universidade Tecnológica Federal do Paraná, 2013.
- GERHARDT, T. E.; SILVEIRA, D. T. **Métodos de pesquisa**. 5. ed. Porto Alegre: Editora da UFRGS, 2009.
- HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Editora Bookman, 2001. ISBN 978-85-7780-086-5.
- HERNANDEZ NETO, A.; FIORELLI, F. A. S. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. **Energy and buildings**, v. 40, p. 2169–2175, 2008.
- KAYTEZ, F. et al. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. **International Journal of Electrical Power Energy Systems**, v. 67, p. 431–438, 2015.
- MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Editora Atlas, 2003.
- MORDJAOUI, M. et al. Electric load forecasting by using dynamic neural network. **International Journal of Hydrogen Energy**, v. 42, n. 28, p. 17655–17663, 2017.
- NIZAMI, S. S. A. K. J.; AL-GARNI, A. Z. Forecasting electric energy consumption using neural networks. **Energy Policy**, v. 23, n. 12, p. 1097–1104, 1995.
- OGCU, G.; DEMIREL, O. F.; ZAIM, S. Forecasting electricity consumption with neural networks and support vector regression. **Procedia - Social and Behavioral Sciences**, v. 58, p. 1576–1585, 2012.
- PINTO, M. O. **Energia elétrica: geração, transmissão e sistemas interligados**. 1. ed. Rio de Janeiro: Editora LTC, 2018. ISBN 978-85-216-2525-4.
- RUAS, G. I. S. et al. Previsão de demanda de energia elétrica utilizando redes neurais artificiais e support vector regression. **3rd Internacional Symposium on Communications, Control and Signal Processing**, 2008.

SANTOS, J. A. A. dos; CHAUCOSKI, Y. Previsão do consumo de energia elétrica na região sudeste: Um estudo de caso usando sarima e lstm. **Revista Cereus**, v. 12, n. 4, 2020.

SCHÖLKOPF, B.; SMOLA, A. J. **Learning with kernel: support vector machines, regularization, optimization, and beyond**. Cambridge, Massachusetts: The MIT Press, 2002.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas: curso prático**. 5. ed. São Paulo: Editora Artliber, 2010.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, n. 14, p. 199–222, 2004.

USHA, T. M.; BALAMURUGAN, S. A. A. Seasonal based electricity demand forecasting using time series analysis. **Circuits and Systems**, v. 7, n. 10, p. 3320–3328, 2016.

VASCONCELOS, B. F. B. **Poder produtivo de métodos de Machine Learning com processos de seleção de variáveis: uma aplicação às projeções de produto de países**. Monografia (Tese de doutorado) — Universidade de Brasília, 2017.

VASCONCELOS, L. P. et al. Uma comparação de técnicas de regressão para a previsão de consumo de energia residencial no cenário nacional. **XLIX Simpósio Brasileiro de Pesquisa Operacional**, v. 13, 2017.

WITTEN, H. I. et al. **Data mining: practical machine learning tools and techniques**. 3. ed. [S.l.]: Morgan Kaufmann, 2016.

ZANIN, A. et al. Perfil do consumo de energia elétrica no abate de frangos de corte: um estudo de caso. **IV Encontro de energia no meio rural**, 2002.

ZENG, Y. R. et al. Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network. **Energy**, v. 127, p. 381–396, 2017.