



GIOVANNA GOUVÊA SPURI DE MIRANDA

**ESTUDO COMPARATIVO DE TÉCNICAS
COMPUTACIONAIS PARA PREVISÃO DE
DEMANDA DE ENERGIA**

LAVRAS – MG

2023

GIOVANNA GOUVÊA SPURI DE MIRANDA

**ESTUDO COMPARATIVO DE TÉCNICAS COMPUTACIONAIS PARA
PREVISÃO DE DEMANDA DE ENERGIA**

Projeto de pesquisa da disciplina Trabalho de Conclusão de Curso apresentada ao Departamento Automática da Universidade Federal de Lavras como parte das exigências do curso de Engenharia de Controle e Automação, para a obtenção do título de Bacharel.

Prof. Dr. Wilian Soares Lacerda

Orientador

LAVRAS – MG

2023

**Ficha catalográfica elaborada pela Coordenadoria de Processos Técnicos
da Biblioteca Universitária da UFLA**

Miranda, Giovanna Gouvêa Spuri

Estudo comparativo de técnicas computacionais para previsão de
demanda de energia / Giovanna Gouvêa Spuri de Miranda. 1^a ed.
rev., atual. e ampl. – Lavras : UFLA, 2023.

106 p. : il.

Orientador: Prof. Dr. Wilian Soares Lacerda.

Bibliografia.

1. TCC. 2. Monografia. 3. Dissertação. 4. Tese. 5. Trabalho
Científico – Normas. I. Universidade Federal de Lavras. II. Título.

CDD-808.066

GIOVANNA GOUVÊA SPURI DE MIRANDA

**ESTUDO COMPARATIVO DE TÉCNICAS COMPUTACIONAIS PARA
PREVISÃO DE DEMANDA DE ENERGIA**

Projeto de pesquisa da disciplina Trabalho de Conclusão de Curso apresentada ao Departamento Automática da Universidade Federal de Lavras como parte das exigências do curso de Engenharia de Controle e Automação, para a obtenção do título de Bacharel.

08 de Dezembro de 2023.

Prof. Dr. Danton Diego Ferreira UFLA
Prof. Dr. Fabio Domingues De Jesus UFLA

Prof. Dr. Wilian Soares Lacerda
Orientador

**LAVRAS – MG
2023**

*À minha avó Dila Thereza Spuri de Miranda
À minha tia Delza Neli Espuri
À Ariany Carolina de Oliveira
Aos meus pais*

AGRADECIMENTOS

À Ariany Carolina de Oliveira por ser uma fonte constante de inspiração, motivação e encorajamento. Sem a sua presença, possivelmente não teria concluído o curso técnico e muito menos começado a graduação. Agradeço imensamente por tudo que fez e faz na minha vida e por toda a influência positiva que proporcionou em minha jornada.

À minha avó Dila, à tia Delza, a minha irmã Gabriela e aos meus pais pelo apoio incondicional e constante cuidado.

Ao Prof. Wilian Soares Lacerda, pelas horas dedicadas e pela disponibilidade em orientar-me ao longo desse trabalho.

À ENACOM pela oportunidade de estágio e pela parceria estabelecida durante a realização deste trabalho.

Ao Endi Daniel pela paciência e por toda ajuda prestada.

À banca examinadora, pelo tempo e dedicação na avaliação do trabalho.

À Universidade Federal de Lavras por propiciar as condições necessárias para a realização deste trabalho.

A todos os professores que contribuíram com seus ensinamentos ao longo da minha formação acadêmica.

RESUMO

Este trabalho constitui um projeto de pesquisa que visa dar início a uma pós graduação e tem como objetivo a análise e comparação das técnicas como regressão, modelos autorregressivos e redes neurais artificiais na previsão da demanda de energia dos consumidores livres de Minas Gerais com dados de consumo de 2014 a 2022 obtidos no site da CCEE (Câmara de Comercialização de Energia Elétrica). A avaliação dos modelos é fundamentada em métricas de desempenho, tais como MAE, MSE, MAPE e R^2 ajustado .

Palavras-chave: Previsão de Demanda; ARIMA; Redes Neurais Artificiais.

ABSTRACT

This work constitutes a research project that aims to start a postgraduate course and aims to analyze and compare techniques such as regression, autoregressive models and artificial neural networks in predicting the energy demand of free consumers in Minas Gerais with data from consumption from 2014 to 2022 obtained from the CCEE (Electric Energy Trading Chamber) website. Model evaluation is based on performance metrics, such as MAE, MSE, MAPE and adjusted R^2

Keywords: Demand Forecast; ARIMA; Artificial neural networks.

LISTA DE FIGURAS

Figura 2.1 – Etapas da previsão	24
Figura 2.2 – Representação em diagrama em blocos do sistema nervoso . .	37
Figura 2.3 – A célula piramidal	38
Figura 2.4 – Modelo não-linear de um neurônio	39
Figura 2.5 – Um outro modelo não-linear de um neurônio	41
Figura 2.6 – Camadas de uma rede	47
Figura 2.7 – Arquitetura <i>feedforward</i> de uma única camada (<i>Perceptron</i>) .	49
Figura 2.8 – Arquitetura <i>feedforward</i> de múltiplas camadas	50
Figura 2.9 – Arquitetura da Rede Neural Recorrente	51
Figura 2.10 – Sistema de aprendizado supervisionado	52
Figura 2.11 – Sistema de aprendizado não-supervisionado	53
Figura 2.12 – Sistema de aprendizado recorrente	53
Figura 2.13 – Quartis	57
Figura 2.14 – Decis	57
Figura 2.15 – Boxplot	61

LISTA DE GRÁFICOS

Gráfico 2.1 – Tipos de tendência	28
Gráfico 2.2 – Componentes de uma série temporal	29
Gráfico 2.3 – Função Degrau	42
Gráfico 2.4 – Função Linear por partes	43
Gráfico 2.5 – Função Sigmóide	44
Gráfico 2.6 – Função Tangente Hiperbólica	45
Gráfico 2.7 – Função ReLU	46
Gráfico 4.1 – Boxplot do consumo energético (GWh) de 2014	83
Gráfico 4.2 – Boxplot do consumo energético (GWh) de 2022	83
Gráfico 4.3 – Comparativo do consumo energético médio (GWh) de cada ramo de atividade no ano de 2014	84
Gráfico 4.4 – Comparativo do consumo energético médio (GWh) de cada ramo de atividade no ano de 2022	85
Gráfico 4.5 – Comparativo do consumo energético total (GWh) de cada ramo de atividade no ano de 2014	86
Gráfico 4.6 – Comparativo do consumo energético total (GWh) de cada ramo de atividade no ano de 2022	86
Gráfico 4.7 – Boxplot do consumo energético (GWh) do ramo de atividade: ‘Metalurgia e Produtos de Metal’ entre 2014 e 2022	88
Gráfico 4.8 – Boxplot do consumo energético (GWh) dos consumidores fi- xos entre 2014 e 2022 do ramo de atividade: ‘Metalurgia e Produtos de Metal’	88
Gráfico 4.9 – Decomposição do consumo total (GWh) do ramo de ativi- dade: ‘Metalurgia e Produtos de Metal’	89
Gráfico 4.10 – Comparação da tendência real do consumo total (GWh) com da equação que representa essa tendência	90
Gráfico 4.11 – Pontos de inflexão da tendência do consumo total (GWh)	90

Gráfico 4.12 – Comparação da sazonalidade real do consumo total (GWh) com da equação que representa essa sazonalidade	91
Gráfico 4.13 – Matriz de correlações entre consumo, precipitação, temperatura e preço	92
Gráfico 4.14 – Correlação entre o consumo total (GWh) do ramo de atividade ‘Metalurgia e Produtos de Metal’ e a temperatura média no estado de MG	93
Gráfico 4.15 – Correlação entre o consumo total (GWh) do ramo de atividade ‘Metalurgia e Produtos de Metal’ e a precipitação média no estado de MG	94
Gráfico 4.16 – Correlação entre o consumo total (GWh) do ramo de atividade ‘Metalurgia e Produtos de Metal’ e o preço médio do submercado Sudeste	94

LISTA DE TABELAS

Tabela 4.1 – Consumo médio dos consumidores livres por ramo de atividade (2014-2022)	84
Tabela 4.2 – Consumo total dos consumidores livres por ramo de atividade (2014-2022)	85
Tabela 4.3 – Porcentagem do consumo total de cada ramo de atividade (2014-2022)	87
Tabela 5.1 – Orçamento estimado para pesquisa em métodos de previsão. .	96
Tabela 6.1 – Cronograma completo	97

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos	16
1.2	Justificativa do trabalho	16
1.3	Organização do trabalho	17
2	REFERENCIAL TEÓRICO	19
2.1	Sistema elétrico brasileiro	19
2.2	Consumidores livres	21
2.3	Previsão de demanda	23
2.4	Séries Temporais	25
2.5	Regressão	30
2.6	Modelo autorregressivo (AR) de média móvel (MA)	32
2.7	Redes Neurais Artificiais	35
2.8	Estatística Descritiva	54
2.9	Avaliação de desempenho e métricas de previsão de demanda de energia	62
2.10	Trabalhos correlacionados	64
3	METODOLOGIA	67
3.1	Conjunto de dados	67
3.2	Recursos Computacionais	68
3.3	Análise dos dados	70
3.4	Modelagem e Previsão	71
3.5	Comparação Estatística dos Métodos	80
4	RESULTADOS PRELIMINARES	83
4.1	Metalurgia e Produtos de Metal	87
5	RECURSOS NECESSÁRIOS	95
5.1	Hardware	95
5.2	Software	95

5.3	Outros Recursos	96
5.4	Orçamento	96
6	CRONOGRAMA	97
7	EQUIPE	99
8	CONCLUSÃO	101
	REFERÊNCIAS	102
	APENDICE A – TABELA DE DADOS DO CONSUMO	
	POR PARCELA DE CARGA - MWh	106

1 INTRODUÇÃO

A geração, transmissão e distribuição de eletricidade são processos complexos que exigem coordenação e planejamento meticulosos. Para fazer esse planejamento existe o Operador Nacional do Sistema Elétrico (ONS) que coordena o Sistema Interligado Nacional (SIN). O SIN é formado por extensas linhas de transmissão de alta tensão e corresponde a 98% de toda energia elétrica consumida no Brasil. Com o sistema interligado caso haja um problema envolvendo uma usina geradora que impeça a injeção de energia elétrica na rede é possível utilizar outras usinas para atender a carga demandada, evitando que ocorra a interrupção no abastecimento de energia elétrica (BARROS; GEDRA; BORELLI, 2014).

Embora tenha inúmeras vantagens, o SIN também apresenta desvantagens que merecem atenção. Entre essas desvantagens, inclui-se o risco de apagões devido a falhas na geração e transmissão de energia, condições climáticas adversas, falta de manutenção, aumento rápido da demanda e desastres naturais. Para mitigar esses riscos e garantir um fornecimento de energia confiável, torna-se essencial a realização de investimentos na modernização da infraestrutura, a diversificação da matriz energética e a implementação de um planejamento adequado, incluindo sistemas de previsão de demanda.

A demanda de energia refere-se à quantidade de energia necessária para atender às necessidades dos consumidores em um determinado período de tempo (COSSI, 2008). Ela está diretamente ligada a fatores como o crescimento populacional, o desenvolvimento industrial, os avanços tecnológicos, os padrões de consumo e os hábitos de vida das pessoas (BEZERRA, 2016). Entretanto, prever a demanda energética não se limita apenas ao seu volume, mas também em como diversificar a matriz energética de forma a obter a melhor eficiência energética considerando os aspectos ambientais (BEZERRA, 2016).

1.1 Objetivos

O objetivo deste estudo é analisar e comparar métodos tradicionais e avançados de previsão de demanda de energia dos consumidores livres de Minas Gerais.

1.1.1 Objetivos específicos

- Prever o consumo por regressão
- Prever o consumo utilizando o método estatístico *autoregressive integrated moving average* (ARIMA)
- Prever o consumo com Redes Neurais Artificiais (RNA)
- Avaliar os modelos de previsão utilizando o Erro Médio Absoluto (MAE), o Erro Médio Quadrático (MSE), o Erro Percentual Absoluto Médio (MAPE) e o coeficiente de determinação ajustado R^2
- Identificar a melhor técnica para prever a demanda

1.2 Justificativa do trabalho

Conforme Martinho (2009), se a demanda exceder a capacidade de geração e distribuição de energia disponível, pode haver sobrecarga no sistema elétrico, levando a apagões ou interrupções no fornecimento de energia, como aconteceu no dia 15 de agosto de 2023 em que houve a interrupção de 16mil MW de carga em todo país afetando pelo menos 29 milhões de pessoas (GLOBO, 2023).

Para suprir a necessidade adicional, pode ser inevitável recorrer a fontes de energia emergenciais, como usinas termelétricas, que geralmente são mais caras do que fontes renováveis e isso pode resultar no aumento dos custos de geração de energia e impacto ambiental (BEZERRA, 2016). Em casos extremos, quando a demanda excede significativamente a capacidade de geração e não há alternativas

viáveis disponíveis, pode haver racionamento, isto é, a redução programada do fornecimento de energia para diferentes regiões ou setores, a fim de equilibrar a oferta e a demanda (SCHWEITZER, 2003).

Em 2001, o Brasil apresentou deficit entre geração e consumo de energia elétrica resultando no maior racionamento de energia elétrica da história do país em termos de abrangência e redução de consumo, tendo duração de junho de 2001 a fevereiro de 2002 (BARDELIN, 2004).

Com base no que foi apresentado fica claro que a previsão de demanda energética é crucial para garantir a estabilidade e eficiência dos sistemas de energia. Ela permite que os fornecedores se preparem adequadamente, evitando desperdícios, interrupções no fornecimento e o aumento no preço devido ao uso de fontes emergenciais.

1.3 Organização do trabalho

Este trabalho está organizado da seguinte forma:

Capítulo 1: Introdução. É comentado sobre o sistema elétrico brasileiro e é discutido a importância da previsão de demanda de energia para a gestão eficiente dos recursos energéticos.

Capítulo 2: Referencial Teórico. É apresentado o referencial teórico necessário para compreender e desenvolver o trabalho. Nele é explorado conceitos fundamentais relacionados ao sistema elétrico brasileiro e a previsão de demanda de energia, como modelos autorregressivos, de regressão e redes neurais artificiais. Além disso, é abordado as métricas de desempenho, como o Erro Médio Absoluto, o Erro Médio Quadrático, o Erro Percentual Absoluto Médio e Coeficiente de Determinação R^2 .

Capítulo 3: Materiais e Metodologia. É exposto os materiais utilizados e a metodologia adotada para desenvolver o sistema de predição de demanda de

energia. Também é feito o detalhamento dos conjuntos de dados utilizados, os critérios de seleção e o processo de implementação das técnicas de previsão.

Capítulo 4: Resultados Preliminares. É apresentado as descobertas obtidas durante a análise dos dados.

Capítulo 5: Recursos Necessários. É detalhado os recursos necessários para conduzir o estudo, incluindo uma descrição dos custos associados aos equipamentos, software e aquisição de dados.

Capítulo 6: Cronograma. É planejado as etapas-chave do projeto e os prazos associados a cada uma delas. O cronograma fornece uma visão geral mais clara do planejamento temporal do estudo, auxiliando na gestão do tempo e no acompanhamento do progresso.

Capítulo 7: Equipe. É apresentado os membros da equipe e suas qualificações.

Capítulo 8: Conclusões. Contém o resumo das principais descobertas e descreve os próximos passos a serem realizados durante o programa de pós-graduação, com o objetivo de alcançar todos os objetivos propostos.

Em sequencia são apresentadas as Referências Bibliográficas e os Apêndices (A, B, etc.) onde são fornecidos os códigos utilizados.

2 REFERENCIAL TEÓRICO

Neste capítulo contém a análise e compreensão da previsão de demanda energética. Sendo assim, há uma breve explicação do sistema elétrico brasileiro, o estudo de modelos de previsão de séries temporais e métricas estatísticas para avaliar o desempenho das diferentes formas de prever a demanda energética.

2.1 Sistema elétrico brasileiro

A partir da década de 1980, a eletricidade passou a ser percebida como um produto que poderia ser negociado e comercializado livremente. Para transformá-la em uma mercadoria atrativa, os setores do mercado de energia elétrica foram segmentados em geração, transmissão, distribuição e comercialização (ABREU, 1999).

A geração refere-se o processo que produz eletricidade a partir da transformação de outras formas de energia, como o calor ou a queda-d'água, em energia elétrica. (OLIVEIRA et al., 2021).

A geração de energia elétrica ocorre normalmente afastada dos centros consumidores. Para que a eletricidade gerada alcance os consumidores, faz-se necessário um complexo sistema de transmissão e distribuição. A transmissão é caracterizada pelas extensas redes de linhas (cabos) que interliga as usinas às regiões consumidoras (BARROS; GEDRA; BORELLI, 2014).

Após percorrer a extensão geográfica existente entre a geração e os centros consumidores por meio das linhas de transmissão, a energia elétrica chega às Estações Transformadoras de Distribuição (ETD). As ETD's são compostas por fios condutores, transformadores e equipamentos diversos de fixação, controle e proteção dos circuitos que rebaixam os níveis de tensão utilizados na transmissão, permitindo, assim, que seja iniciada a etapa de distribuição (BARROS; GEDRA; BORELLI, 2014).

Para coordenar e controlar a operação das instalações de geração e transmissão de energia elétrica existe o Operador Nacional do Sistema Elétrico (ONS). Ele controla o Sistema Interligado Nacional (SIN) que é uma grandiosa infraestrutura que conecta diversas regiões do Brasil, possibilitando que a energia produzida em uma parte do país possa ser consumida em outra. Ele integra usinas geradoras de energia, linhas de transmissão e subestações, formando uma complexa rede que abrange quase todo o território nacional. Este sistema é dividido em quatro subsistemas: Sul, Sudeste/Centro-Oeste, Nordeste e Norte. Essa interligação permite que a energia gerada em regiões com excedente possa ser enviada para regiões com demanda maior (FERREIRA, 2006).

Ao mesmo tempo em que a interligação permite que seja compensada uma falha, também pode gerar um efeito prejudicial ao sistema na situação em que um desligamento não consiga ser compensado, sobrecarregando outras linhas e outras usinas, que por sua vez também chegam ao seu limite, provocando também seu desligamento. Assim, em um efeito dominó de sobrecarga e desligamentos várias linhas são desligadas, causando o desabastecimento em uma grande área do país (BARROS; GEDRA; BORELLI, 2014).

Segundo o Globo (2023) no Brasil já teve 10 apagões significativos. Em 1999 cerca de 60% do território nacional ficou no escuro após uma descarga elétrica atingir uma subestação de uma companhia energética importante no estado de São Paulo. Em 2001 uma falha humana na subestação de Cachoeira Paulista (SP) provocou um blecaute que atingiu o Rio de Janeiro, Espírito Santo e parte de Minas Gerais. Em 2009 aconteceu um dos maiores apagões do século, a quantidade de energia desligada com o apagão foi de 28 GW no Brasil, afetando 18 estados, e de 980 MW no Paraguai, quase 60 milhões de pessoas ficaram sem energia elétrica por até seis horas, devido à interrupção de três linhas de alta tensão provenientes de Itaipu. Em 2011 47,7 milhões de pessoas foram afetadas, cerca de 90% do nordeste ficou sem energia elétrica por causa dos efeitos de uma falha na subes-

tação localizada no município de Jatobá na divisa entre Bahia e Pernambuco. Em 2012 houve três apagões, 3 de outubro teve uma falha no transformador em Itaipu que afetou cinco Estados (Paraná, Rio de Janeiro, Minas Gerais, Acre, Rondônia) e parte do Centro-Oeste. Em 25 de outubro o apagão cortou a energia em nove Estados do Nordeste e parte da Região Norte por quase três horas e em dezembro de 2012, um problema na hidrelétrica de Itumbiara, atingiu seis Estados. Em 2014 o Sul, Sudeste, Norte e o Centro-Oeste ficaram sem energia por quase duas horas. Em 2018 cerca 70 milhões de pessoas ficaram sem energia elétrica. E no dia 15 de agosto de 2023, em decorrência da abertura da interligação Norte-Sudeste, houve interrupção 16 mil MW de carga em todo o país afetando pelo menos 29 milhões de pessoas.

Dentro deste contexto, a previsão de demanda elétrica torna-se fundamental para garantir a estabilidade e confiabilidade do sistema elétrico. A previsão de demanda permite que os operadores do sistema antecipem possíveis picos de consumo e, assim, planejem a distribuição de energia de forma mais eficiente. Ao prever com precisão quando e onde a demanda será maior, é possível alocar recursos de geração e transmissão de forma a evitar sobrecargas e consequentes desligamentos.

2.2 Consumidores livres

A comercialização de energia é um processo crucial no setor energético que envolve a venda e compra de energia elétrica entre geradores, comercializadores e consumidores.

Este processo é regulado por órgãos governamentais para garantir a segurança e a eficiência do mercado energético e diversas leis e portarias foram estabelecidas para regulamentar e orientar o setor.

Atualmente existem dois ambientes de comercialização: o Ambiente de Contratação Regulada (ACR), onde as distribuidoras compram energia por meio de

leilões para atender aos consumidores cativos, e o Ambiente de Contratação Livre (ACL), que permite a negociação direta entre fornecedores e consumidores livres ou especiais (Brasil, 1995). O foco desse trabalho é o Ambiente de Contratação Livre, pois esse mercado está aumentando cada vez mais devido a alterações nas leis.

2.2.1 Lei nº 9.074 de 1995

A Lei nº 9.074/95 (Brasil, 1995) introduziu o conceito do consumidor livre e estabeleceu as condições para que um consumidor pudesse se tornar livre, bem como as obrigações e direitos que essa categoria possui. Os consumidores livres são agentes que têm a liberdade de escolher seu fornecedor de energia, podendo negociar diretamente as condições de fornecimento, como preço e quantidade.

De acordo com o Art. 15 §2º (Brasil, 1995) ‘Os consumidores com carga igual ou superior a 3.000 kW, atendidos em tensão igual ou superior a 69 kV, poderão optar pela compra de energia elétrica a qualquer concessionário, permissionário ou autorizado de energia elétrica do mesmo sistema interligado.’

2.2.2 Portaria nº 514 de 2018.

Em 27 de Dezembro a portaria nº 514 (Subsecretaria de Planejamento, Acompanhamento e Gestão, 2018) estabeleceu novos limites para entrar no ambiente livre:

§1º A partir de 1º de julho de 2019, os consumidores com carga igual ou superior a 2.500 kW, atendidos em qualquer tensão, poderão optar pela compra de energia elétrica a qualquer concessionário, permissionário ou autorizado de energia elétrica do Sistema Interligado Nacional.

§2º A partir de 1º de janeiro de 2020, os consumidores com carga igual ou superior a 2.000 kW, atendidos em qualquer tensão, poderão optar pela compra de energia elétrica a qualquer concessionário, permissionário ou autorizado de energia elétrica do Sistema Interligado Nacional.

§3º A partir de 1º de janeiro de 2021, os consumidores com carga igual ou superior a 1.500 kW, atendidos em qualquer tensão, poderão optar pela compra de energia elétrica a qualquer concessionário, permissionário ou autorizado de energia elétrica do Sistema Interligado Nacional.

§4º A partir de 1º de janeiro de 2022, os consumidores com carga igual ou superior a 1.000 kW, atendidos em qualquer tensão, poderão optar pela compra de energia elétrica a qualquer concessionário, permissionário ou autorizado de energia elétrica do Sistema Interligado Nacional.

§5º A partir de 1º de janeiro de 2023, os consumidores com carga igual ou superior a 500 kW, atendidos em qualquer tensão, poderão optar pela compra de energia elétrica a qualquer concessionário, permissionário ou autorizado de energia elétrica do Sistema Interligado Nacional.

Essa regulamentação introduz um novo desafio na previsão de demanda elétrica dos consumidores livres, uma vez que facilita a entrada de consumidores de pequeno porte no ambiente livre que somados aumentam significativamente o consumo no ACL.

2.3 Previsão de demanda

A previsão tem por objetivo antecipar eventos futuros, utilizando dados históricos e análises correntes, sendo essencial para indivíduos, organizações e agências governamentais. É uma ferramenta que auxilia na tomada de decisões, ao permitir avaliar antecipadamente as possíveis consequências de uma ação, e, assim, direcionar escolhas mais assertivas em diversas situações, como em investimentos, planejamentos e estratégias de vendas (ARMSTRONG, 2001).

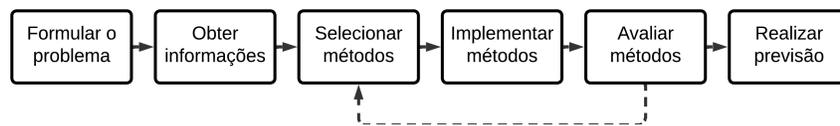
As técnicas de previsão podem ser classificadas em dois tipos principais: quantitativas e qualitativas. As técnicas qualitativas são subjetivas, por outro lado, as técnicas quantitativas se baseiam em análises de dados históricos, utilizando métodos como séries temporais e regressões, procurando padrões nas informações passadas para antecipar o futuro (SOUZA, 2021).

A escolha do método de previsão apropriado varia de acordo com a situação e a natureza dos dados disponível, sendo necessário assim a análise detalhada dos dados. Essa análise permite avaliar como diferentes fatores influenciam o resultado previsto. Segundo Armstrong (2001) para previsões de longo prazo, métodos econométricos podem ser mais adequados, enquanto as previsões de curto prazo normalmente são beneficiadas pelos métodos de extrapolação.

Dada a complexidade e a variabilidade dos problemas que se busca prever, é essencial a aplicação criteriosa de princípios de previsão, que orientem na escolha e aplicação dos métodos mais adequados, evitando práticas deficientes e contribuindo para decisões mais acertadas e embasadas (ARMSTRONG, 2001).

Conforme a Fig. 2.1 para realizar a previsão de algum dado é necessário seguir 6 etapas:

Figura 2.1 – Etapas da previsão



Fonte: Adaptado de Armstrong (2001)

Etapa 1 - Formular o problema: Definição clara do que se deseja alcançar com a previsão (ARMSTRONG, 2001) e (SOUZA, 2021).

Etapa 2 - Obter informações: Coleta de dados históricos relevantes para previsão. Neste etapa, considera-se utilizar a série temporal mais longa disponível para uma análise mais aprofundada e precisa (SOUZA, 2021).

Etapa 3 - Selecionar métodos: Com base nos dados coletados e analisados deve-se selecionar a(s) técnica(s) de previsão mais adequada no objetivo definido (ARMSTRONG, 2001) e (SOUZA, 2021).

Etapa 4 - Implementar métodos: Realizar previsões para curto, médio ou longo prazo utilizando a(s) técnica(s) selecionada(s) (ARMSTRONG, 2001) e (SOUZA, 2021).

Etapa 5 - Avaliar métodos: Após realizar as previsões, avalia-se as incertezas e precisão do(s) modelo(s) implementados. Caso a avaliação não tenha sido satisfatória, deve-se selecionar outros métodos ou melhorar os métodos anteriores (ARMSTRONG, 2001) e (SOUZA, 2021).

Etapa 6 - Realizar previsão: Com o(s) modelo(s) satisfatório(s) gera-se os resultados (ARMSTRONG, 2001) e (SOUZA, 2021).

2.4 Séries Temporais

Segundo Morettin e Toloí (2018) uma série temporal é qualquer conjunto de observações ordenadas no tempo. Formalmente, denota-se uma trajetória qualquer $Z^{(j)}(t)$ em que para cada t fixo tem-se os valores de uma variável aleatória $Z(t)$ que terá certa distribuição de probabilidades, sendo assim a série temporal é uma parte de uma trajetória, dentre muitas que poderiam ter sido observadas.

O objetivo desse trabalho ao obter a série temporal é fazer previsões da série $Z(t)$, com o conhecimento de observações da série de entrada $X(t)$ conforme a Eq. 2.1.

$$Z(t) = \sum_{\tau=0}^{\infty} v(\tau)X(t - \tau) \quad (2.1)$$

onde $X(t)$ é a série de entrada, $Z(t)$ é a série de saída, $v(t)$ é uma função de transferência e t é tempo (MORETTIN; TOLOI, 2018).

Ao analisar séries temporais, é essencial entender suas características e componentes. A seguir será explorado esses conceitos.

2.4.1 Estacionariedade

Uma série temporal é considerada estacionária quando suas propriedades estatísticas, como média e variância, permanecem constantes ao longo do tempo. Em outras palavras, uma série estacionária se desenvolve aleatoriamente ao redor

de uma média constante, indicando uma forma de equilíbrio estável (MORETTIN; TOLOI, 2018).

Como a maioria dos procedimentos de análise estatística de séries temporais supõe que estas sejam estacionárias e a maior parte das séries na prática apresentam alguma forma de não estacionariedade é necessário transformar os dados originais.

A transformação mais comum consiste em tomar diferenças sucessivas da série original, até se obter uma série estacionária. A n-ésima diferença de $Z(t)$ é dada descrita pela Eq. 2.2 onde Δ representa o operador de diferença (MORETTIN; TOLOI, 2018).

$$\Delta^n Z(t) = \Delta[\Delta^{n-1} Z(t)] \quad (2.2)$$

Para determinar se a série é estacionária pode ser feito o teste ADF (*Augmented Dickey-Fuller*). Esse teste consiste no(a):

1. **Formulação da hipótese nula e alternativa:** A hipótese nula (H_0) é definida como a situação em que a série temporal apresenta uma raiz unitária, tornando-a não estacionária. Enquanto isso, a hipótese alternativa (H_1) postula que a série temporal adquire estacionariedade após a aplicação de diferenciações apropriadas, ou seja, depois de passar por um processo de 'd' diferenças.
2. **Cálculo da Estatística do Teste:** A estatística de teste do ADF compara a variação da série original com a variação após a diferenciação. Quanto mais a série for estacionária, menor será a variação após a diferenciação, levando a uma estatística de teste mais negativa.
3. **Comparação com Valores Críticos:** A estatística do teste de ADF é comparada a valores críticos de acordo com o nível de significância escolhido (geralmente 5% ou 1%). Se a estatística do teste for menor (mais negativa)

que o valor crítico correspondente, a hipótese nula é rejeitada em favor da hipótese alternativa, indicando que a série é estacionária.

4. Interpretação dos resultados:

- (a) Se o valor da estatística do teste de ADF for menor que os valores críticos, a série é considerada estacionária, e nenhuma ou pouca diferenciação é necessária.
- (b) Se o valor da estatística do teste for maior que os valores críticos, a série é considerada não estacionária, e a ordem de diferenciação 'd' deve ser aumentada até que a série se torne estacionária.

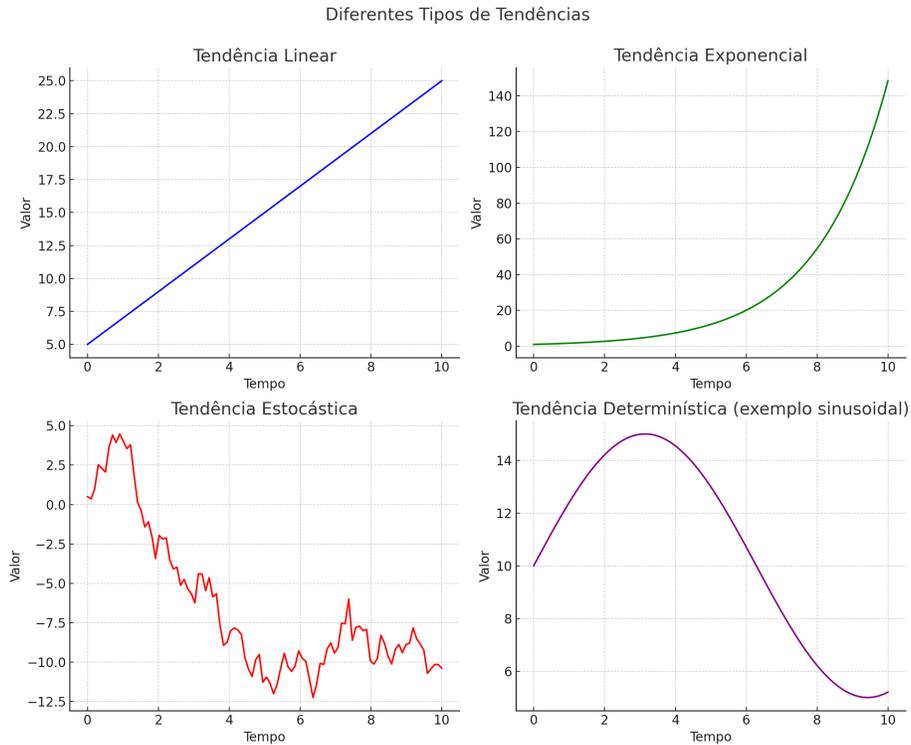
2.4.2 Tendência

Tendência refere-se a um padrão de longo prazo em uma série temporal, sem considerar as flutuações sazonais e irregulares. Pode ser linear, curvilínea, estocásticas ou determinística conforme mostra a Fig. 2.1 (BOX et al., 2015).

- **Linear:** Representada por uma linha reta que pode ter uma inclinação positiva (crescente) ou negativa (decrecente).
- **Curvilínea:** Caracterizada por uma curva que pode ser côncava (se a série cresce a uma taxa crescente) ou convexa (se a série cresce a uma taxa decrescente).
- **Estocásticas:** Não tem uma forma definida, já que é influenciada por mudanças aleatórias ao longo do tempo.
- **Determinística:** Pode assumir uma variedade de formas, dependendo do padrão subjacente que se repete ao longo do tempo.

Dentro da tendência pode haver movimentos ascendentes ou descendentes que ocorrem de forma cíclica dentro desses movimentos são denominados de ciclo. (SOUSA et al., 2021)

Gráfico 2.1 – Tipos de tendência



Fonte: Próprio autor

2.4.3 Sazonalidade

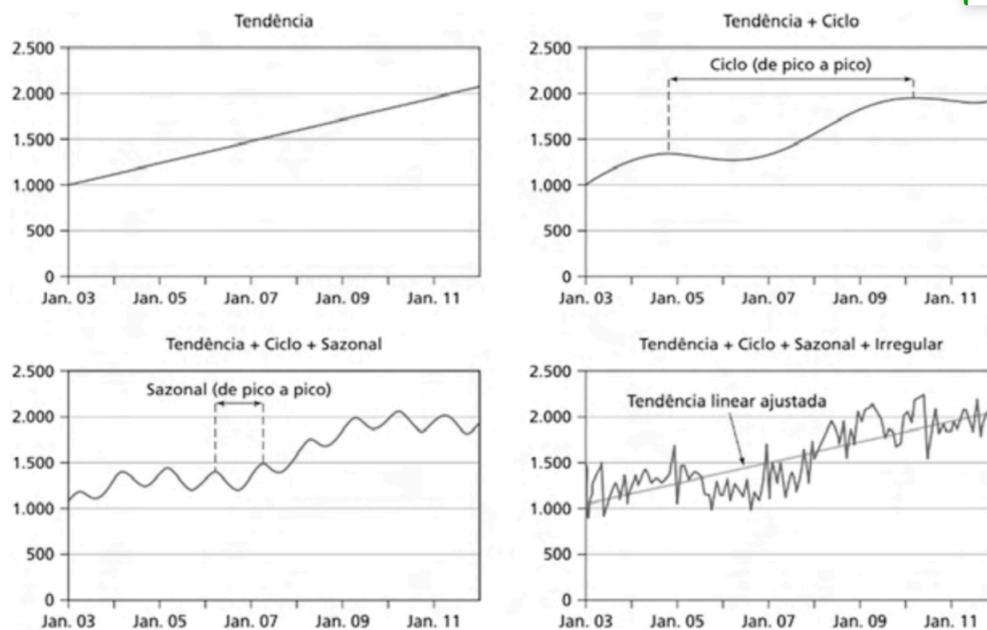
Sazonalidade refere-se a flutuações que ocorrem em intervalos regulares, geralmente devido a fatores externos, como o clima ou feriados. Estas flutuações são previsíveis e podem ser repetidas em um período específico (CHATFIELD, 2016).

2.4.4 Resíduo

O resíduo é a diferença entre o valor observado e o valor previsto por um modelo. É utilizado para avaliar a precisão de um modelo de previsão (MORETTIN; TOLOI, 2018).

O resíduo também pode ser chamado de ruído ou irregularidade. Ele é um erro aleatório presente no modelo de série temporal e como pode ser observado na Fig. 2.2 tudo que não representar uma tendência, um ciclo ou uma sazonalidade é considerado um erro aleatório, ou seja, uma irregularidade (SOUSA et al., 2021).

Gráfico 2.2 – Componentes de uma série temporal



Fonte: Sousa et al. (2021)

2.4.5 Modelos de Séries Temporais

Para isolar os componentes sazonalidade (S), tendência (T), ciclo (C) e irregularidade (I), são utilizadas técnicas estatísticas. Dentre vários modelos para isolar os componentes, destaca-se os modelos aditivo e multiplicativo (SOUSA et al., 2021).

Matematicamente o modelo aditivo é descrito pela Eq. 2.3.

$$X(t) = T + S + C + I \quad (2.3)$$

Já o modelo multiplicativo é descrito pela Eq. 2.4.

$$X_t = T \times S \times C \times E \quad (2.4)$$

Existem outros modelos avançados, como os modelos ARIMA, que combinam técnicas de médias móveis e autoregressão que serão explorados a seguir.

2.5 Regressão

Regressão é uma técnica estatística usada para modelar a relação entre uma ou mais variáveis independentes (nomeada também de características, preditores ou variáveis explicativas) e uma variável dependente (denominadas como variável de resposta ou alvo). O principal objetivo da regressão é entender como as variáveis independentes influenciam a variável dependente e, com base nessa análise, fazer previsões ou estimativas.

2.5.1 Linear

Regressão Linear é uma técnica estatística que busca estabelecer uma relação linear entre uma ou mais variáveis independentes e uma variável dependente.

Para uma única variável independente, a equação que descreve a regressão linear é dada pela Eq. 2.5. Nessa equação, y é a variável dependente que está sendo prevista, x_1 é a variável independente, β_0 é o intercepto da regressão (ou seja, onde a linha de regressão corta o eixo vertical) e β_1 é o coeficiente da variável independente x_1 .

$$y = \beta_0 + \beta_1 x_1 \quad (2.5)$$

E para múltiplas variáveis independentes tem-se a Eq. 2.6.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.6)$$

onde y é a variável dependente que deseja-se prever, x_1, x_2, \dots, x_p são as variáveis independentes, β_0 é o intercepto da regressão e $\beta_1, \beta_2, \dots, \beta_p$ são os coeficientes das variáveis independentes.

Como a regressão linear assume que a relação entre as variáveis é estritamente linear e, em alguns casos, as relações não podem ser adequadamente modeladas por uma linha reta faz-se necessário o uso da regressão polinomial.

2.5.2 Polinomial

A regressão polinomial é uma extensão da regressão linear que permite modelar relações não lineares entre as variáveis. Em vez de ajustar uma linha reta aos dados, a regressão polinomial ajusta um polinômio de grau n aos dados. A regressão polinomial de grau n é expressa pela Eq. 2.7.

$$y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \dots + \beta_nx^n \quad (2.7)$$

Nessa equação, y é a variável dependente que está sendo prevista, x é a variável independente, e $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ são os coeficientes do polinômio. O grau do polinômio, representado por n , determina a complexidade da relação que está sendo modelada. Quanto maior o grau, mais flexível o modelo se torna, mas também corre o risco de se ajustar demais aos dados de treinamento (*overfitting*).

A regressão polinomial permite capturar padrões mais complexos do que uma simples regressão linear, no entanto, no contexto de séries temporais, é importante considerar também os modelos autorregressivos, pois eles reconhecem a influência dos valores passados sobre os valores futuros, tornando-os essenciais para a previsão de tendências e padrões em dados com dependência temporal.

2.6 Modelo autorregressivo (AR) de média móvel (MA)

Modelos Autorregressivos (AR) são uma abordagem comum para modelar e prever séries temporais (CHATFIELD, 2016; BOX et al., 2015). A característica principal desse modelo é que ele usa observações anteriores da própria série temporal para prever observações futuras. Ou seja, a previsão de um ponto no tempo depende das observações passadas da mesma série temporal. Um modelo AR de ordem p (denotado como $AR(p)$) é descrito pela Eq. 2.8.

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (2.8)$$

onde X_t representa o valor da série temporal no tempo t , c é uma constante (ou termo de interceptação), $\phi_1, \phi_2, \dots, \phi_p$ são os coeficientes associados aos valores passados da série, p é a ordem da autorregressão, indicando quantos valores passados estão sendo considerados e ε_t é o erro aleatório no tempo (t).

Para ajustar um modelo AR aos dados, é necessário estimar os coeficientes $\phi_1, \phi_2, \dots, \phi_p$ e a constante c com base nos dados observados. Existem várias abordagens para realizar esse ajuste, sendo a mais comum o uso de métodos de estimação, como o Método dos Mínimos Quadrados.

2.6.1 Método dos Mínimos Quadrados

O Método dos Mínimos Quadrados é uma técnica de ajuste de parâmetros usada para encontrar os melhores valores dos coeficientes de um modelo, de modo a minimizar a soma dos quadrados dos erros entre as previsões do modelo e os dados observados (Eq. 2.9).

$$L = \sum_{t=1}^N (X_t - \hat{X}_t)^2 \quad (2.9)$$

onde L é a função de perda, N é o número de observações na série temporal, X_t é o valor real da série no tempo t e \hat{X}_t é o valor previsto pelo modelo no tempo t .

O objetivo é encontrar os valores dos coeficientes $\phi_1, \phi_2, \dots, \phi_p$ que minimizam a função de perda L . Uma vez que o modelo é ajustado, ele pode ser usado para fazer previsões futuras, com as previsões dependendo dos valores passados da série.

2.6.2 Modelo de média móvel (MA)

O modelo de Média Móvel (MA) descreve como os valores de uma série temporal podem ser modelados com base em uma combinação linear de valores passados dos resíduos ou erros anteriores. Um modelo MA de ordem q (denotado como $MA(q)$) é descrito pela Eq. 2.10.

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2.10)$$

onde X_t é o valor da série temporal no momento t , μ é a média da série temporal (parâmetro constante), ε_t é o erro no momento t , representando a variação não explicada pela média, os parâmetros $\theta_1, \theta_2, \dots, \theta_q$ são coeficientes da média móvel para os q períodos anteriores e q é a ordem da média móvel, indicando quantos lags (atrasos) dos erros passados são considerados.

Para ajustar um modelo MA aos dados, é necessário determinar a ordem (q) do modelo, para isso pode-se utilizar a busca em grade. Essa técnica envolve a exploração de um conjunto predefinido de valores candidatos para encontrar o melhor parâmetro ou combinação de parâmetros. Ao testar diversas combinações de valores candidatos, essa abordagem permite identificar a configuração que resulta no melhor ajuste aos dados observados, com base em métricas de desempenho definidas. O procedimento da busca em grade segue as etapas:

1. **Especificação dos possíveis valores:** Na primeira etapa, uma faixa de valores possíveis para o parâmetro é definida, representando as opções a serem exploradas.

2. **Iteração Através dos Valores Candidatos:** Para cada valor possível deve ser feito a iteração com os passos subsequentes.

- (a) **Ajuste do Modelo:** O modelo é ajustado e treinado com um dos valores possíveis.
- (b) **Avaliação do Desempenho:** O desempenho do modelo é avaliado utilizando métricas de desempenho, como erro médio quadrático (MSE).
- (c) **Seleção do Melhor Valor:** O valor que resulta no melhor desempenho é selecionado.

2.6.3 Autoregressive Moving Average (ARMA)

O modelo ARMA combina os modelos autorregressivos (AR) e média móvel (MA) com o objetivo de capturar a estrutura de autocorrelação e média móvel presente nos dados de séries temporais. Um modelo ARMA de ordem de autorregressão p e ordem de média móvel q (denotado como $ARMA(p, q)$) é descrito pela Eq. 2.11.

$$\begin{aligned} X_t = & c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} \\ & + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \end{aligned} \quad (2.11)$$

onde X_t é o valor da série temporal no tempo t , c é a constante ou intercepto na equação, representando a componente média da série temporal ao longo do tempo. Os termos autorregressivos, $\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p}$, refletem a influência dos valores passados da série temporal, ponderados pelos coeficientes autorregressivos ϕ_i . O termo ε_t é o erro no tempo t , representando a componente não explicada ou o ruído branco na série temporal. Os termos de médias móveis,

$\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$, refletem a influência dos erros passados, onde θ_i são os coeficientes de médias móveis.

O objetivo da modelagem ARMA é estimar os coeficientes autorregressivos, os coeficientes da média móvel e a constante, para que o modelo seja capaz de capturar a estrutura de autocorrelação e média móvel presente na série temporal. Com um modelo ARMA bem ajustado, é possível fazer previsões futuras e analisar o comportamento da série temporal ao longo do tempo.

2.6.4 Autoregressive Integrated Moving Average (ARIMA)

O modelo ARIMA é uma extensão do modelo ARMA e foi desenvolvido para lidar com séries temporais não estacionárias. Para atingir esse objetivo, ele incorpora o conceito de integração ao modelo ARMA. A etapa de integração envolve a diferenciação da série temporal original, a fim de torná-la estacionária, conforme explicado na sessão 2.4.1. O número de diferenças necessárias para alcançar a estacionariedade é denotado por 'd'. A série diferenciada, representada como X'_t , é obtida por meio da Eq. 2.12.

$$X'_t = X_t - X_{t-d} \quad (2.12)$$

O modelo ARIMA pode ser expresso conforme a Eq. 2.13.

$$\begin{aligned} X'_t = & c + \phi_1 X'_{t-1} - \phi_2 X'_{t-2} - \dots - \phi_p X'_{t-p} \\ & + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \end{aligned} \quad (2.13)$$

2.7 Redes Neurais Artificiais

Redes neurais artificiais (RNAs) são modelos computacionais inspirados no processamento de informações no cérebro humano. Elas pertencem ao campo de aprendizado de máquina e constroem modelos a partir de dados de entrada e

saída. Haykin (2001) compara RNAs ao cérebro, destacando que ambos adquirem conhecimento através de aprendizagem e usam conexões neuronais (pesos sinápticos) para armazenar este conhecimento. (HAYKIN, 2001) ainda cita como benefícios da redes neurais:

1. **Generalização:** RNAs processam entradas não vistas durante o treinamento.
2. **Não-linearidade:** Permitem modelagem de relações complexas.
3. **Mapeamento de Entrada-Saída:** Aprende mapeamentos a partir de exemplos rotulados.
4. **Adaptabilidade:** Ajustam-se a mudanças ambientais.
5. **Resposta a Evidências:** Fornece confiabilidade em decisões de classificação.
6. **Informação Contextual:** Tratam informações distribuídas na rede.
7. **Tolerância a Falhas:** Degradam-se suavemente sob condições adversas.
8. **Implementação em VLSI:** Adequadas para tecnologias de integração de alta escala.
9. **Analogia Neurobiológica:** Motivadas pela estrutura e funcionamento do cérebro, fornecendo insights para a eletrônica e a tecnologia VLSI.

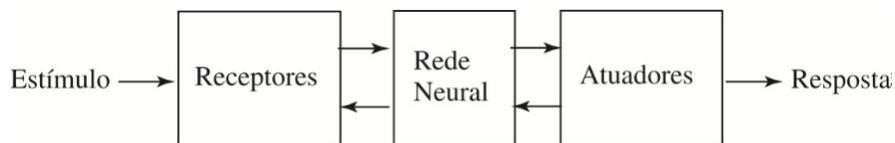
Sabendo da inspiração neurobiológica, é pertinente fazer uma breve análise da estrutura e organização do cérebro humano.

2.7.1 Cérebro humano

O sistema nervoso humano pode ser descrito como um conjunto de três estágios: Receptores, Rede Neural Nervosa e Atuadores, conforme ilustrado na

Fig. 2.2. No coração deste sistema está o cérebro, que funciona como uma rede neural. Esta rede está constantemente recebendo e interpretando informações, e tomando decisões com base na entrada. Os receptores transformam estímulos do corpo ou do ambiente externo em impulsos elétricos, que são enviados ao cérebro. Em resposta, os atuadores transformam os impulsos elétricos do cérebro em ações visíveis.

Figura 2.2 – Representação em diagrama em blocos do sistema nervoso

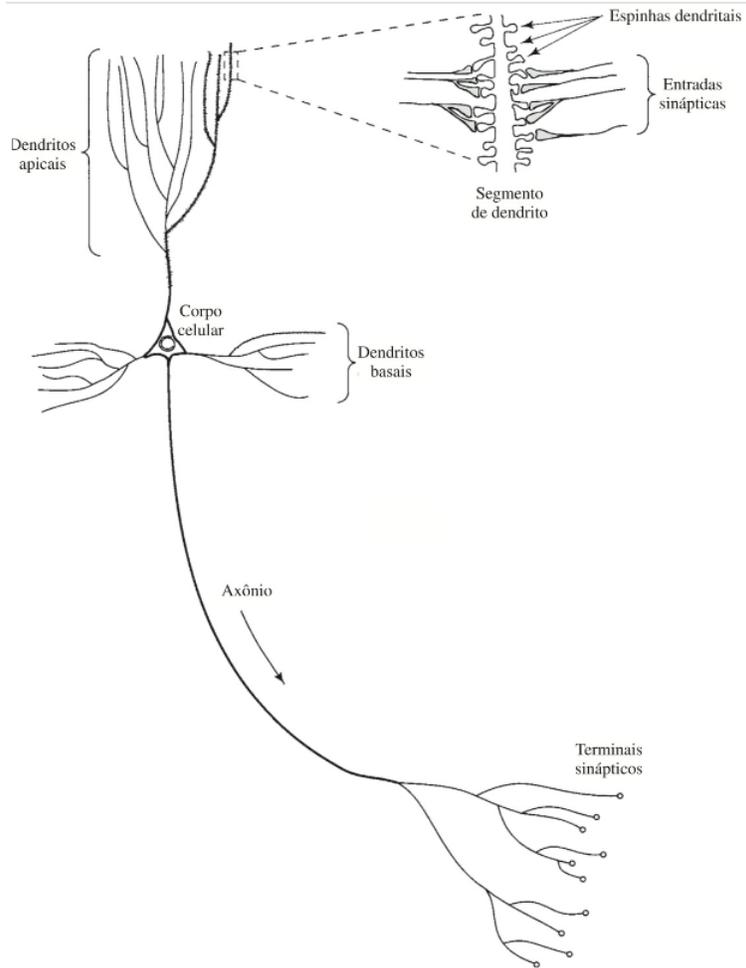


Fonte: Haykin (2001)

Os neurônios vêm em uma variedade de formas e tamanhos. Um exemplo é a célula piramidal (Fig. 2.3), que é um tipo comum de neurônio cortical. Esta célula pode receber muitos contatos sinápticos e se conectar a muitas outras células.

Os neurônios comunicam-se através de potenciais de ação, ou impulsos, que são pulsos breves de tensão. Esta comunicação começa quando um neurônio é estimulado e gera um potencial de ação. Este impulso elétrico viaja pelo axônio, uma extensão longa e fina do neurônio. Ao chegar ao final do axônio, o impulso chega a estruturas chamadas terminais sinápticos. Esses terminais estão próximos dos dendritos de outro neurônio. Os dendritos são extensões ramificadas que saem do corpo celular do neurônio e recebem informações de outros neurônios. Na junção entre o axônio de um neurônio e o dendrito de outro substâncias químicas chamadas neurotransmissores são liberadas. Estes neurotransmissores atravessam a sinapse e se ligam a receptores específicos no dendrito do próximo neurônio. Isso pode resultar na geração de um novo potencial de ação no neurônio receptor, continuando a transmissão da informação.

Figura 2.3 – A célula piramidal



Fonte: Haykin (2001)

As sinapses são fundamentais para a comunicação neural e podem converter um sinal elétrico pré-sináptico em um sinal químico e, em seguida, de volta em um sinal elétrico pós-sináptico. Este processo de comunicação e modulação é aprimorado pela plasticidade, que refere-se à sua capacidade de mudar e adaptar-se. Em cérebros adultos, isso envolve a criação de novas conexões sinápticas e a modificação das existentes.

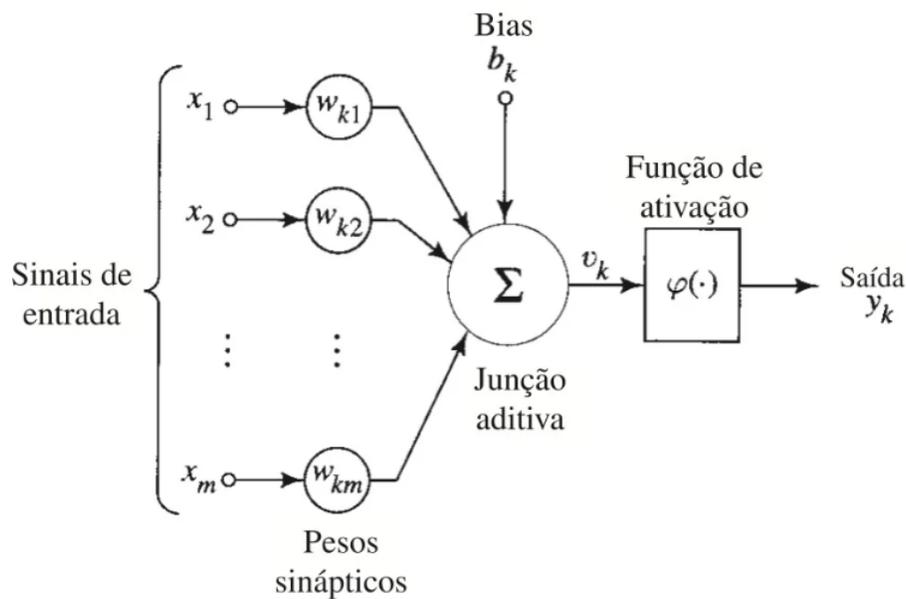
O cérebro opera de forma mais lenta do que os dispositivos eletrônicos modernos, mas compensa essa lentidão com uma quantidade surpreendente de

neurônios e sinapses, tornando-o extremamente eficiente. A eficiência energética do cérebro é notavelmente superior à dos melhores computadores da atualidade.

2.7.2 Neurônio Artificial

Um neurônio é uma unidade de processamento essencial na operação de redes neurais. A Fig. 2.4 ilustra o modelo de um neurônio, que serve como base para a construção de redes neurais artificiais. Esse modelo tem três componentes principais:

Figura 2.4 – Modelo não-linear de um neurônio



1. **Sinapses ou elos de conexão:** Cada sinapse é caracterizada por um peso. Um sinal x_j de entrada é multiplicado pelo peso sináptico w_{kj} . Nota-se que o peso sináptico de um neurônio artificial pode variar entre valores positivos e negativos, diferentemente das sinapses biológicas.
2. **Somador:** Ele agrega os sinais de entrada ponderados.

3. **Função de ativação:** Limita a amplitude da saída do neurônio. Geralmente, a amplitude é normalizada para o intervalo $[0,1]$ ou $[-1,1]$.

Além disso, o modelo incorpora um bias b_k , que ajusta a entrada da função de ativação. Dependendo do valor do bias (positivo ou negativo), a entrada líquida para a função de ativação é aumentada ou diminuída, respectivamente.

Matematicamente, o neurônio k é representado pelas Eqs. 2.14 e 2.15.

$$u_k = \sum_{j=1}^m w_{kj}x_j \quad (2.14)$$

$$y_k = \phi(u_k + b_k) \quad (2.15)$$

onde, x_1, x_2, \dots, x_m são os sinais de entrada, $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos sinápticos, u_k é a saída do somador, b_k é o bias, $\phi(\cdot)$ é a função de ativação, e y_k é a saída do neurônio. O bias b_k resulta em uma transformação afim na saída v_k descrita pela Eq. 2.16.

$$v_k = u_k + b_k \quad (2.16)$$

O bias b_k é um parâmetro externo. Ele pode ser incorporado nas equações conforme demonstrado nas Eqs. 2.17 e 2.18.

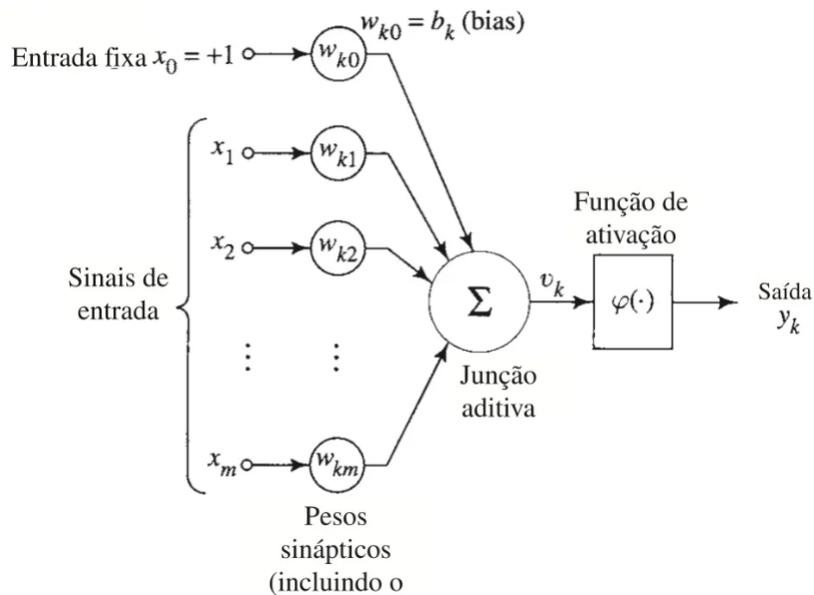
$$v_k = \sum_{j=0}^m w_{kj}x_j \quad (2.17)$$

$$y_k = \phi(v_k) \quad (2.18)$$

Nessa formulação, uma nova sinapse é adicionada com entrada $x_0 = +1$ e peso $w_{k0} = b_k$. Assim, o modelo do neurônio k pode ser alternativamente representado como na Fig. 2.5, onde o efeito do bias é considerado adicionando um novo

sinal de entrada e um novo peso sináptico igual ao bias b_k . Embora os modelos pareçam diferentes, eles são matematicamente equivalentes.

Figura 2.5 – Um outro modelo não-linear de um neurônio



2.7.3 Tipos de Função de Ativação

A função de ativação, denotada por $\phi(v)$, determina a saída de um neurônio com base em seu potencial ativação local v . Existem cinco funções de ativação fundamentais: Degrau, Linear e Sigmóide, Tangente Hiperbólica e ReLU.

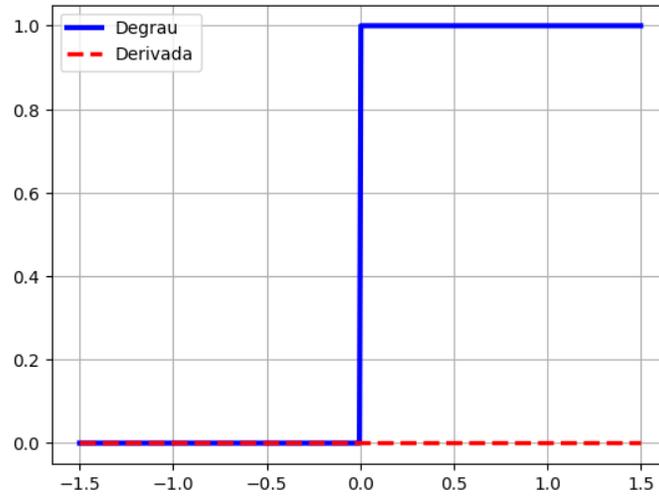
2.7.3.1 Função Degrau

A função degrau é representada na Fig. 2.3, é dada pela Eq. 2.19.

$$\phi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases} \quad (2.19)$$

Esta é frequentemente chamada de função de Heaviside na engenharia. A saída do neurônio k que usa esta função é definida pela Eq. 2.20.

Gráfico 2.3 – Função Degrau



Fonte: Próprio autor

$$y_k = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases} \quad (2.20)$$

onde v_k é definido pela Eq. 2.21.

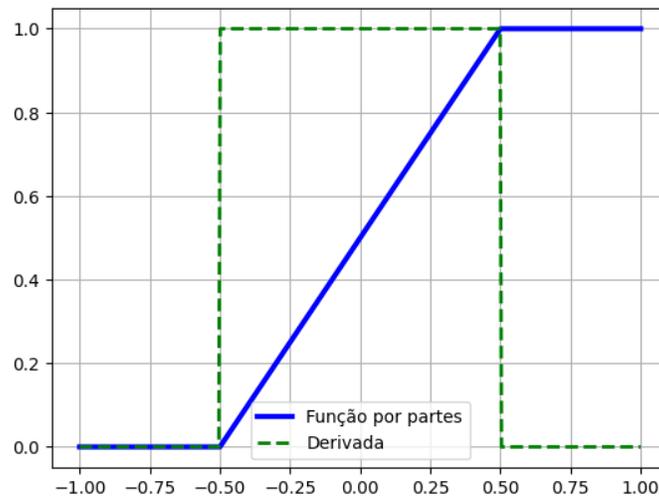
$$v_k = \sum_{j=1}^m w_{kj}x_j + b_k \quad (2.21)$$

Este neurônio destaca a propriedade de ‘tudo ou nada’ do modelo, isso refere-se à resposta binária de um neurônio a um estímulo: ou ele dispara um potencial de ação em amplitude máxima ou não dispara nada, sem estados intermediários. Esta característica garante a transmissão clara e consistente de informações nas redes neurais.

2.7.3.2 Função Linear por Partes

A função linear por partes é representada na Fig. 2.4 e é descrita pela Eq. 2.22.

Gráfico 2.4 – Função Linear por partes



Fonte: Próprio autor

$$\phi(v) = \begin{cases} 1 & \text{se } v \geq +\frac{1}{2} \\ v & \text{se } +\frac{1}{2} > v > -\frac{1}{2} \\ 0 & \text{se } v \leq -\frac{1}{2} \end{cases} \quad (2.22)$$

Ela pode ser vista como uma aproximação de um amplificador não-linear. Existem duas variações notáveis desta função:

- Quando a região linear opera sem saturação, ela se torna um combinador linear.
- Quando a amplificação da região linear é infinitamente grande, ela se torna uma função degrau.

A função de ativação linear por partes é simples e flexível, com uma derivada fácil de calcular, tornando-a benéfica para certas implementações. No entanto, suas descontinuidades e limitações na representação de dados complexos, juntamente com problemas potenciais de saturação, podem apresentar desafios em redes neurais profundas e algoritmos específicos.

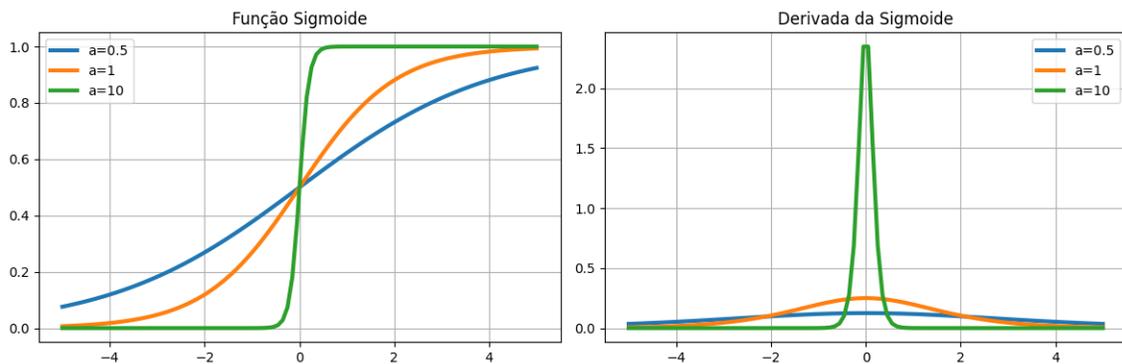
2.7.3.3 Função Sigmóide

A função sigmóide é representada na Fig. 2.5 e é descrita pela Eq. 2.23.

$$\phi(v) = \frac{1}{1 + \exp(-av)} \quad (2.23)$$

O parâmetro a determina a inclinação da função. As funções sigmóides com diferentes inclinações são ilustradas na Fig. 2.5. Nota-se que quando a tende ao infinito, a função sigmóide converge para uma função de limiar (degrau).

Gráfico 2.5 – Função Sigmóide



Fonte: Próprio autor

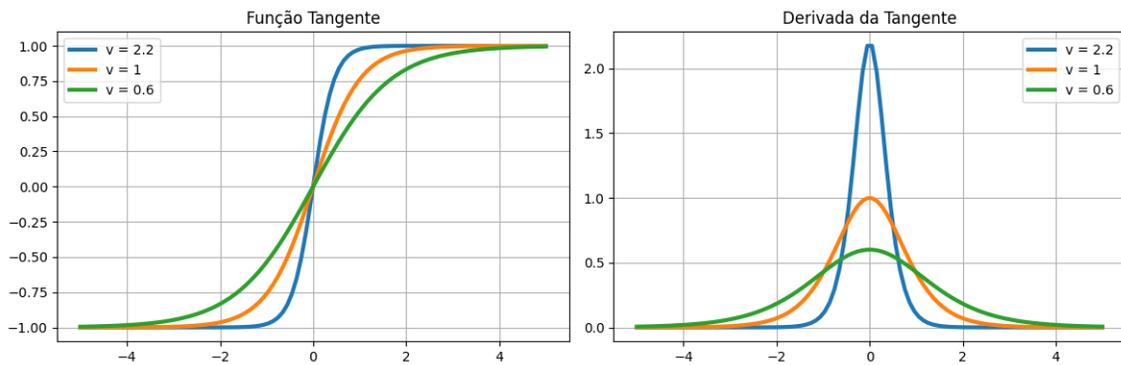
Além disso, a função sigmóide é diferenciável, o que é uma propriedade crucial na teoria de redes neurais.

2.7.3.4 Função Tangente Hiperbólica

A função tangente hiperbólica é representada na Fig. 2.6 e é descrita pela Eq. 2.24.

$$\phi(v) = \tanh(v) = \frac{e^v - e^{-v}}{e^v + e^{-v}} \quad (2.24)$$

Gráfico 2.6 – Função Tangente Hiperbólica



Fonte: Próprio autor

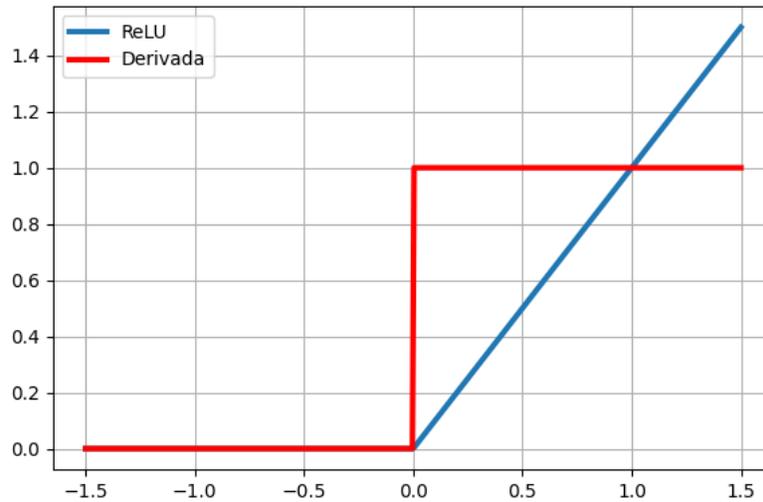
A função tangente hiperbólica mapeia um número real v para um valor no intervalo de -1 a 1, tornando-a uma função útil para normalizar a saída de neurônios em redes neurais. Esta função é semelhante à função sigmóide, mas tem a vantagem de ser centrada em zero, o que pode facilitar o treinamento de redes neurais. Assim como a função sigmóide, a função tangente hiperbólica é diferenciável.

2.7.3.5 Função ReLU

A função ReLU é representada na Fig. 2.7 e é descrita pela Eq. 2.25.

$$\phi(v) = \max(0, v) \quad (2.25)$$

Gráfico 2.7 – Função ReLU



Fonte: Próprio autor

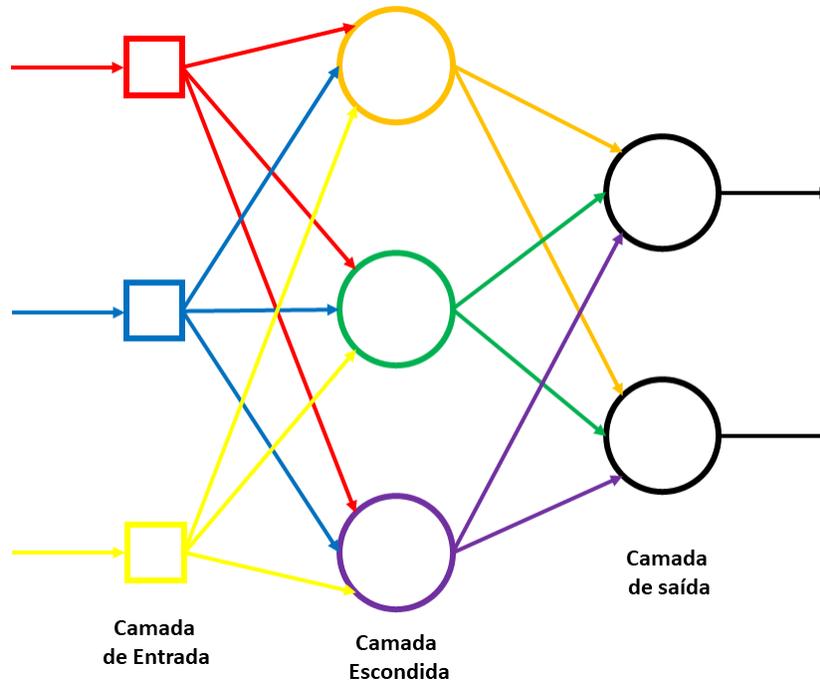
A função ReLU é zero para valores negativos de v e aumenta linearmente para valores positivos de v . A simplicidade da função ReLU e a sua capacidade de introduzir não-linearidade nas redes neurais a tornam uma escolha popular em arquiteturas de redes profundas.

2.7.4 Arquitetura da Rede

As redes neurais artificiais são a base de muitos sistemas de aprendizado de máquina e inteligência artificial. Elas são compostas por várias camadas interconectadas de neurônios (ou unidades de processamento), cada uma com seu próprio conjunto de parâmetros ajustáveis. Normalmente as camadas são classificadas conforme a Fig. 2.6 em:

- **Camada de Entrada:** É a primeira camada de uma rede neural e serve para receber dados brutos ou características de entrada. Cada neurônio nesta camada corresponde a uma característica ou dimensão dos dados de entrada. A camada de entrada transmite os valores para as camadas ocultas;

Figura 2.6 – Camadas de uma rede



Fonte: Adaptado de Lacerda (2023a)

- **Camadas Intermediárias ou Ocultas:** É a camada que conecta a camada de entrada à camada de saída e onde a maior parte do processamento é feita. Nela cada neurônio em uma camada oculta é conectado a todos os neurônios da camada anterior e da camada seguinte. A presença de múltiplas camadas ocultas permite que a rede aprenda representações mais abstratas e complexas dos dados, tornando-a capaz de realizar tarefas mais sofisticadas, como classificação de imagens ou processamento de linguagem natural.
- **Camada de Saída:** É a última camada de uma rede neural e produz os resultados finais ou previsões com base no processamento dos dados de entrada. O número de neurônios na camada de saída geralmente corresponde ao número de classes ou categorias nas quais a rede deve fazer previsões. A saída de cada neurônio na camada de saída representa a probabilidade da en-

trada pertencer à classe correspondente. A classe com a maior probabilidade geralmente é escolhida como a previsão final da rede.

À estrutura organizacional da rede, ou seja, o número de camadas, o número de neurônios em cada camada, as conexões entre neurônios e como essas camadas estão dispostas, podem variar bastante, mas os tipos mais comuns de arquitetura são: *Feedforward* de uma única camada, *Feedforward* de múltiplas camadas e Redes Neurais Recorrentes.

2.7.4.1 *Feedforward* de uma única camada (*Perceptron*)

A a arquitetura *feedforward* de uma única camada é uma rede neural simples com apenas camada de entrada e saída (Fig. 2.7), adequada para tarefas simples e linearmente separáveis, mas inadequada para tarefas mais complexas que exigem a capacidade de aprendizado de representações não lineares.

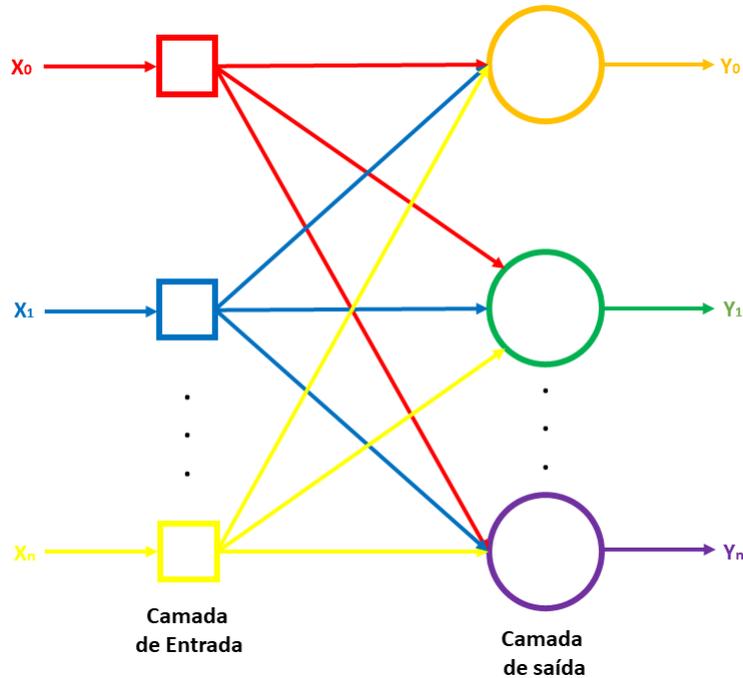
Neste tipo de rede, todas as unidades de processamento estão dispostas em uma única camada, e as conexões fluem em uma única direção, da entrada para a saída. Essas redes são frequentemente usadas para problemas simples de classificação ou regressão, quando a relação entre as entradas e as saídas é linear ou quase linear.

2.7.4.2 *Feedforward* de múltiplas camadas

Uma rede neural *feedforward* de múltiplas camadas, também conhecida como MLP (*Multilayer Perceptron*), é uma arquitetura mais avançada e poderosa em comparação com uma rede neural de camada única. Ela é caracterizada por ter uma ou mais camadas ocultas entre a camada de entrada e a camada de saída (Fig. 2.8).

A operação dessa arquitetura é baseada em uma série de camadas de neurônios interconectados. Começa com uma camada de entrada que recebe dados brutos ou características de entrada. Esses dados são propagados através de uma ou

Figura 2.7 – Arquitetura *feedforward* de uma única camada (*Perceptron*)

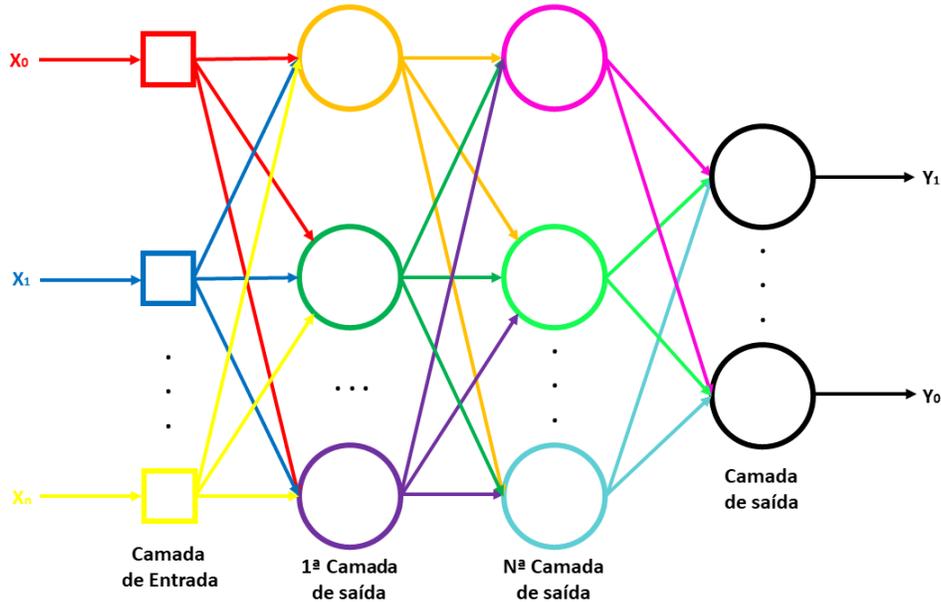


Fonte: Adaptado de Lacerda (2023a)

mais camadas ocultas, onde cada neurônio aplica transformações não lineares às entradas. A presença de camadas ocultas permite que o MLP aprenda representações complexas e abstratas dos dados, tornando-o adequado para tarefas de alta complexidade. Finalmente, os resultados são produzidos na camada de saída, com uma função de ativação apropriada para a tarefa.

Durante o treinamento, os pesos das conexões entre neurônios são ajustados por meio de algoritmos de otimização, como a retropropagação (*backpropagation*), para minimizar uma função de perda. Esse processo de ajuste iterativo permite que o MLP aprenda a mapear as entradas para as saídas desejadas.

Figura 2.8 – Arquitetura feedforward de múltiplas camadas



Fonte: Adaptado de Lacerda (2023a)

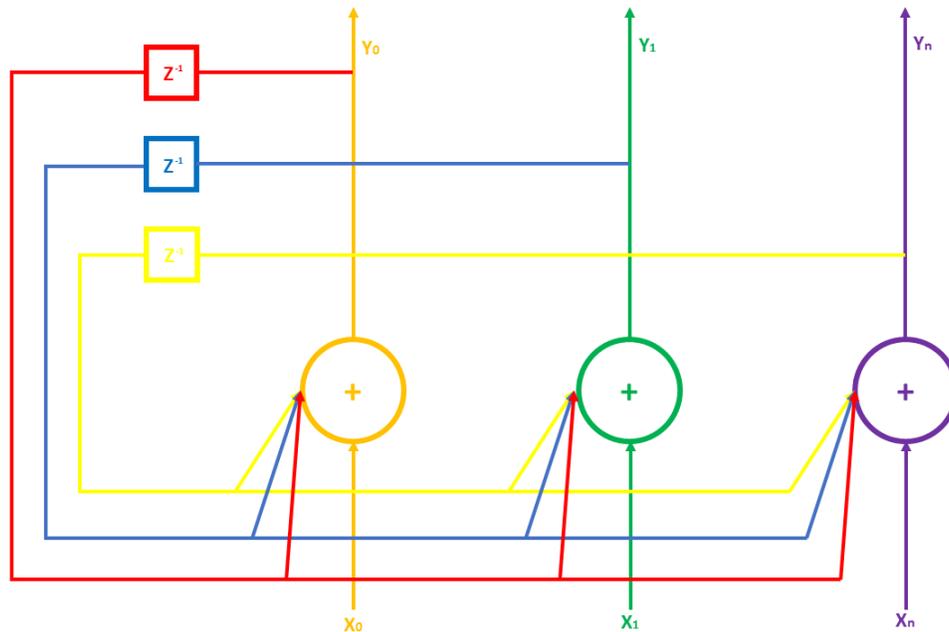
2.7.4.3 Redes Neurais Recorrentes

Ao contrário das redes *feedforward*, as redes recorrentes (RNNs) possuem conexões retroativas, o que permite que informações passem de uma etapa de tempo para a próxima conforme pode ser visualizado na Fig. 2.9. Isso torna as RNNs adequadas para tarefas que envolvem sequências de dados, como processamento de linguagem natural, previsão de séries temporais e muito mais. As RNNs podem lidar com entradas de comprimentos variáveis e manter estados anteriores em memória.

2.7.5 Processos de Aprendizagem

Uma das características mais essenciais das redes neurais é sua capacidade de aprender com o ambiente e, assim, aprimorar seu desempenho. Esse processo ocorre por meio de iterações nas quais os pesos da rede são ajustados, o que é

Figura 2.9 – Arquitetura da Rede Neural Recorrente



Fonte: Adaptado de Lacerda (2023a)

conhecido como treinamento. O aprendizado é concluído quando a rede neural atinge uma solução geral para uma classe específica de problemas.

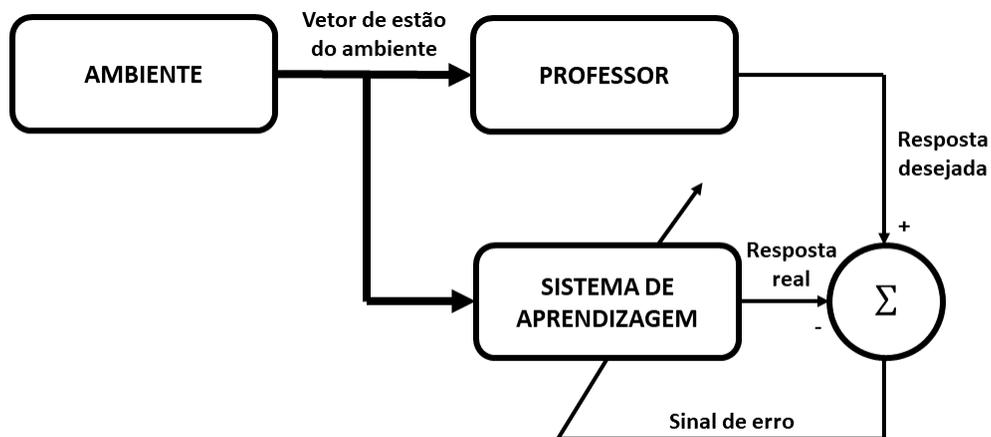
No processo de treinamento é essencial contar com algoritmos de aprendizado, que consistem em conjuntos bem definidos de regras para resolver problemas de aprendizado. Existem diversos tipos de algoritmos de aprendizado, cada um adaptado a modelos particulares de redes neurais, os três principais são Supervisionado, Não-Supervisionado e o Aprendizado por Reforço

2.7.5.1 Supervisionado

No aprendizado supervisionado, o modelo é treinado usando um conjunto de dados que contém exemplos de entrada e saída correspondente (Fig. 2.10). O objetivo é que o modelo aprenda a mapear corretamente as entradas para as saídas. Isso é feito ajustando os parâmetros do modelo de acordo com a diferença entre as

previsões do modelo e os valores reais do conjunto de treinamento. O aprendizado supervisionado é comumente usado em tarefas de classificação e regressão.

Figura 2.10 – Sistema de aprendizado supervisionado



Fonte: Adaptado de Lacerda (2023b)

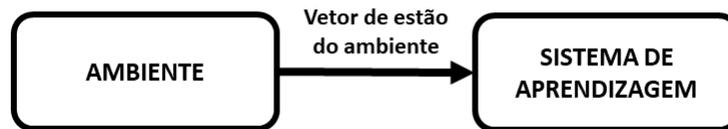
Para realizar o aprendizado supervisionado deve-se:

1. Ter o conjunto de dados de treinamento, que contém dados de entrada juntamente com as saídas desejadas correspondentes.
2. Aplicar o algoritmo de aprendizado ao conjunto de dados de treinamento para ajustar os parâmetros de um modelo. O objetivo é que o modelo aprenda a mapear as entradas para as saídas desejadas, de modo que, com o tempo, possa fazer previsões precisas para novos exemplos.
3. Avaliar o desempenho do modelo utilizando um conjunto de dados de teste que não foi aplicado durante o treinamento. O objetivo é verificar o quão bem o modelo generaliza para novos dados não vistos.
4. Realizar as previsões de dados não rotulados depois que o modelo apresentar bom desempenho

2.7.5.2 Não-supervisionado

No aprendizado não supervisionado (Fig. 2.11), o modelo é treinado em dados que não possuem rótulos de saída. Em vez disso, o objetivo é descobrir estruturas, padrões ou agrupamentos nos dados. Isso inclui técnicas como *clustering* (agrupamento), redução de dimensionalidade e geração de dados sintéticos. O aprendizado não supervisionado é usado em tarefas como segmentação de mercado, análise de dados e muito mais.

Figura 2.11 – Sistema de aprendizado não-supervisionado



Fonte: Adaptado de Lacerda (2023b)

2.7.5.3 Por reforço

O aprendizado por reforço envolve um agente que toma ações em um ambiente e recebe recompensas ou penalidades com base nessas ações. O objetivo do agente é aprender a realizar ações que maximizem as recompensas ao longo do tempo.

Figura 2.12 – Sistema de aprendizado recorrente



Fonte: Adaptado de Lacerda (2023b)

Conforme pode ser observado na Fig. 2.12 o sistema de aprendizado recorrente é composto por:

- **Agente:** É a entidade que está sendo treinada para tomar decisões.
- **Ambiente:** É o contexto no qual o agente opera e no qual suas ações têm impacto.
- **Ações:** O que o agente pode realizar ações dentro do ambiente.
- **Estado:** Descreve as condições atuais do ambiente em um determinado momento. O agente observa o estado para tomar decisões.
- **Recompensa:** É um sinal numérico que o agente recebe após executar uma ação em um determinado estado. O objetivo do agente é maximizar a recompensa cumulativa ao longo do tempo.

O processo de aprendizado por reforço envolve o agente interagindo com o ambiente em episódios repetidos, tomando ações e recebendo recompensas. O agente usa essas interações para aprender uma estratégia ótima, que lhe permite tomar decisões que maximizam a recompensa ao longo do tempo.

2.8 Estatística Descritiva

A estatística descritiva descreve e sintetiza as principais características observadas em um conjunto de dados por meio de tabelas, gráficos e medidas-resumo, permitindo a compreensão do comportamento dos dados (BELFIORE, 2015).

As informações contidas em um conjunto de dados podem ser resumidas por meio de medidas numéricas adequadas, chamadas medidas-resumo. Essas medidas têm como objetivo principal a representação do comportamento da variável por meio de seus valores centrais e não centrais, e suas dispersões ou formas de distribuição dos seus valores em torno da média (BELFIORE, 2015).

As medidas-resumo que serão apresentadas a seguir são: medidas de posição ou localização e medidas de dispersão ou variabilidade.

2.8.1 Medidas de posição ou localização

As medidas de posição ou localização fornecem valores que caracterizam o comportamento de uma série de dados, indicando a posição ou localização dos dados em relação ao eixo dos valores assumidos pela variável ou característica em estudo. Essas medidas são subdivididas em medidas de tendência central (média, mediana e moda) e medidas separatrizes (quartis, decis e percentis) (BELFIORE, 2015).

2.8.1.1 Média aritmética simples

A média aritmética simples é a soma do total de valores de uma determinada variável dividida pelo número total de observações (Eq. 2.26).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.26)$$

2.8.1.2 Média aritmética ponderada

No cálculo da média aritmética simples, todas as ocorrências têm a mesma importância ou peso. Quando se deseja atribuir diferentes pesos (p_i) para cada valor i da variável \bar{x} , utiliza-se a média aritmética ponderada descrita pela Eq. 2.27.

$$\bar{x} = \frac{\sum_{i=1}^n x_i p_i}{\sum_{i=1}^n p_i} \quad (2.27)$$

2.8.1.3 Mediana para dados discretos e contínuos não agrupados

A mediana (Md) é uma medida de localização do centro da distribuição de um conjunto de dados ordenados de forma crescente. Seu valor separa a série

em duas partes iguais, de modo que 50% dos elementos são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana. A mediana da variável \bar{x} pode ser calculada conforme a Eq. 2.28.

$$Md(x) = \begin{cases} \frac{1}{2} \cdot [x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}], & \text{se } n \text{ for par} \\ x_{(\frac{n+1}{2})}, & \text{se } n \text{ for ímpar} \end{cases} \quad (2.28)$$

em que n é o número total de observações e $x_1 \leq \dots \leq x_n$, tal que x_1 é a menor observação ou o valor do primeiro elemento e x_n é a maior observação ou o valor do último elemento.

2.8.1.4 Moda

Considere um conjunto de observações x_1, x_2, \dots, x_n de uma determinada variável. A moda é o valor que aparece com maior frequência.

2.8.2 Medidas separatrizes

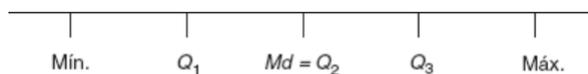
A utilização apenas de medidas de tendência central pode não ser o método mais adequado para analisar o conjunto de dados, uma vez que estes são afetados por valores extremos e assim não é possível ter uma ideia clara da dispersão e da simetria dos dados. Como alternativa, pode-se utilizar as medidas separatrizes, como quartis, decis e percentis. O 2º quartil (Q_2), 5º decil (D_5) ou 50º percentil (P_{50}) correspondem à mediana, sendo, portanto, medidas de tendência central.

2.8.2.1 Quartis

Os quartis ($Q_i, i = 1, 2, 3$) são medidas de posição que dividem um conjunto de dados, ordenados em forma crescente, em quatro partes com dimensões iguais, conforme pode ser visto na Fig. 2.13.

Assim, o 1º Quartil (Q_1 ou 25º percentil) indica que 25% dos dados são inferiores a Q_1 ou que 75% dos dados são superiores a Q_1 . O 2º Quartil (Q_2 ou 5º

Figura 2.13 – Quartis



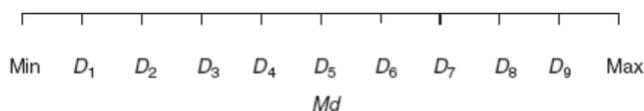
Fonte: Belfiore (2015)

decil ou 50º percentil) corresponde à mediana, indicando que 50% dos dados são inferiores ou superiores a Q_2 . Já o 3º Quartil (Q_3 ou 75º percentil) indica que 75% dos dados são inferiores a Q_3 ou que 25% dos dados são superiores a Q_3 .

2.8.2.2 Decis

Os decis (D_i , $i = 1, 2, \dots, 9$) são medidas de posição que dividem um conjunto de dados, ordenados em forma crescente, em 10 partes iguais conforme demonstrado na Fig. 2.14.

Figura 2.14 – Decis



Fonte: Belfiore (2015)

Dessa forma, o 1º decil (D_1 ou 10º percentil) indica que 10% dos dados são inferiores a D_1 ou que 90% dos dados são superiores a D_1 . O 2º decil (D_2 ou 20º percentil) indica que 20% dos dados são inferiores a D_2 ou que 80% dos dados são superiores a D_2 . E assim sucessivamente, até o 9º decil (D_9 ou 90º percentil), indicando que 90% dos dados são inferiores a D_9 ou que 10% dos dados são superiores a D_9 .

2.8.2.3 Percentis

Os percentis (P_i , $i = 1, 2, \dots, 99$) são medidas de posição que dividem um conjunto de dados, ordenados em forma crescente, em 100 partes iguais. Desta

maneira, o 1º percentil (P_1) indica que 1% dos dados é inferior a P_1 ou que 99% dos dados são superiores a P_1 . O 2º percentil (P_2) indica que 2% dos dados são inferiores a P_2 ou que 98% dos dados são superiores a P_2 . E assim, sucessivamente, até o 99º percentil (P_{99}), indicando que 99% dos dados são inferiores a P_{99} ou que 1% dos dados é superior a P_{99} .

2.8.3 Medidas de dispersão ou variabilidade

As medidas de dispersão mais comuns referem-se à amplitude, ao desvio-médio, à variância, ao desvio-padrão, ao erro-padrão e ao coeficiente de variação (CV).

2.8.3.1 Amplitude

A medida mais simples de variabilidade é a amplitude total, ou simplesmente amplitude (A), que representa a diferença entre o maior e menor valor do conjunto de observações (Eq. 2.29).

$$A = x_{max} - x_{min} \quad (2.29)$$

2.8.3.2 Desvio-médio

O desvio é a diferença entre cada valor observado e a média da variável. Assim, para dados populacionais, seria representado por $(x_i - \mu)$, e para dados amostrais por $(x_i - \bar{x})$. O desvio-médio, ou desvio-médio absoluto, representa a média aritmética dos desvios absolutos (em módulo).

2.8.3.3 Variância

A variância é uma medida de dispersão ou variabilidade que avalia quanto os dados estão dispersos em relação à média aritmética. Assim, quanto maior a variância, maior a dispersão dos dados.

A variância populacional (σ^2) é calculada a partir de todos os elementos de uma população, conforme Eq. 2.31. Ela mede a média dos quadrados das diferenças entre cada valor individual e a média populacional.

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \quad (2.30)$$

onde N é o tamanho da população e μ é a média populacional.

Por outro lado, a variância amostral (S^2) é uma estimativa da variância populacional, calculada a partir de uma amostra representativa da população, conforme a Eq. 2.31.

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (2.31)$$

onde n é o tamanho da amostra e \bar{x} é a média amostral.

A relação entre a variância amostral (S^2) e a variância populacional (σ^2) é dada pela Eq. 2.32.

$$s^2 = \frac{n}{n - 1} \sigma^2 \quad (2.32)$$

2.8.3.4 Desvio-padrão

Como a variância considera a média dos desvios quadrados, ele tende a ser muito grande e de difícil interpretação. Para resolver esse problema, extrai-se a raiz quadrada da variância e essa medida é conhecida como desvio-padrão. É possível obter o desvio padrão populacional e amostral conforme as Eqs. 2.33 e 2.34.

$$\sigma = \sqrt{\sigma^2} \quad (2.33)$$

$$S = \sqrt{S^2} \quad (2.34)$$

2.8.3.5 Erro-padrão

O erro-padrão é o desvio-padrão da média. É obtido dividindo-se o desvio-padrão pela raiz quadrada do tamanho da população ou amostra conforme as Eqs. 2.35 e 2.36.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \quad (2.35)$$

$$S_{\bar{x}} = \frac{S}{\sqrt{N}} \quad (2.36)$$

Quanto maior o número de medições, melhor será a determinação do valor médio (maior precisão), devido à compensação dos erros aleatórios.

2.8.3.6 Coeficiente de variação

O coeficiente de variação (*CV*) é uma medida de dispersão relativa que fornece a variação dos dados em relação à média. Quanto menor for o seu valor, mais homogêneos serão os dados, ou seja, menor será a dispersão em torno da média. Pode ser calculado conforme as Eqs. 2.37 e 2.38.

$$CV = \frac{\sigma}{\mu} \cdot 100(\%) \quad (2.37)$$

$$CV = \frac{S}{\bar{x}} \cdot 100(\%) \quad (2.38)$$

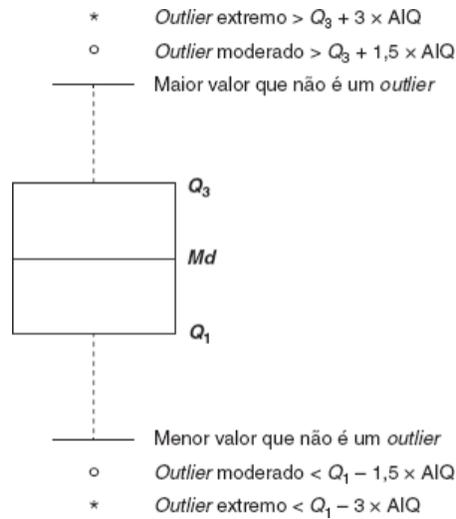
2.8.4 Outliers

Um conjunto de dados pode conter algumas observações que apresentam um grande afastamento das restantes ou são inconsistentes. Essas observações são designadas *outliers*, ou ainda por valores atípicos, discrepantes, anormais ou extremos.

É possível visualizar os *outliers* dos dados por meio do *boxplot* conforme exemplificado na Fig. 2.15. Para gerar esse gráfico primeiro é calculada a ampli-

tude interquartil (AIQ) que corresponde à diferença entre o terceiro e o primeiro quartil, conforme a Eq. 2.39

Figura 2.15 – Boxplot



Fonte: Belfiore (2015)

$$AIQ = Q_3 - Q_1 \quad (2.39)$$

O valor da AIQ é o comprimento da caixa. Quaisquer valores situados abaixo de Q_1 ou acima de Q_3 por mais $1,5 \times AIQ$ são considerados *outliers* moderados e serão representados por círculos, podendo ainda serem aceitos na população com alguma suspeita. Assim, o valor x^0 de uma variável é considerado um *outlier* moderado quando $X^0 < Q_1 - 1,5 \cdot AIQ$ e $X^0 > Q_3 + 1,5 \cdot AIQ$

Ou ainda, quaisquer valores situados abaixo de Q_1 ou acima de Q_3 por mais $3 \times AIQ$ serão considerados *outliers* extremos e serão representados por asteriscos. Assim, o valor x^* de uma variável é considerado um *outlier* extremo quando $X^* < Q_1 - 3 \cdot AIQ$ e $X^* > Q_3 + 3 \cdot AIQ$

2.9 Avaliação de desempenho e métricas de previsão de demanda de energia

A avaliação de desempenho e o uso de métricas adequadas são fundamentais na previsão de demanda de energia para medir a precisão e a eficácia dos modelos utilizados. A seguir será definido como calcular o erro médio absoluto (MAE), erro médio quadrático (MSE), raiz do erro médio quadrático (RMSE), erro percentual absoluto médio (MAPE) e coeficiente de determinação (R^2) conforme descrito por Chicco, Warrens e Jurman (2021).

Chicco, Warrens e Jurman (2021) define X_i como o valor previsto, Y_i como o valor real, m é a quantidade de amostras, a média dos valores reais (\bar{Y}) é definida conforme a Eq. 2.40 e soma total média dos quadrados (MST) é descrita conforme a Eq. 2.41.

$$\bar{Y} = \frac{1}{m} \sum_{i=1}^m Y_i \quad (2.40)$$

$$MST = \frac{1}{m} \sum_{i=1}^m (Y_i - \bar{Y})^2 \quad (2.41)$$

2.9.1 Erro médio absoluto (MAE)

O Erro Médio Absoluto (MAE, do inglês *Mean Absolute Error*) mede a diferença média entre as previsões e os valores reais da demanda de energia, conforme a Eq. 2.42. Esse método fornece uma medida da magnitude média dos erros, ignorando sua direção (se é positivo ou negativo). Sendo assim, normalmente, um valor de MAE mais baixo indica uma melhor precisão do modelo, pois significa que as previsões estão mais próximas dos valores reais em média.

$$MAE = \sum_{i=1}^m \frac{|X_i - Y_i|}{m} \quad (2.42)$$

(Melhor valor = 0; Pior valor = $+\infty$)

2.9.2 Erro médio quadrático (MSE)

O Erro Médio Quadrático (MSE, do inglês *Mean Square Error*) é obtido somando todos os erros quadrados e dividindo pelo número total de observações (Eq. 2.43) e assim mede o quanto, em média, as previsões estão desviando dos valores reais ao quadrado, então quanto menor o valor do MSE, melhor é o desempenho do modelo.

$$MSE = \sum_{i=1}^m \frac{(X_i - Y_i)^2}{m} \quad (2.43)$$

(Melhor valor = 0; Pior valor = $+\infty$)

2.9.3 Erro percentual absoluto médio (MAPE)

O Erro Percentual Absoluto Médio (MAPE, do inglês *Mean Absolute Percentage Error*) é uma métrica utilizada para avaliar a precisão de um modelo em problemas de previsão ou regressão. Ela calcula a média dos erros percentuais absolutos entre as previsões e os valores reais (Eq. 2.44), dessa forma um valor de MAPE mais baixo indica uma melhor precisão do modelo, pois significa que as previsões estão mais próximas dos valores reais, em média, em termos percentuais. No entanto, é importante observar que o MAPE pode ser sensível a valores próximos a zero.

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i - X_i}{Y_i} \right| \quad (2.44)$$

(Melhor valor = 0; Pior valor = $+\infty$)

2.9.4 Coeficiente de determinação (R^2)

O Coeficiente de Determinação (R^2) representa a soma de todas as diferenças percentuais absolutas para todas as amostras (Eq. 2.45). Essa métrica pode

variar entre 0 e 1, sendo que um valor mais próximo de 1 indica um ajuste melhor do modelo aos dados observados.

$$R^2 = 1 - \frac{MSE}{MST} = 1 - \frac{\sum_{i=1}^m (X_i - Y_i)^2}{\sum_{i=1}^m (\bar{Y} - Y_i)^2}, \quad 0 < R^2 < 1 \quad (2.45)$$

(Melhor valor = 1; Pior valor = 0)

2.9.4.1 Coeficiente de determinação ajustado (R^2)

O Coeficiente de Determinação Ajustado (R^2 Ajustado) é uma métrica estatística utilizada para avaliar a qualidade do ajuste de um modelo aos dados observados. Enquanto o R^2 padrão mede a proporção da variabilidade da variável dependente explicada pelo modelo, o R^2 ajustado leva em consideração o número de preditores no modelo, penalizando a inclusão de variáveis irrelevantes. A fórmula para o cálculo do R^2 Ajustado é dada pela Eq. 2.46.

$$R_{\text{ajustado}}^2 = 1 - \frac{(1 - R^2) \cdot (n - 1)}{n - k - 1} \quad (2.46)$$

onde R^2 é o Coeficiente de Determinação padrão, n é o número de observações e k é o número de preditores no modelo.

(Melhor valor = 1; Pior valor = $-\infty$)

2.10 Trabalhos correlacionados

Os estudos de Villamagna (2013), Zeng, Liu e Yu (2019), Divina et al. (2019), Al-bayaty et al. (2019) e Moon et al. (2019) têm como foco comum a previsão de consumo de energia em diferentes contextos, utilizando uma variedade de abordagens e metodologias. Cada um desses estudos apresenta características distintas e resultados específicos que contribuem para a compreensão e o aprimoramento da previsão de consumo de energia.

No estudo comparativo conduzido por Villamagna (2013), que abordou a previsão de consumo de energia elétrica na Universidade Federal de Lavras (UFLA), mostrou que as Redes Neurais Artificiais demonstraram maior eficácia, evidenciando sua capacidade de aprender padrões complexos e lidar com a não linearidade e não estacionariedade nas séries temporais em questão, enquanto a metodologia *Box e Jenkins* se mostrou eficaz na representação da série, capturando tendências e sazonalidades. Portanto, tanto as Redes Neurais Artificiais quanto a abordagem de *Box e Jenkins* são métodos válidos para prever séries temporais de consumo e demanda de energia elétrica, podendo ser utilizados em conjunto para aprimorar a escolha do modelo e, assim, enriquecer ainda mais os trabalhos que dependem de previsões para tomadas de decisão.

O estudo de Zeng, Liu e Yu (2019) realiza uma comparação de quatro métodos de previsão de consumo de energia em edifícios, abrangendo desde modelos matemáticos até técnicas de aprendizado de máquina. O estudo conclui que, em cenários de uso de energia complexos, métodos como *Support Vector Machine* (SVM) e *Multivariate Linear Regression* (MLR) geralmente superam as previsões com RNAs e *Gaussian Process Regression* (GPR), fornecendo resultados mais precisos e adequados.

Por sua vez, Divina et al. (2019) concentra-se na previsão de carga horária de energia elétrica em edifícios, comparando uma variedade de abordagens, incluindo modelos tradicionais e baseados em aprendizado de máquina. Os resultados indicam que métodos de aprendizado de máquina, como *Random Forests*, *Generalized Boosted Regression Model* e *Extreme Gradient Boosting*, tendem a apresentar bom desempenho na previsão de carga horária.

No estudo de Al-bayat et al. (2019), os autores exploram algoritmos de aprendizado de máquina para prever a demanda de energia de curto prazo em uma cidade específica. Eles concluem que entre os quatro modelos considerados, dois algoritmos preditivos, a Rede Neural Artificial (RNA) e as Árvores de Decisão

(DTs), demonstraram as melhores capacidades preditivas, alcançando taxas de erro mais baixas.

Por fim, Moon et al. (2019) concentra-se na comparação de modelos de redes neurais artificiais com diferentes funções de ativação e números de camadas ocultas para prever o consumo de energia em edifícios. O estudo destaca a importância dos hiperparâmetros nas redes neurais, com o modelo baseado em SELU e cinco camadas ocultas demonstrando um desempenho superior.

Em termos de comparação, todos os estudos compartilham o objetivo de aprimorar a previsão de consumo de energia, mas diferem em relação às abordagens e ênfases. Eles destacam a relevância de escolher os métodos certos, a influência dos hiperparâmetros nas redes neurais e a necessidade de modelos eficazes em cenários específicos.

3 METODOLOGIA

A metodologia adotada foi baseada nas etapas definidas no Capítulo 2, sessão 2.3.

3.1 Conjunto de dados

Os dados foram coletados do site¹ da CCEE (Câmara de Comercialização de Energia Elétrica), uma entidade responsável por viabilizar e supervisionar as transações de compra e venda de energia elétrica no Brasil.

3.1.1 Estrutura dos Dados

Os dados baixados são individuais e apresentam os resultados públicos de cada agente. Eles possuem uma visão mensal do consumo de energia elétrica, com registros datados de 2014 a 2022. O arquivo fornecido no formato Excel (.xlsx) possui diversas abas (Veja Ap. A), sendo a tabela ‘Total de consumo por parcela de carga - MWh’ da aba ‘003 Consumo’ a tabela de interesse. As principais colunas desta aba incluem:

- **Mês:** Representa a data de referência do consumo.
- **Sigla do Perfil e Nome Empresarial:** Informações que identificam a distribuidora de energia.
- **Carga:** Nome ou designação da carga.
- **Submercado:** Indica a região do submercado onde o consumo foi registrado.
- **Ramo de Atividade:** Especifica o setor ou indústria ao qual pertence o consumo.

¹ <<https://www.ccee.org.br/web/guest/dados-e-analises/consumo>>, Acesso em 23 de outubro de 2023.

- **Estado:** Designa o estado em que o consumo ocorreu.
- **Consumo no Ambiente Livre da parcela de carga:** Representa o consumo de energia no ambiente livre, medido em MWh.
- **Consumo de energia ajustado da parcela cativa da carga:** Reflete o consumo de energia ajustado para cargas cativas, também em MWh.
- **Consumo de energia ajustado de uma parcela de carga:** Indica o consumo de energia ajustado total, medido em MWh.

Após baixar a planilha e extrair os dados da aba de interesse realizou-se uma inspeção para garantir a integridade e completude dos dados. Nessa etapa foi realizado a exclusão das linhas com dados faltantes na coluna ‘Ramo de Atividade’ e ‘Consumo no Ambiente Livre da parcela de carga’ e foi feito a seleção dos dados de consumo apenas para o estado de Minas Gerais . Em seguidas os dados foram salvos no formato csv.

3.2 Recursos Computacionais

A linguagem de programação escolhida foi o *Python* ², pois é um ferramenta gratuita e versátil amplamente utilizada em tarefas de aprendizado de máquina e análise de dados.

Optou-se por realizar as implementações e análises no *Google Colab*, aproveitando os benefícios da computação em nuvem para colaboração e execução eficiente de código. Além disso, para análise dos dados e implementação das técnicas de previsão foram selecionadas as seguintes bibliotecas:

- **NumPy** ³: é amplamente utilizado para operações numéricas e manipulação de *arrays*. No código, é usado para trabalhar com dados numéricos, realizar

² <<https://www.python.org/>> Acesso em 22 de novembro de 2023.

³ <<https://numpy.org/>>, Acesso em 22 de novembro de 2023.

manipulações matriciais e cálculos necessários, principalmente para ajuste de modelos lineares e manipulação de dados temporais.

- ***Pandas***⁴: desempenha um papel central na manipulação e análise de dados tabulares, ele possibilita a organização eficiente dos conjuntos de dados de demanda de energia, facilitando operações como filtragem, agregação e limpeza dos dados. No código, é utilizado para criar e manipular *DataFrames*, estruturas de dados essenciais para organizar e analisar séries temporais.
- ***Seaborn***⁵: é utilizada para criar visualizações estatísticas atraentes e informativas. No código, destaca-se na criação de gráficos que revelam correlações entre variáveis e *heatmaps* que proporcionam uma análise visual de padrões nos dados.
- ***Matplotlib***⁶: também é uma biblioteca para criação de gráficos. No código, é usado para gerar visualizações das séries temporais e gráficos de comparação entre valores reais e previstos.
- ***TensorFlow***⁷: é uma biblioteca de código aberto para aprendizado de máquina. No código, é usado para construir, treinar e fazer previsões com modelos de redes neurais artificiais.
- ***Scikit-learn***⁸: é empregado para a implementação de algoritmos de aprendizado de máquina. No código, destaca-se no treinamento de modelos de regressão linear, no cálculo de métricas de erro e na divisão dos dados em conjuntos de treino e teste.

⁴ <<https://pandas.pydata.org/>>, Acesso em 22 de novembro de 2023.

⁵ <<https://seaborn.pydata.org/>>, Acesso em 22 de novembro de 2023.

⁶ <<https://matplotlib.org/>>, Acesso em 22 de novembro de 2023.

⁷ <<https://www.tensorflow.org/>>, Acesso em 22 de novembro de 2023.

⁸ <<https://scikit-learn.org/stable/>>, Acesso em 22 de novembro de 2023.

- ***Statsmodels***⁹: é utilizado para modelagem estatística. No código, é usado para ajustar modelos ARIMA e SARIMAX.
- ***SciPy***¹⁰: é uma biblioteca que se integra ao *NumPy* e fornece funcionalidades adicionais para otimização, estatística e álgebra linear. No código, é utilizado para otimizar parâmetros em modelos ARIMA e SARIMAX.
- ***Itertools***¹¹: é uma biblioteca que fornece ferramentas para criar e manipular iteradores. No código, é empregado para criar iteradores para gerar combinações de parâmetros durante a busca em grade. Sua utilização simplifica a implementação de estratégias de otimização.
- ***Holidays***¹²: é uma biblioteca para lidar com feriados. No código, é utilizado para obter feriados brasileiros para consideração na análise de séries temporais.
- ***Calendar***¹³: é essencial para operações relacionadas ao calendário, sendo usada no código para obter o último dia do mês para uma data específica, facilitando manipulações temporais.

Para acessar ao *Google Colab*, utilizou-se um notebook *Dell Inc. Inspiron* 15 equipado com 16GB de RAM, 4 núcleos de processamento, uma capacidade total de armazenamento de 512GB e executando o sistema operacional *Linux Ubuntu* 23.04.

3.3 Análise dos dados

Após a coleta, uma análise inicial foi realizada para compreender a estrutura e características dos dados. Dado que cada ramo de atividade possui um perfil

⁹ <<https://www.statsmodels.org/stable/index.html>>, Acesso em 22 de novembro de 2023.

¹⁰ <<https://www.scipy.org/>>, Acesso em 22 de novembro de 2023.

¹¹ <<https://docs.python.org/3/library/itertools.html>>, Acesso em 22 de novembro de 2023.

¹² <<https://pypi.org/project/holidays/>>, Acesso em 22 de novembro de 2023.

¹³ <<https://docs.python.org/3/library/calendar.html>>, Acesso em 22 de novembro de 2023.

de consumo distinto, para simplificar as comparações subsequentes, decidiu-se focar em um único ramo. Assim, selecionou-se o ramo de atividade: 'Metalurgia e Produtos de Metal', pois ele é o ramo que apresentou o maior consumo total em todos os anos. Além disso, optou-se por analisar os dados referentes ao estado de Minas Gerais (MG), uma vez que as informações de precipitação e temperatura que seriam utilizadas nas análises subsequentes pertencem a esse estado.

3.3.1 Análise Exploratória dos Dados (EDA)

Utilizando a linguagem de programação *Python*² e as bibliotecas descritas na sessão 3.2, realizou-se uma análise exploratória aprofundada dos dados. As principais análises incluíram:

- Perfil de consumo por ramo de atividade
- Decomposição da série em tendência, sazonalidade e resíduo
- Evolução do consumo médio e total ao longo dos anos
- Correlação do consumo com a precipitação, temperatura, preço e variáveis sazonais

Essas análises forneceram *insights* importantes sobre as tendências e fatores que influenciam o consumo de energia.

3.4 Modelagem e Previsão

Inicialmente serão empregadas as técnicas de Regressão, ARIMA e RNA para prever o consumo total em GWh do ano de 2022. Todas as previsões serão realizadas com a linguagem de programação *Python*², utilizando as bibliotecas (descritas em sessão 3.2) com parâmetros e métodos específicos de cada técnica.

3.4.1 Regressão Linear

Para realizar a previsão por regressão linear será utilizada o *LinearRegression* da biblioteca *scikit-learn*⁸ (*sklearn*). Esta técnica é utilizada para realizar regressões lineares simples (Eq. 2.5) e múltiplas (Eq. 2.6) em análise de dados e aprendizado de máquina.

O método *LinearRegression* modela a relação linear entre uma ou mais variáveis independentes, também chamadas de características (por exemplo, temperatura, precipitação, variáveis sazonais entre outros) e uma variável dependente (consumo). A regressão linear assume que essa relação pode ser representada por uma combinação linear das características, onde o objetivo é encontrar os coeficientes que minimizam o erro quadrático médio entre as previsões do modelo e os valores reais.

3.4.1.1 Parâmetros

Os principais parâmetros *LinearRegression()* são:

- ***fit_intercept* (booleano, padrão=*True*):** Este parâmetro indica se o modelo deve calcular o valor do intercepto (coeficiente independente) durante o treinamento. Se for definido como *True*, o modelo calculará o intercepto. Se for definido como *False*, o modelo assumirá que o intercepto é zero.
- ***normalize* (booleano, padrão=*False*):** Quando definido como *True*, o modelo normalizará as variáveis independentes antes de realizar a regressão. Isso significa que as variáveis serão dimensionadas para ter média zero e desvio padrão igual a um. Normalizar as variáveis pode ser útil quando as características estão em escalas diferentes.
- ***copy_X* (booleano, padrão=*True*):** Este parâmetro controla se uma cópia das variáveis independentes deve ser feita antes de realizar a regressão. Quando definido como *True*, os dados originais não são modificados. No

entanto, em situações de grande volume de dados, definir isso como `False` pode economizar memória, desde que os dados não sejam mais utilizados após o treinamento.

- **`n_jobs` (inteiro ou `None`, padrão=`None`):** Especifica o número de núcleos da CPU a serem usados durante o cálculo dos coeficientes. Se definido como `None`, será usado um único núcleo. Isso pode acelerar o processo de treinamento em máquinas com vários núcleos disponíveis.

3.4.1.2 Métodos

Os principais métodos do `LinearRegression()` são:

- **`fit(X, y)`:** Este método é usado para treinar o modelo de regressão linear. `X` representa as variáveis independentes (ou características) e `y` representa a variável dependente (ou alvo).
- **`predict(X)`:** Uma vez que o modelo é treinado, pode-se usar este método para fazer previsões com base em novos dados. `X` é o conjunto de características dos novos dados e o método retornará as previsões correspondentes.
- **`score(X, y)`:** Este método calcula o coeficiente de determinação (R^2) do modelo, que é uma medida da qualidade do ajuste do modelo aos dados de treinamento. Quanto mais próximo de 1, melhor o ajuste.

3.4.2 Média Móvel Integrada Autoregressiva (ARIMA)

Para realizar a previsão pela média móvel integrada autoregressiva será utilizado o comando `ARIMA` para prever o consumo sem influência de variáveis externa e para prever o consumo com as variáveis independentes `SARIMAX`. Ambos comandos são da biblioteca `statsmodels`⁹.

3.4.2.1 ARIMA

O comando ARIMA é uma ferramenta na análise de séries temporais e previsão que combina as três componentes da técnica ARIMA descrita na sessão 2.6.4.

3.4.2.1.1 Parâmetros

Os principais parâmetros do comando **ARIMA()** são:

- **endog** (obrigatório): Representa a série temporal que deseja-se modelar. É uma sequência ou *array-like* contendo os dados da série temporal.
- **order** (obrigatório): Uma tupla contendo três valores: (p, d, q), onde:
 - p (ordem do componente AR - *Autoregressive*): Especifica quantos lags dos próprios valores da série temporal estão incluídos no modelo.
 - d (ordem de diferenciação - *Integrated*): Indica quantas vezes a série temporal deve ser diferenciada para torná-la estacionária.
 - q (ordem do componente MA - *Moving Average*): Especifica quantos lags dos erros anteriores estão incluídos no modelo.
- **textit** (opcional): Permite incluir variáveis exógenas no modelo, o que pode ser útil em alguns casos.
- **trend** (opcional): Permite especificar o componente de tendência do modelo ARIMA. Pode-se escolher entre "c"(constante), "t"(tendência linear), ou "ct"(constante e tendência linear).
- **seasonal**: Permite lidar com componentes sazonais na série temporal. Define a presença ou ausência de componentes sazonais no modelo.

- ***start_params***: Esse parâmetro é usado para definir os valores iniciais dos parâmetros do modelo. Pode ser útil para auxiliar na convergência do algoritmo de otimização durante o ajuste do modelo.

3.4.2.2 Métodos

Os principais métodos do comando **ARIMA()** são:

- ***fit()***: É usado para ajustar o modelo ARIMA aos dados da série temporal especificados no objeto *endog*. O resultado do método *fit()* é um objeto que contém informações sobre o modelo ajustado, incluindo os parâmetros estimados.
- ***forecast(steps)***: É usado para fazer previsões futuras com base no modelo ARIMA ajustado. O argumento *steps* especifica o número de etapas de tempo no futuro para prever. O método retorna as previsões.
- ***predict(start, end)***: Este método permite fazer previsões dentro de um intervalo específico de tempo. Pode-se especificar os índices de início (*start*) e fim (*end*) para as previsões desejadas.
- ***summary()***: Este método fornece um resumo estatístico do modelo ARIMA ajustado, incluindo informações sobre os parâmetros estimados, estatísticas de ajuste, erros-padrão, entre outros.
- ***get_prediction(start, end, dynamic)***: Este método é usado para obter previsões com intervalos de confiança para um intervalo específico de tempo. Pode-se especificar os índices de início (*start*) e fim (*end*) e também usar o argumento *dynamic* para controlar como as previsões são feitas.
- ***forecast(steps, alpha)***: Este método é semelhante ao método *forecast()*, mas permite especificar um nível de significância (*alpha*) para calcular intervalos de confiança em torno das previsões.

3.4.2.3 SARIMAX

O SARIMAX é uma extensão do modelo ARIMA que incorpora componentes sazonais em sua estrutura, o que significa que esse modelo pode acomodar não apenas a autocorrelação e a diferenciação, mas também variáveis exógenas que podem influenciar a série temporal.

3.4.2.3.1 Parâmetros

Os principais parâmetros do comando **SARIMAX()** são:

- **endog (obrigatório):** Este é o parâmetro obrigatório e representa a série temporal que deseja-se modelar.
- **order (p, d, q):** Estes parâmetros são usados para especificar as ordens não sazonais do modelo SARIMA.
 - **p:** Ordem do componente auto-regressivo (AR) não sazonal.
 - **d:** Ordem de diferenciação não sazonal.
 - **q:** Ordem do componente de média móvel (MA) não sazonal.
- **seasonal_order (P, D, Q, S):** Estes parâmetros são usados para especificar as ordens sazonais do modelo SARIMA.
 - **P:** Ordem do componente auto-regressivo sazonal (SAR).
 - **D:** Ordem de diferenciação sazonal.
 - **Q:** Ordem do componente de média móvel sazonal (SMA).
 - **S:** Período sazonal, indicando o número de observações em um ciclo sazonal
- **trend:** Este parâmetro permite especificar o componente de tendência na série temporal. As opções incluem:

- **n**: Nenhum componente de tendência.
 - **c**: Tendência constante.
 - **t**: Tendência linear.
 - **ct**: Tendência constante mais tendência linear.
- **exog**: Este parâmetro permite incluir variáveis exógenas, que são fatores externos que podem influenciar a série temporal. É uma matriz ou DataFrame que contém as variáveis exógenas.
 - **measurement_error**: Um booleano que indica se há um erro de medição na série temporal.
 - **time_varying_regression**: Um booleano que indica se os coeficientes das variáveis exógenas são permitidos para variar com o tempo.
 - **mle_regression**: Um booleano que indica se a regressão é ajustada pelo método de máxima verossimilhança.
 - **simple_differencing**: Um booleano que controla se a diferenciação sazonal deve ser aplicada após a diferenciação não sazonal.
 - **enforce_invertibility**: Um booleano que impõe a invertibilidade do modelo durante o ajuste.
 - **enforce_stationarity**: Um booleano que impõe a estacionariedade do modelo durante o ajuste.

3.4.2.4 Métodos

Os principais métodos do comando **SARIMAX()** são:

- **fit()**: Ajusta o modelo SARIMAX à série temporal fornecida. Ele estima os parâmetros do modelo com base nos dados de entrada.

- ***forecast()***: Este método é usado para fazer previsões futuras com base no modelo SARIMAX ajustado. Permite especificar o número de períodos de previsão desejados.
- ***predict()***: É usado para fazer previsões futuras, mas pode fornecer mais informações, como intervalos de confiança para as previsões.
- ***get_prediction()***: Permite obter previsões e intervalos de confiança para um intervalo de datas específico na série temporal. Ele fornece uma maneira mais flexível de fazer previsões para períodos específicos.
- ***get_forecast()***: É usado para obter previsões e intervalos de confiança, mas geralmente é focado em previsões de longo prazo.
- ***summary()***: Este método fornece um resumo das estatísticas do modelo SARIMAX ajustado, incluindo informações sobre os parâmetros estimados, erros padrão, estatísticas de ajuste e testes diagnósticos.
- ***plot_diagnostics()***: É útil para visualizar os diagnósticos do modelo, como gráficos de resíduos, histograma de resíduos, autocorrelação de resíduos e outros gráficos que ajudam a avaliar a qualidade do modelo ajustado.

3.4.3 Redes Neurais

Para realizar a previsão do consumo com RNA será utilizado *Keras*¹⁴ da biblioteca *TensorFlow*⁷.

3.4.3.1 Parâmetros

Os principais parâmetros do *keras.Sequential()* são:

¹⁴ <<https://keras.io/>>, Acesso em 22 de novembro de 2023.

- **layers (opcional):** Uma lista de camadas a serem adicionadas ao modelo sequencial. Isso permite criar o modelo com todas as camadas desejadas no momento da criação.
- **name (opcional):** Um nome opcional para o modelo. Isso pode ser útil para identificar o modelo quando se trabalha com vários modelos diferentes em um projeto.

3.4.3.2 Métodos

Os principais métodos do *keras.Sequential()* são:

- **add(layer):** Adiciona uma camada ao modelo sequencial. Este método pode ser usado para construir o modelo camada por camada.
- **compile(optimizer, loss, metrics):** Define as configurações de treinamento do modelo. São fornecidos o otimizador, a função de perda e as métricas a serem usadas para treinar e avaliar o modelo.
 - **optimizer:** O otimizador a ser usado para ajustar os pesos da rede durante o treinamento.
 - **loss:** A função de perda a ser otimizada durante o treinamento, que é específica para a tarefa.
 - **metrics:** Uma lista de métricas usadas para avaliar o desempenho do modelo durante o treinamento.
- **fit(x, y, epochs, batch_size, validation_data):** Treina o modelo usando os dados de treinamento.
 - **x:** Refere-se aos dados de entrada que são utilizados para treinar o modelo.

- **y**: Este é o segundo parâmetro e refere-se aos rótulos de saída associados aos dados de entrada x . Os rótulos são as respostas ou valores-alvo que o modelo tentará prever durante o treinamento.
- **epochs**: Este parâmetro define o número de épocas de treinamento. Uma época é uma passagem completa pelos dados de treinamento.
- **batch_size**: Refere-se ao número de exemplos de treinamento usados em cada etapa de atualização dos pesos do modelo. Usar um tamanho de lote maior pode acelerar o treinamento, mas também requer mais memória. O tamanho do lote influencia a estabilidade do treinamento e a convergência do modelo.
- **validation_data (opcional)**: Refere-se ao conjunto de dados de validação, que é usado para avaliar o desempenho do modelo durante o treinamento. O conjunto de validação é útil para monitorar o *overfitting* (ajuste excessivo) e garantir que o modelo está generalizando bem. Ele é usado para calcular métricas de validação, como a função de perda e métricas definidas anteriormente no método *compile*.
- **predict(x)**: Faz previsões usando o modelo treinado. Deve-se fornecer os dados de entrada x , e o método retorna as previsões correspondentes.
- **summary()**: Exibe um resumo da arquitetura do modelo, incluindo o número de parâmetros em cada camada e o número total de parâmetros.

3.5 Comparação Estatística dos Métodos

Após prever o consumo com as técnicas de regressão, ARIMA e RNA será realizada uma comparação estatística entre o valor real e o valor previsto para determinar qual técnica apresentou melhor desempenho na previsão do consumo de energia. Métricas como MAE, MSE, MAPE e R^2 *ajustado* serão utilizadas para avaliar e comparar o desempenho de cada modelo. Na plataforma da linguagem

*Python*² será utilizado o módulo *metrics* da biblioteca *Scikit-Learn*⁸ para obter essas métricas para cada modelo.

3.5.1 Erro Médio Absoluto

Para calcular o MAE será utilizado o comando *mean_absolute_error*¹⁵ que retorna o resultado da Eq. 2.42 que recebe como parâmetros *y_true* (Array ou lista contendo os valores verdadeiros) e *y_pred* (Array ou lista contendo os valores previstos pelo modelo).

3.5.2 Erro Médio Quadrático

Para calcular o MSE será utilizado o comando *mean_squared_error*¹⁶. O calculo é feito conforme a Eq. 2.43 e possui como parâmetros *y_true* (valores verdadeiros) e *y_pred* (valores previstos pelo modelo).

3.5.3 Erro Percentual Absoluto Médio

Para calcular o MAPE conforme a Eq. 2.44 será utilizado o comando *mean_absolute_percentage_error*¹⁷ que recebe como parâmetros *y_true* (valores verdadeiros) e *y_pred* (valores previstos pelo modelo).

3.5.4 Coeficiente de Determinação Ajustado (R^2)

Para calcular o R^2 ajustado será utilizado o comando *r2_score*¹⁸. O calculo é feito conforme a Eq. 2.46 e possui como parâmetros *y_true* (valores verdadeiros) e *y_pred* (valores previstos pelo modelo).

¹⁵ <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_absolute_error.html>

¹⁶ <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_squared_error.html>

¹⁷ <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_absolute_percentage_error.html#sklearn.metrics.mean_absolute_percentage_error>

¹⁸ <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.r2_score.html#sklearn.metrics.r2_score>

4 RESULTADOS PRELIMINARES

Nessa seção contém a análise dos dados de consumo. O objetivo é compreender a dinâmica do consumo ao longo do tempo e identificar se há tendência, sazonalidade, ou se tem alguma variável que influencia na demanda de energia.

Os dados analisados abrangem os anos de 2014 a 2022 e incluem informações sobre o consumo, precipitação, temperatura e preço de 15 ramos de atividade do estado de Minas Gerais.

Inicialmente realizou-se a análise do *boxplot* conforme pode ser visto nas Figs. 4.1 e 4.2, esses gráficos revelam padrões distintos de consumo entre os setores. Nota-se, que o setor de ‘Metalurgia e Produtos de Metal’ possui consumo na ordem de 140GWh, enquanto muitos outros setores não ultrapassam 20GWh.

Gráfico 4.1 – Boxplot do consumo energético (GWh) de 2014

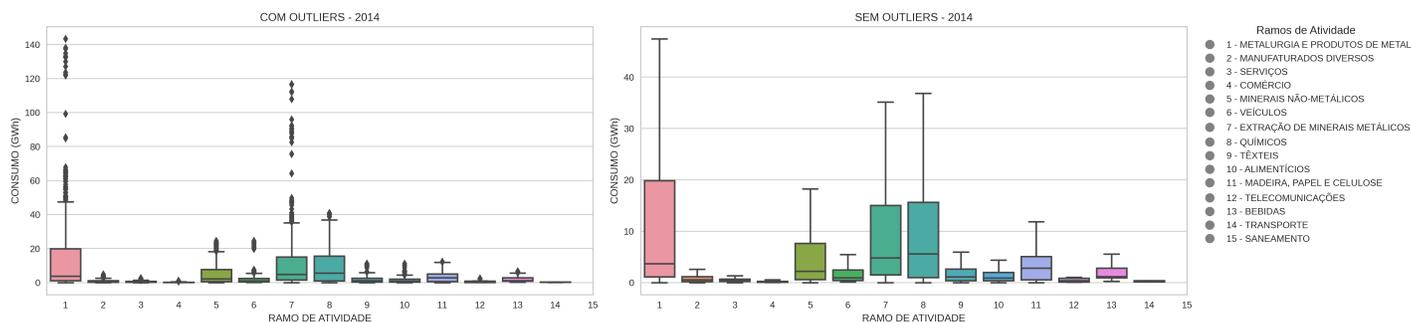
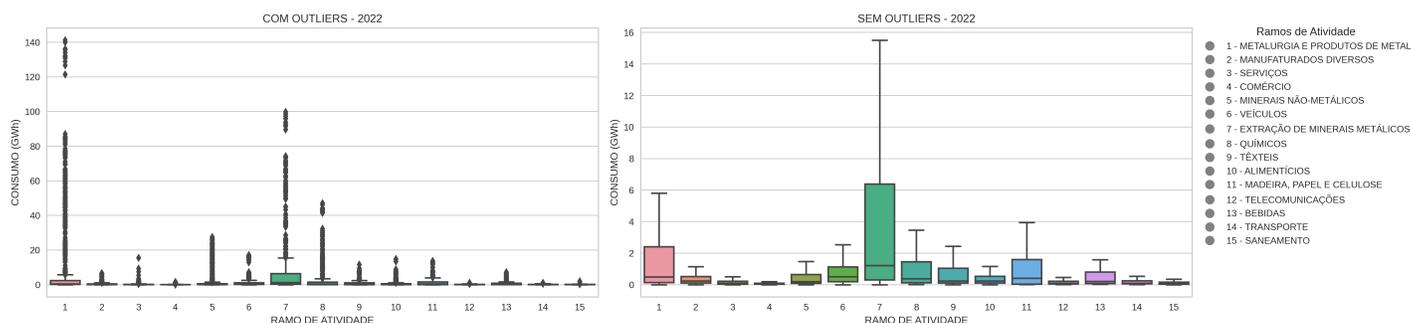


Gráfico 4.2 – Boxplot do consumo energético (GWh) de 2022



Ao comparar o *boxplot* de 2014 (Fig.4.1) com o de 2022 (Fig.4.2) percebe-se que a quantidade de *outliers* aumentou e que mesmo dentro do mesmo ramo houve uma variação no perfil de consumo. Para entender melhor essa variação tem-se as Tabs. 4.1 e 4.2 as Figs. 4.5 e 4.6.

A tabela do consumo médio dos consumidores livres por ramo de atividade (4.1) juntamente com os gráficos 4.3 e 4.4 mostram que para todos os ramos de atividade o consumo médio diminuiu de 2014 para 2022.

Tabela 4.1 – Consumo médio dos consumidores livres por ramo de atividade (2014-2022)

Ramo de Atividade	2014 (MWh)	2015 (MWh)	2016 (MWh)	2017 (MWh)	2018 (MWh)	2019 (MWh)	2020 (MWh)	2021 (MWh)	2022 (MWh)
ALIMENTÍCIOS	1.590,72	1.571,64	1.275,82	987,29	901,31	843,62	731,80	622,61	552,12
BEBIDAS	1.787,07	1.644,98	1.332,68	1.037,57	1.018,56	1.014,96	922,83	879,72	860,65
COMÉRCIO	261,10	245,55	211,14	186,49	172,61	154,64	128,23	115,85	106,49
MINERAIS METÁLICOS	17.012,72	19.310,46	15.684,21	13.467,58	12.968,87	10.779,39	9.219,24	9.640,37	9.283,28
MADEIRA	4.857,66	4.790,52	3.825,31	2.876,05	2.722,02	2.341,43	2.003,48	1.934,63	1.977,30
MANUFATURADOS DIV.	1.427,08	1.330,91	1.011,14	773,52	720,80	655,07	555,75	538,09	495,92
METALURGIA	12.634,07	11.161,80	9.739,72	7.099,85	6.903,63	5.710,32	4.893,21	4.695,70	4.236,88
MINERAIS NÃO-MET.	4.506,62	4.151,39	3.070,13	2.025,83	1.874,14	1.686,42	1.385,35	1.263,03	1.096,41
QUÍMICOS	6.232,75	6.164,20	5.144,01	3.592,98	3.810,88	3.247,91	2.855,37	2.729,16	2.502,65
SANEAMENTO	4.331,01	3.820,72	2.456,59	1.503,79	1.372,32	1.094,13	789,08	502,12	406,22
SERVIÇOS	871,88	851,76	747,55	557,20	467,76	410,78	298,41	259,59	322,08
TELECOMUNICAÇÕES	376,87	379,14	358,70	324,38	304,99	280,73	254,28	220,73	201,50
TRANSPORTE	4.119,03	4.058,51	3.734,57	2.783,81	2.290,69	1.934,32	1.324,14	1.056,67	915,65
TÊXTEIS	1.778,25	1.635,33	1.349,48	1.089,05	1.036,35	1.000,56	833,21	892,06	791,23
VEÍCULOS	2.183,91	1.861,62	1.575,98	1.431,27	1.383,02	1.341,69	998,25	1.060,86	1.031,11

Gráfico 4.3 – Comparativo do consumo energético médio (GWh) de cada ramo de atividade no ano de 2014

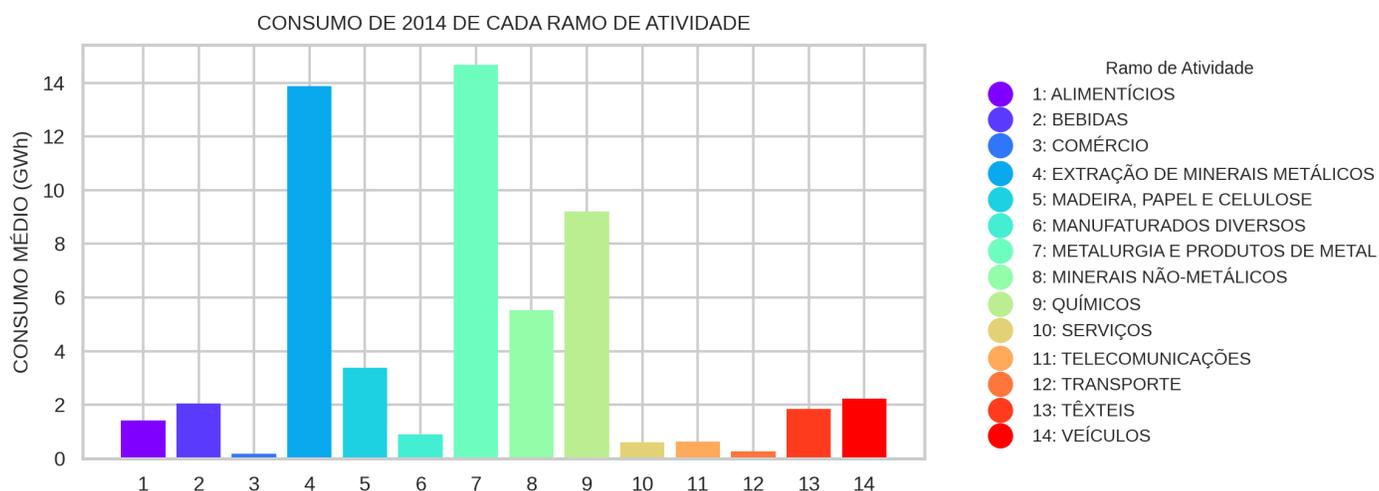
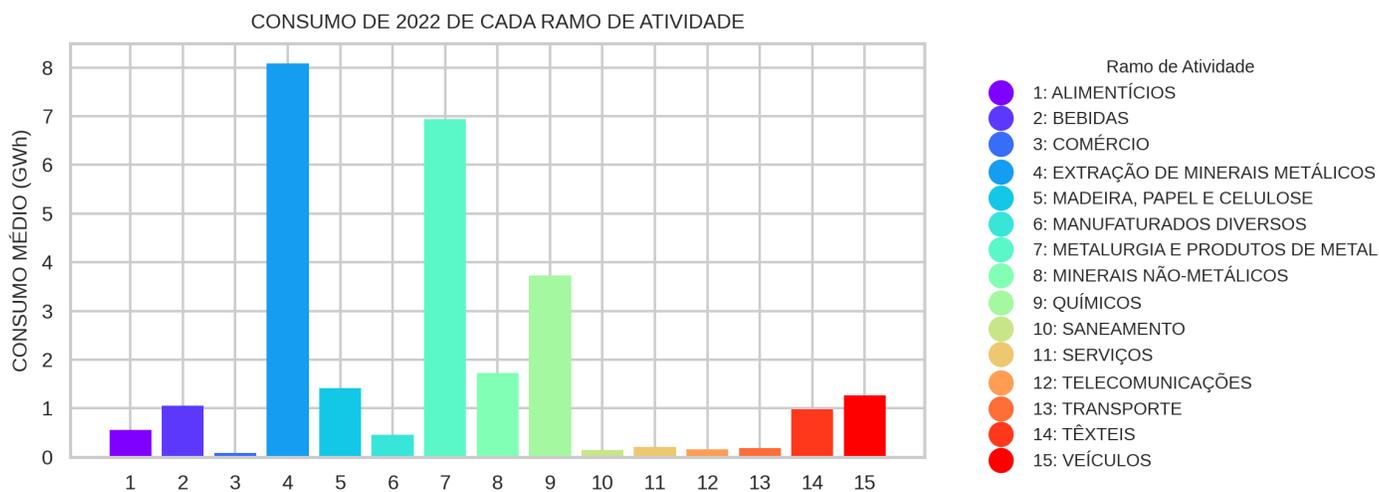


Gráfico 4.4 – Comparativo do consumo energético médio (GWh) de cada ramo de atividade no ano de 2022



Entretanto ao analisar a tabela do consumo total dos consumidores livres por ramo de atividade (4.2) e comparar os gráficos 4.5 e 4.6 tem-se que o consumo total aumentou. Isso pode ser explicado pelas alterações na lei conforme descrito na seção 2.2. Como o consumo exigido para ser um consumidor livre está diminuindo, a quantidade de consumidores de pequeno porte está aumentando fazendo com que a média caia, mas o total aumente.

Tabela 4.2 – Consumo total dos consumidores livres por ramo de atividade (2014-2022)

Ramo de Atividade	2014 (MWh)	2015 (MWh)	2016 (MWh)	2017 (MWh)	2018 (MWh)	2019 (MWh)	2020 (MWh)	2021 (MWh)	2022 (MWh)
ALIMENTÍCIOS	7.177.309,72	7.121.121,34	9.131.064,80	13.841.859,72	14.850.950,38	16.525.681,48	18.238.618,91	20.149.501,88	21.654.333,70
BEBIDAS	1.274.177,86	1.187.676,67	1.399.316,20	1.725.487,03	1.999.437,61	2.290.765,44	2.413.210,60	2.547.670,75	2.732.554,26
COMÉRCIO	2.340.786,32	2.381.305,78	3.497.476,07	6.401.866,52	7.545.312,13	8.245.944,07	9.491.515,45	11.626.644,37	12.983.321,74
MINERAIS METÁLICOS	11.840.850,21	13.440.081,44	11.574.943,66	12.120.817,64	12.644.651,62	12.008.239,36	11.791.414,12	13.101.268,36	13.367.918,77
MADEIRA	8.481.470,37	8.191.788,54	9.372.003,30	10.606.857,79	10.983.358,47	11.142.852,49	11.255.539,60	12.538.365,07	14.139.684,04
MANUFATURADOS DIV.	7.753.350,70	7.240.154,50	8.698.874,50	12.079.319,98	12.904.553,25	13.777.392,94	14.531.744,80	16.551.226,56	17.308.659,25
METALURGIA	40.744.874,82	36.353.982,25	39.231.604,96	40.547.237,98	43.520.505,39	42.610.425,83	43.838.268,19	48.811.813,95	49.846.835,65
MINERAIS NÃO-MET.	13.920.957,25	13.649.768,72	14.260.768,05	15.305.156,42	15.742.768,33	16.747.854,26	17.805.848,42	19.992.427,59	19.939.368,64
QUÍMICOS	15.002.233,02	15.096.127,28	15.766.384,22	16.042.634,11	18.955.302,90	17.389.330,99	17.694.729,50	19.603.562,89	20.163.834,84
SANEAMENTO	1.091.413,47	962.822,50	1.193.904,71	1.849.664,21	1.899.288,86	2.390.683,66	3.107.395,06	4.096.759,47	4.622.012,24
SERVIÇOS	4.743.010,52	4.801.355,29	6.470.078,07	9.811.688,71	10.475.081,40	11.315.708,55	11.036.906,67	13.207.251,19	19.867.341,60
TELECOMUNICAÇÕES	846.072,02	855.347,85	1.006.160,37	1.591.715,09	1.748.188,63	1.901.959,55	1.971.178,21	2.160.731,95	2.250.977,80
TRANSPORTE	1.779.421,24	1.753.275,38	1.721.636,32	1.789.988,59	1.850.881,50	2.259.283,27	2.093.468,21	2.211.619,21	2.309.257,43
TÊXTEIS	4.221.560,98	3.856.105,61	4.242.779,18	5.077.156,09	5.250.159,62	5.374.029,31	5.083.412,35	6.083.827,47	5.852.748,78
VEÍCULOS	5.191.160,61	4.462.309,97	4.813.031,94	6.115.799,51	6.529.230,36	7.172.689,13	6.216.087,80	7.356.027,65	7.517.830,73

Gráfico 4.5 – Comparativo do consumo energético total (GWh) de cada ramo de atividade no ano de 2014

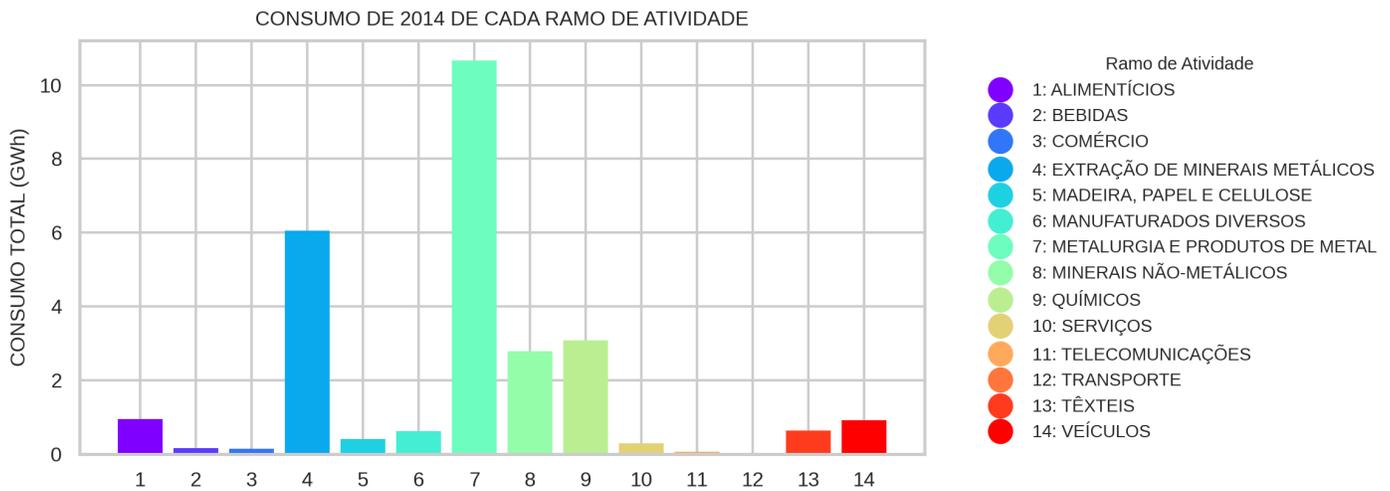
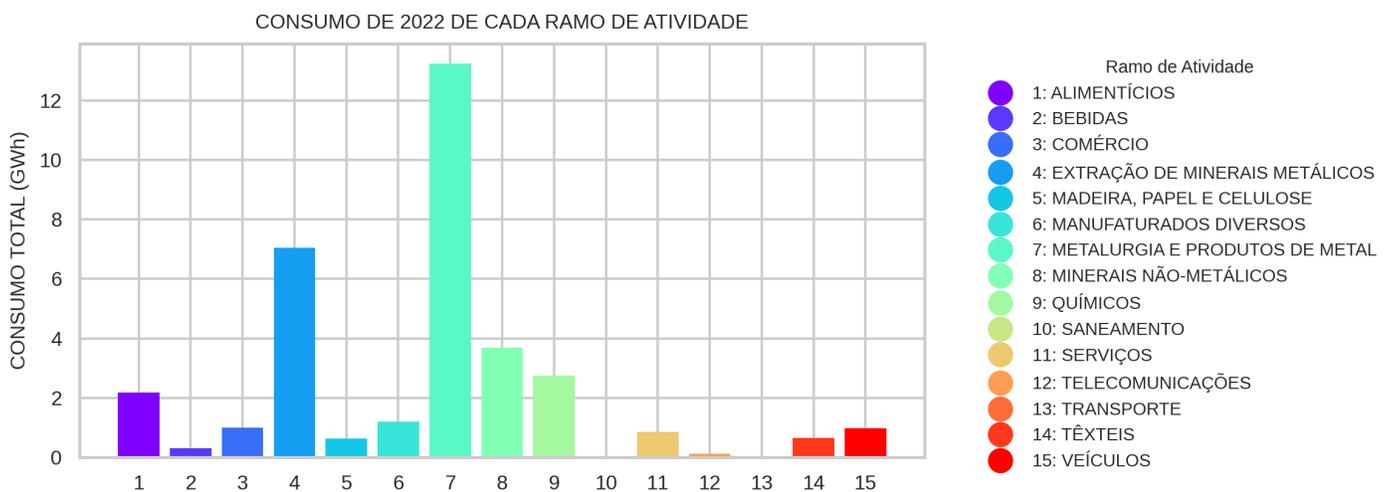


Gráfico 4.6 – Comparativo do consumo energético total (GWh) de cada ramo de atividade no ano de 2022



Dado que cada ramo de atividade apresenta um perfil de consumo distinto, recomenda-se realizar previsões individualizadas para cada um. Para isso será

utilizado os dados do consumo total, uma vez que, conforme observado, a média está em declínio e o interesse é em entender a demanda real de consumo.

De acordo com a tabela 4.3 os ramos de atividade ‘Metalurgia e Produtos de Metal’, ‘Químicos’ e ‘Extração de minerais metálicos’ representam mais de 50% do consumo total em todos os anos desde 2014 até 2022. Sendo o ramo ‘Metalurgia e Produtos de Metal’ maior que ‘Químicos’ e ‘Extração de minerais metálicos’ juntos, então, para fins de simplificação, esse trabalho concentrará no ramo ‘Metalurgia e Produtos de Metal’.

Tabela 4.3 – Porcentagem do consumo total de cada ramo de atividade (2014-2022)

	2014 (%)	2015 (%)	2016 (%)	2017 (%)	2018 (%)	2019 (%)	2020 (%)	2021 (%)	2022 (%)
METALURGIA	32,2	30,0	29,6	26,2	26,1	24,9	24,8	24,4	23,2
QUÍMICOS	11,9	12,4	11,9	10,4	11,4	10,2	10,0	9,8	9,4
MINERAIS NÃO-MET.	11,0	11,2	10,8	9,9	9,4	9,8	10,1	10,0	9,3
MINERAIS MET.	9,4	11,1	8,7	7,8	7,6	7,0	6,7	6,5	6,2
MADEIRA	6,7	6,8	7,1	6,8	6,6	6,5	6,4	6,3	6,6
MANUFATURADOS	6,1	6,0	6,6	7,8	7,7	8,0	8,2	8,3	8,1
ALIMENTÍCIOS	5,7	5,9	6,9	8,9	8,9	9,7	10,3	10,1	10,1
VEÍCULOS	4,1	3,7	3,6	3,9	3,9	4,2	3,5	3,7	3,5
SERVIÇOS	3,8	4,0	4,9	6,3	6,3	6,6	6,3	6,6	9,3
TÊXTEIS	3,3	3,2	3,2	3,3	3,1	3,1	2,9	3,0	2,7
COMERCIO	1,9	2,0	2,6	4,1	4,5	4,8	5,4	5,8	6,1
TRANSPORTE	1,4	1,4	1,3	1,2	1,1	1,3	1,2	1,1	1,1
BEBIDAS	1,0	1,0	1,1	1,1	1,2	1,3	1,4	1,3	1,3
SANEAMENTO	0,9	0,8	0,9	1,2	1,1	1,4	1,8	2,0	2,2
TELECOMUNICAÇÕES	0,7	0,7	0,8	1,0	1,0	1,1	1,1	1,1	1,0

4.1 Metalurgia e Produtos de Metal

Ao observar o gráfico 4.7, é evidente que o número de *outliers* vem crescendo ao longo dos anos e como dito anteriormente, esse fenômeno pode ser atribuído ao aumento de consumidores de pequeno porte.

Dado que *outliers* podem distorcer a análise, uma estratégia eficaz para compreender o padrão de consumo deste setor é considerar, na base de dados, apenas os consumidores classificados como livres desde 2014. A Fig. 4.8 mostra o boxplot do consumo total do ramo de atividade ‘Metalurgia e Produtos de Metal’ depois da filtragem e como esperado a quantidade de *outliers* diminuiu.

Gráfico 4.7 – Boxplot do consumo energético (GWh) do ramo de atividade: ‘Metalurgia e Produtos de Metal’ entre 2014 e 2022

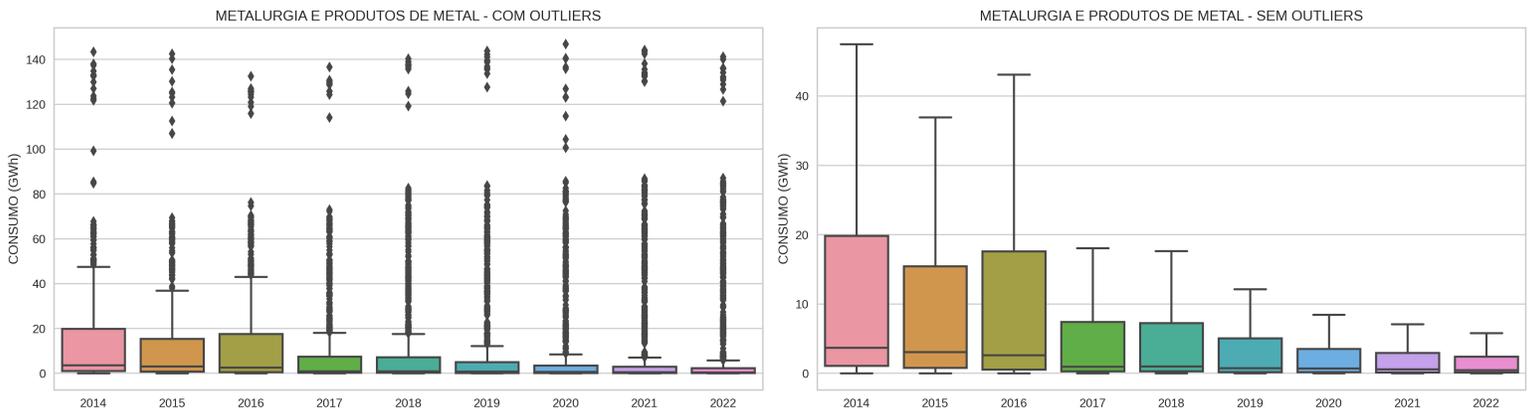
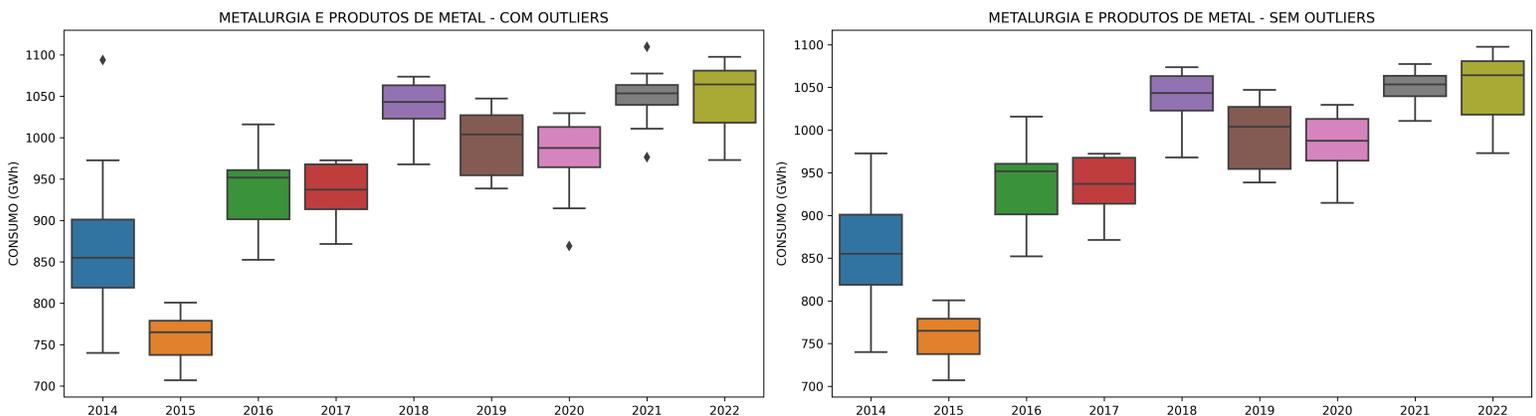


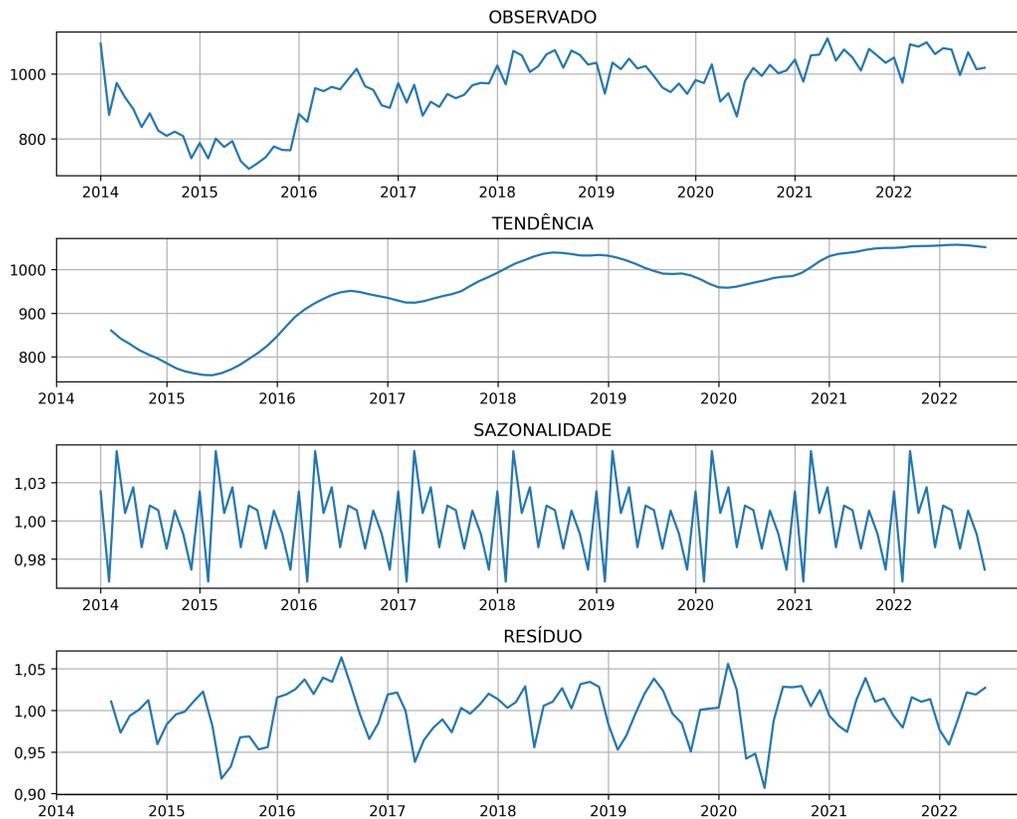
Gráfico 4.8 – Boxplot do consumo energético (GWh) dos consumidores fixos entre 2014 e 2022 do ramo de atividade: ‘Metalurgia e Produtos de Metal’



4.1.1 Decomposição

A série do consumo total (GWh) para o ramo de atividade ‘Metalurgia e Produtos de Metal’ foi decomposta em tendência, sazonalidade e resíduo pelo método multiplicativo conforme pode ser observado na Fig. 4.9.

Gráfico 4.9 – Decomposição do consumo total (GWh) do ramo de atividade: ‘Metalurgia e Produtos de Metal’



4.1.1.1 Tendência

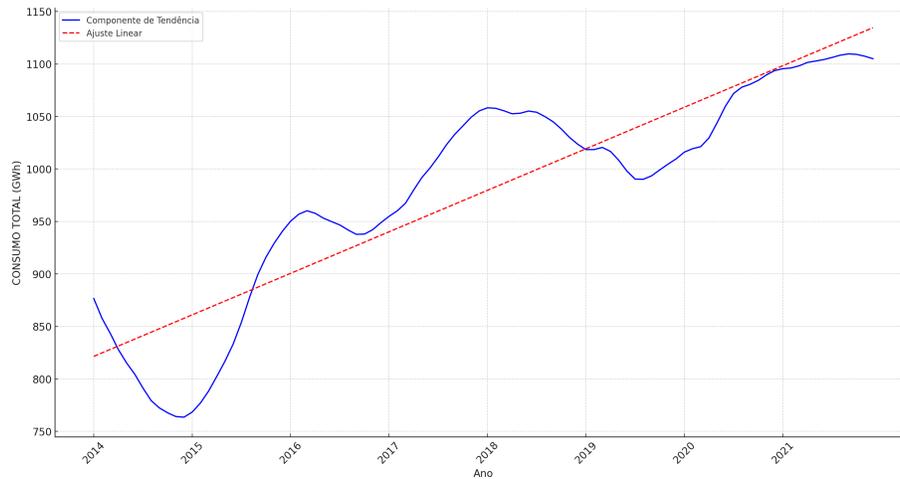
A tendência é uma componente de uma série temporal que indica um padrão geral ou direção ao longo do tempo. A análise revelou uma tendência crescente no consumo (Veja Fig. 4.10). A Eq. 4.1 descreve a tendência.

$$y(t) = 822115,51 + 91,52t \quad (4.1)$$

onde $y(t)$ é o consumo no tempo t .

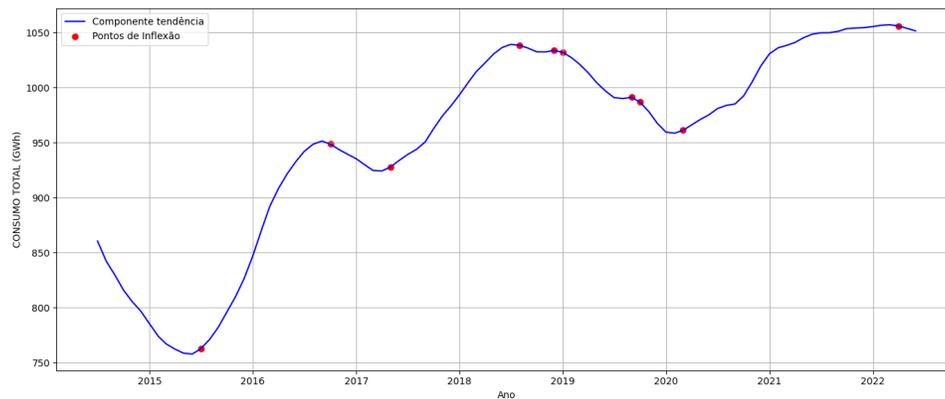
Essa equação indica que há um aumento de aproximadamente 0,091 GWh no consumo a cada mês, partindo de um valor base de 822,115 GWh. Anali-

Gráfico 4.10 – Comparação da tendência real do consumo total (GWh) com da equação que representa essa tendência



sando a inclinação da reta tem-se que o consumo total tem um aumento médio de aproximadamente 0,00959% por mês, ou 0,11506% por ano, porém, ao analisar a segunda derivada da série de tendência tem-se que há vários pontos de inflexão (mudança de sinal, de positivo para negativo ou vice-versa), ou seja, há locais onde taxa de crescimento pode ter diferente da calculada.

Gráfico 4.11 – Pontos de inflexão da tendência do consumo total (GWh)



4.1.1.2 Sazonalidade

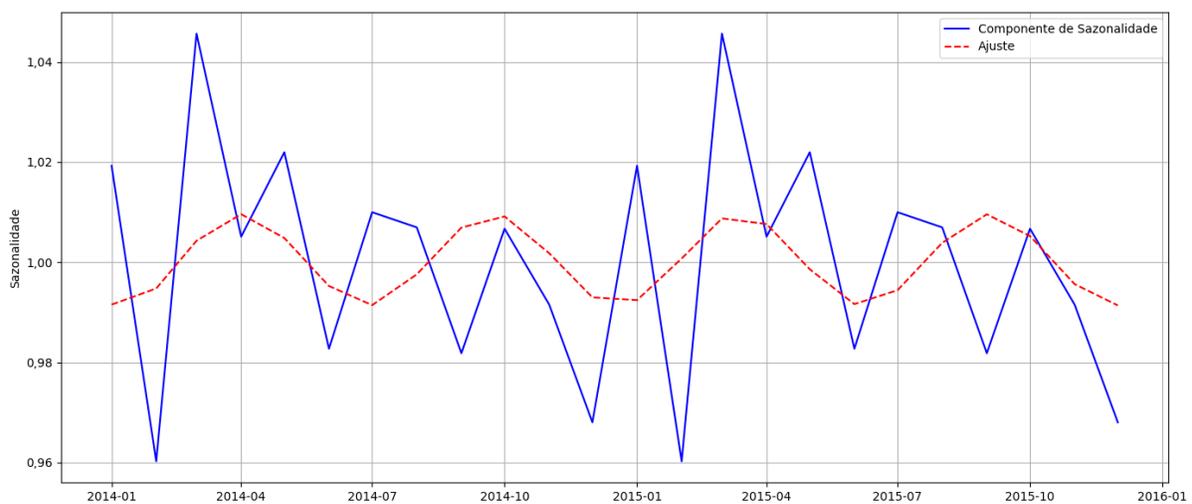
Sazonalidade refere-se a flutuações regulares e previsíveis que ocorrem em um período fixo. No contexto deste conjunto de dados, identificou-se uma sazonalidade senoidal. A Eq. 4.1 descreve a sazonalidade.

$$y(t) = 1,00013 + 0,00972 \sin(1,06265t + 1,43129) - 0,01820 \cos(1,06265t) \quad (4.2)$$

onde $y(t)$ representa o componente sazonal do consumo no tempo t .

Ao observar a Fig. 4.12 nota-se que os meses de março e maio apresentam os valores sazonais mais altos, março com um consumo cerca de 4,57% acima da média e maio com e 2,20% acima. Por outro lado, fevereiro e dezembro revelam níveis de consumo mais baixos. Em fevereiro, o consumo é cerca de 4,0% inferior à média, enquanto dezembro apresenta uma diminuição de cerca de 3,2% no consumo.

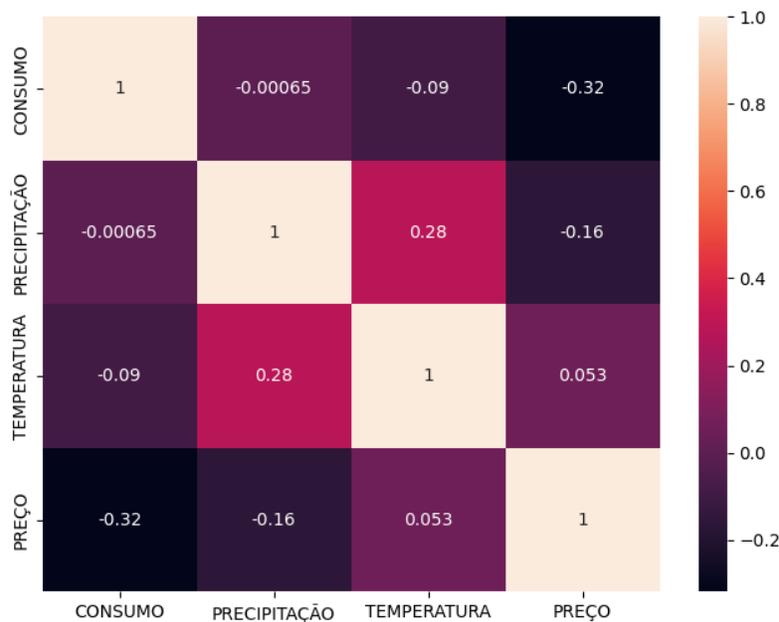
Gráfico 4.12 – Comparação da sazonalidade real do consumo total (GWh) com da equação que representa essa sazonalidade



4.1.2 Correlações

A Fig. 4.13 expõe a matriz de correlação de Pearson das variáveis: consumo, precipitação, temperatura e preço.

Gráfico 4.13 – Matriz de correlações entre consumo, precipitação, temperatura e preço



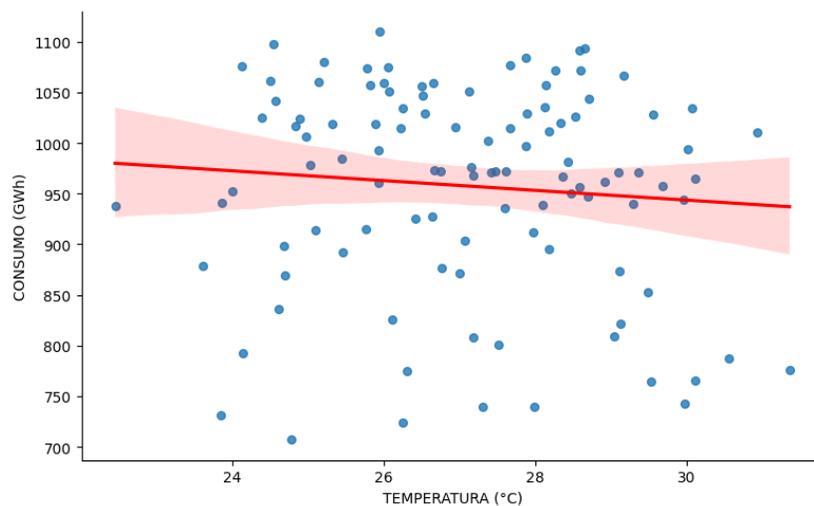
O coeficiente de correlação de Pearson entre "Consumo" e "Temperatura" é aproximadamente -0,09. O sinal negativo indica uma correlação negativa, ou seja, à medida que a temperatura aumenta, o consumo diminui, e vice-versa. O valor absoluto de 0,09 é relativamente baixo, o que sugere uma correlação fraca entre as duas variáveis.

Para uma análise mais detalhada das correlações tem-se as Figs. 4.14, 4.15 e 4.16, esses gráficos são compostos pelos pontos de dispersão que representam o conjunto de dados, pela linha de regressão que indica uma relação positiva entre as variáveis se a inclinação da reta for positiva, relação negativa se a inclinação for negativa e não há correlação se a reta estiver na horizontal. Esse tipo de gráfico é

composto também por uma faixa que mostra a incerteza associada à estimativa da linha de regressão, quanto mais larga essa faixa, maior é a incerteza.

Examinando o gráfico de dispersão entre "Consumo" e "Temperatura" (Fig. 4.14) tem-se a comprovação de que há uma leve tendência descendente, porém a correlação entre essas variáveis não é forte.

Gráfico 4.14 – Correlação entre o consumo total (GWh) do ramo de atividade 'Metalurgia e Produtos de Metal' e a temperatura média no estado de MG



O coeficiente entre "Consumo" e "Precipitação" é $-0,00065$, esse valor é extremamente próximo de 0, o que sugere que não há uma correlação linear significativa entre essas variáveis. O gráfico 4.15 corrobora com essa análise.

Já a correlação entre "Consumo" e "Preço" é moderada e apresenta um coeficiente de $-0,32$. Isso indica que à medida que o preço aumenta, o consumo tende a diminuir (e vice-versa). Essa fato pode ser averiguado no gráfico 4.16.

Gráfico 4.15 – Correlação entre o consumo total (GWh) do ramo de atividade ‘Metalurgia e Produtos de Metal’ e a precipitação média no estado de MG

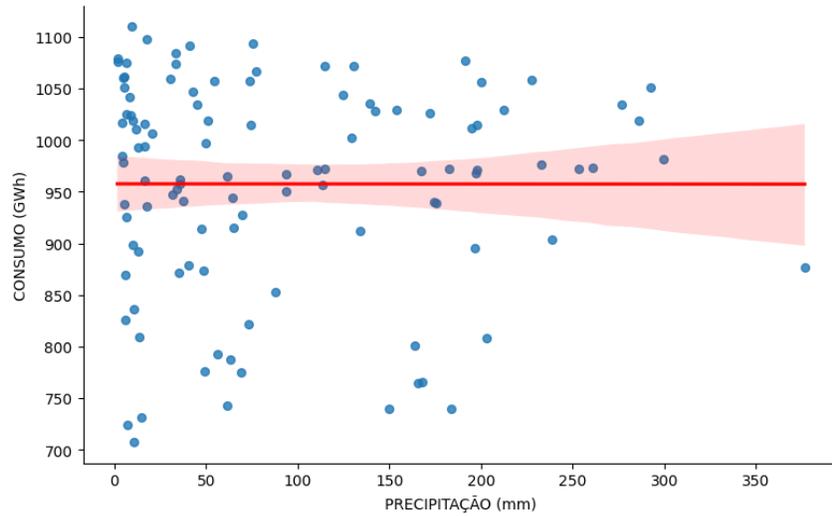
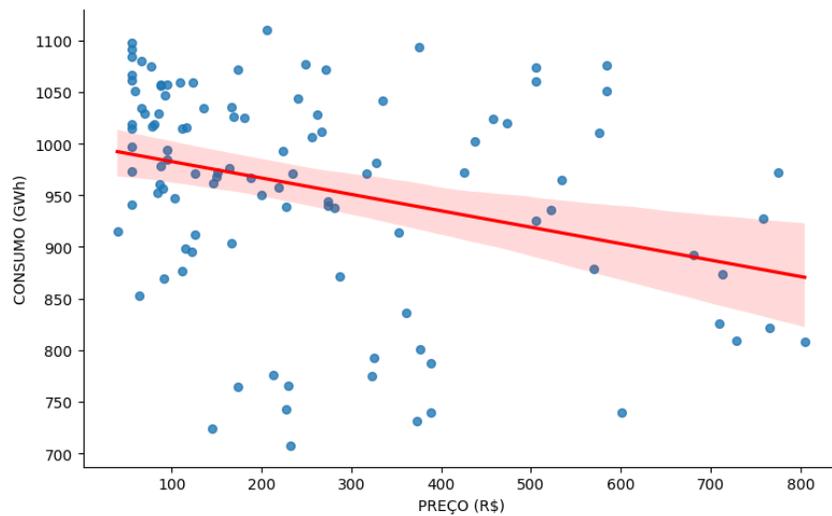


Gráfico 4.16 – Correlação entre o consumo total (GWh) do ramo de atividade ‘Metalurgia e Produtos de Metal’ e o preço médio do submercado Sudeste



5 RECURSOS NECESSÁRIOS

Este capítulo detalha os recursos fundamentais para a realização o estudo comparativo sobre métodos de previsão.

5.1 Hardware

- **Notebook/Computador Pessoal:** É essencial possuir um dispositivo com capacidade de processamento adequada. Recomenda-se um computador com pelo menos 16GB de RAM. Um exemplo é o notebook *Dell Inc. Inspiron 15*, equipado com o processador *Ryzen 5*, que possui 4 núcleos, cada um com 2 *threads*, atingindo uma frequência máxima de 2100 MHz. Além disso, apresenta arquitetura *x86_64 (64-bit)* e uma capacidade total de armazenamento de 512GB.

5.2 Software

- **Linux:** O Ubuntu 23.04 oferece um ambiente estável e confiável para desenvolvimento e pesquisa.
- **Python:** Linguagem de programação versátil e com ampla variedade de bibliotecas para análise e previsão de dados, como *pandas*, *numpy*, *scikit-learn* e *TensorFlow*.
- **Google Colab:** Plataforma interativa que facilita a escrita e execução de códigos *Python*, além de permitir análises em tempo real e visualização de dados.
- **GitHub:** Plataforma de hospedagem de código-fonte que utiliza *Git*. Permite o controle de versão e a colaboração, facilitando o gerenciamento e compartilhamento de projetos de pesquisa.

- **LaTeX:** Sistema de preparação de documentos para apresentação de resultados, criação de relatórios e artigos acadêmicos com alta qualidade tipográfica.

5.3 Outros Recursos

- **Bases de Dados:** Acesso a conjuntos de dados. Os dados necessários são fornecidos gratuitamente pela CCEE.

5.4 Orçamento

A Tab. 5.1 detalha os recursos essenciais adquiridos para a pesquisa em métodos de previsão. Ela apresenta uma lista dos itens, suas descrições, quantidades, preços unitários e o custo total associado a cada recurso. Vale salientar que, todos os recursos já foram obtidos e não há previsão de gastos futuros para a continuidade da pesquisa.

Tabela 5.1 – Orçamento estimado para pesquisa em métodos de previsão.

Item	Descrição	Quant.	Preço Un (R\$)	Preço Total (R\$)
1	Dell Inc. Inspiron 15	1	3.598,00	3.598,00
2	Python	1	0,00	0,00
3	Google Colab	1	0,00	0,00
4	GitHub	1	0,00	0,00
5	LaTeX	1	0,00	0,00
6	Bases de dados	1	0,00	0,00
Total				3.598,00

6 CRONOGRAMA

Nesta seção é apresentado o cronograma detalhado (Tab. 6.1) para concluir a pesquisa do estudo comparativo de métodos de previsão. Ele foi elaborado considerando as diversas fases do projeto, desde a revisão bibliográfica até as etapas que devem ser executadas para obtenção do título de mestre.

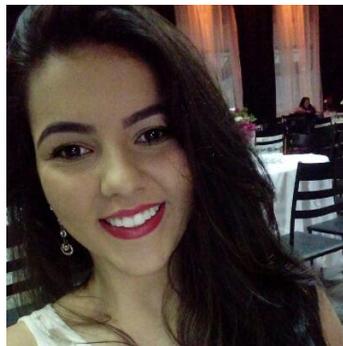
Tabela 6.1 – Cronograma completo

Etapa	2024											
	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1			X	X	X	X						
2				X	X	X	X					
3								X	X	X	X	X
4							X	X	X	X	X	
5												X

Etapa	2025											
	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
6	X	X	X	X	X	X	X	X	X	X	X	
7	X	X	X	X	X	X	X	X	X	X	X	
8												X

1. Cursar 16 créditos em disciplinas da pós-graduação (1º semestre de 2024).
2. Revisão bibliográfica do mestrado.
3. Cursar 16 créditos em disciplinas da pós-graduação (2º semestre de 2024).
4. Implementar os modelos de previsão e gerar resultados.
5. Qualificação.
6. Aprimorar resultados.
7. Redação da Dissertação.
8. Defesa da Dissertação.

7 EQUIPE



Giovanna G. Spuri de Miranda

Universidade Federal de Lavras (UFLA)

ENACOM

giovanna.miranda1@estudante.ufla.br



Técnica em Mecatrônica pelo Centro Federal de Educação Tecnológica de Minas Gerais - CEFET/MG (2018), graduanda em Engenharia de Controle e Automação na pela Universidade Federal Lavras (UFLA). Estagiária na empresa ENACOM



Prof. Dr. Wilian Soares Lacerda

Universidade Federal de Lavras (UFLA)

Departamento de Automática

lacerda@ufla.br



Técnico em Eletrônica pelo Centro Federal de Educação Tecnológica de Minas Gerais - CEFET/MG (1986), possui graduação em Engenharia Elétrica pela Universidade Federal de Minas Gerais (1991), mestrado em Engenharia Elétrica (área Automática) pela Universidade Federal de Minas Gerais (1994) e doutorado em Engenharia Elétrica (linha Engenharia da Computação) pela Universidade Federal de Minas Gerais (2006). Desenvolve pesquisa na área de Sistemas Embarcados e Inteligência Computacional, atuando principalmente nos seguintes temas: Redes Neurais Artificiais, Computação Evolutiva, Hardware Evolutivo, Hardware Reconfigurável. Desenvolve protótipos de sistemas embarcados para aplicações

específicas utilizando: microcontroladores, hardware reconfigurável (FPGA), sensores e atuadores diversos.



Endi Daniel Coelho Silva

ENACOM

Especialista em otimização

endi.silva@enacom.com.br



Cientista da Computação formado pelo IFNMG - Instituto Federal de Minas Gerais Campus Montes Claros. Mestre pela Universidade de São Paulo no campus São Carlos.

8 CONCLUSÃO

Este trabalho teve por objetivo a construção de um projeto de pesquisa voltado para a análise e previsão de demanda de energia. Até o presente momento, realizou-se estudo inicial dos dados disponíveis.

Para a continuação desta pesquisa, é fundamental considerar a análise da base de dados completa, sem excluir nenhum cliente. Essa abordagem permitirá uma visão mais abrangente e realista da demanda de energia, embora exija o desenvolvimento de modelos mais robustos capazes de lidar com valores discrepantes.

É relevante também incorporar análises mais aprofundadas, como a avaliação da taxa de crescimento de novos consumidores ao longo do tempo. A partir desses dados, é possível inferir qual será o valor de consumo dos novos consumidores e somá-lo ao consumo da base de dados congelada. Essa técnica leva em consideração a expansão da base de clientes e oferece uma visão mais precisa da demanda futura, contribuindo para um planejamento mais sólido e abrangente.

Entretanto, é importante ressaltar que a análise da base de dados completa pode ser desafiadora devido à presença de valores discrepantes e flutuações imprevisíveis. Portanto, para a continuação deste estudo, é essencial estudar e desenvolver modelos robustos capazes de lidar com essas complexidades, garantindo que as previsões sejam confiáveis e úteis para os fins de gerenciamento da demanda de energia. Ao considerar a totalidade dos clientes, independentemente de sua presença contínua na base de dados, este projeto de pesquisa pode fornecer uma visão mais completa e precisa da dinâmica da demanda de energia e contribuir para um planejamento energético mais eficaz.

REFERÊNCIAS

ABREU, Y. V. de. **A reestruturação do setor elétrico brasileiro: questões e perspectivas**. Dissertação (Mestrado) — Programa de Pós-graduação em Energia da Universidade de São Paulo (USP), 1999.

AL-BAYATY, H. et al. City scale energy demand forecasting using machine learning based models: a comparative study. In: **Proceedings of the Second International Conference on Data Science, E-Learning and Information Systems**. [S.l.: s.n.], 2019. p. 1–9.

ARMSTRONG, J. S. **Principles of forecasting: a handbook for researchers and practitioners**. [S.l.]: Springer, 2001. v. 30.

BARDELIN, C. E. A. **Os efeitos do racionamento de energia elétrica ocorrido no Brasil em 2001 e 2002 com ênfase no consumo de energia elétrica**. Dissertação (Mestrado) — Escola Politécnica, 2004.

BARROS, B. F. D.; GEDRA, R. L.; BORELLI, R. **Geração, transmissão, distribuição e consumo de energia elétrica**. [S.l.]: Saraiva Educação SA, 2014.

BELFIORE, P. **Estatística - Aplicada à Administração, Contabilidade e Economia com Excel e SPSS**. E-book. [S.l.]: Grupo GEN, 2015. ISBN 9788595155596.

BEZERRA, F. N. R. **Sustentabilidade da matriz energética brasileira**. Dissertação (Mestrado) — Programa de Pós-Graduação em Economia Rural da Universidade Federal do Ceará (UFC), Fortaleza-CE, 2016.

BISHOP, C. M. **Neural Networks for Pattern Recognition**. New York: Oxford University Press, 1997. 482 p.

BOX, G. E. et al. **Time Series Analysis: Forecasting and Control**. [S.l.]: John Wiley & Sons, 2015.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2. ed. Rio de Janeiro: Livros Técnicos e Científicos, 2007. 238 p.

Brasil. **Lei nº 9.074, de 7 de julho de 1995**. 1995. Diário Oficial da União. Acesso em: 10 set. 2023. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/19074cons.htm>.

BUENO, R. D. L. d. S. **Econometria de Séries Temporais - 2ª edição revista e atualizada**. Cengage Learning Brasil, 2018. E-book. ISBN 9788522128259. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788522128259/>>.

CHATFIELD, C. **The Analysis of Time Series: An Introduction**. [S.l.]: Chapman and Hall/CRC, 2016.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. **PeerJ Computer Science**, PeerJ Inc., v. 7, p. e623, 2021.

COSSI, A. M. **Planejamento de redes de distribuição de energia elétrica de média e baixa tensão**. Tese (Doutorado) — Programa de Pós-graduação em Energia da Universidade Estadual Paulista (UNESP), 2008.

DIVINA, F. et al. A comparative study of time series forecasting methods for short term electric energy consumption prediction in smart buildings. **Energies**, MDPI, v. 12, n. 10, p. 1934, 2019.

FERREIRA, R. V. **Previsão de demanda: um estudo de caso para o sistema interligado nacional**. Dissertação (Mestrado) — Programa de Pós-Graduação em Engenharia Elétrica da Universidade Feral de Minas Gerais (UFMG), 2006.

GLOBO, O. **Relembre outros apagões que marcaram o Brasil**. 2023. Acessado em: 14 de setembro de 2023. Disponível em: <<https://oglobo.globo.com/brasil/noticia/2023/08/15/relembre-outros-apagoes-que-marcaram-o-brasil.ghtml>>.

HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2001.

LACERDA, W. S. **Arquiteturas das RNA's**. 2023. Slide de aula. Disciplina: Introdução a Redes Neurais Artificiais, Universidade Federal de Lavras.

LACERDA, W. S. **Treinamento MLP**. 2023. Slide de aula. Disciplina: Introdução a Redes Neurais Artificiais, Universidade Federal de Lavras.

MARTINHO, E. **Distúrbios da energia elétrica**. [S.l.]: Saraiva Educação SA, 2009.

MOON, J. et al. A comparative analysis of artificial neural network architectures for building energy consumption forecasting. **International Journal of Distributed Sensor Networks**, SAGE Publications Sage UK: London, England, v. 15, n. 9, p. 1550147719877616, 2019.

MORETTIN, P. A.; TOLOI, C. M. **Análise de séries temporais: modelos lineares univariados**. [S.l.]: Editora Blucher, 2018.

NETO, A. S. **Estatística e introdução à econometria - 2ª edição**. Editora Saraiva, 2013. E-book. ISBN 9788502199835. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788502199835/>>.

NEVES, C. d.; ROSSI, J. W. **Econometria e Séries Temporais com Aplicações à Dados da Economia Brasileira**. Grupo GEN, 2014. E-book. ISBN 978-85-216-2685-5. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/978-85-216-2685-5/>>.

OLIVEIRA, I. C. D. et al. **Geração de Energia Elétrica**. [S.l.]: E-book, 2021.

SCHWEITZER, A. **O racionamento de energia elétrica no Brasil no período 2001-2002**. Dissertação (Mestrado) — Graduação em Ciências Econômicas da Universidade Federal de Santa Catarina, Florianópolis - SC, 2003. Monografia.

SOUSA, A. R. S. et al. **Análise de séries temporais**. [S.l.]: Grupo A, 2021. ISBN 9786556902876.

SOUZA, T. M. de. **Previsão de preços e demandas de produtos do varejo utilizando técnicas de aprendizado de máquina**. 106 p. Dissertação (Mestrado) — Universidade Federal de Lavras, Lavras, MG, 2021. Acesso em: 19 set. 2023. Disponível em: <http://repositorio.ufla.br/bitstream/1/46869/1/DISSERTA%2087%2083O_Previs%20a%20de%20pre%20a%20os%20e%20demandas%20de%20produtos%20do%20varejo%20utilizando%20t%20c%20a%20n%20c%20de%20aprendizado%20de%20m%20a%20l%20quina.pdf>.

SPIEGEL, M. R.; STEPHENS, L. J. **Estatística**. Grupo A, 2009. E-book. ISBN 9788577805204. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788577805204/>>.

Subsecretaria de Planejamento, Acompanhamento e Gestão. **Portaria nº 514, de 16 de novembro de 2018**. 2018. Diário Oficial do Distrito Federal. Acesso em: 19 set. 2023. Disponível em: <https://www.sinj.df.gov.br/sinj/Norma/075b64fe392842daae14a8c7d79f119c/Portaria_514_16_11_2018.html>.

VILLAMAGNA, M. R. **Seleção de Modelos de Séries Temporais e Redes Neurais Artificiais na Previsão de Consumo e Demanda de Energia Elétrica**. 2013. 74 f. 118 p. Dissertação (Mestrado) — Programa de Pós-graduação em Estatística e Experimentação Agropecuária da Universidade Federal de Lavras, Lavras, MG, 2013. Acesso em: 05 nov. 2023. Disponível em: <http://repositorio.ufla.br/bitstream/1/1117/1/DISSERTA%2087%2083O_Sele%20a%20de%20modelos%20de%20s%20a%20ries%20temporais%20e%20redes%20neurais%20artificiais%20na%20previs%20a%20de%20consumo%20e%20demanda%20de%20energia%20el%20trica.pdf>.

ZENG, A.; LIU, S.; YU, Y. Comparative study of data driven methods in building electricity use prediction. **Energy and Buildings**, Elsevier, v. 194, p. 289–300, 2019.

APÊNDICE A – TABELA DE DADOS DO CONSUMO POR PARCELA DE CARGA - MWH

Mês	Cód. Parâ	Sigla	Nome Empresarial	Cód. Carga	Carga	CNPJ da Carga	Cidade	Estado	Ramo de Atividade	Sobrecarga	Data de Medição	Cód. Desc	Sig. Desc.	Quantidade de Carga	Consumo em Ambiente	Consumo da parcela cabine	Consumo de uma parcela
Jan/20	36	VZM	NEXA S.A.	48	VZM21UZ.FORCA.MG	421668100883	JUÍZ DE FORA	MG	MINÉRIAS E PRODUTOS DE METAIS	SUDESTE	01/08/2001	1141	CEMIG DISTRIB	51,790	27104,034	-	27104,034
Jan/20	36	VZM	NEXA S.A.	7707	VZM2VZ.ZANTEM.MG	421666100964	-	MG	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/01/2005	1141	CEMIG DISTRIB	21,670	22.946,842	-	22.946,842
Jan/20	36	VZM	NEXA S.A.	925823	VZM2AG.PARACATU	4216661001421	ARRIPIANA	MG	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/01/2005	1141	CEMIG DISTRIB	7,500	5.795,402	-	5.795,402
Jan/20	39	CVBD	VALE S.A.	144	VALE.FABRINA	3359251001649	FABRINA	MG	SERVÇOS	SUDESTE	01/09/2019	20	ENERGISA MT	429,854	-	-	429,854
Jan/20	39	CVBD	VALE S.A.	7464	AGUA LIMPA MORLA	3359251001349	RIO BRACCCABA	MG	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/02/2004	1140	CEMIG CERCADO	110,200	53.890,473	-	53.890,473
Jan/20	39	CVBD	VALE S.A.	7383	FABRICA NOVA MARAFANA	33592510041268	MARAFANA	MG	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/02/2005	1141	CEMIG DISTRIB	8,200	4416,710	-	4416,710
Jan/20	39	CVBD	VALE S.A.	70552	VALE S S05SEGO	3359251000901	-	MG	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/01/2008	48	ELERONORTE	61,500	31.130,359	-	31.130,359
Jan/20	39	CVBD	VALE S.A.	70552	VALE CARBONS	3359251000704	PARACAMERAS	PA	EXTR.AÇADO DE MINÉRIAS METÁLICAS	NORTE	01/01/2008	48	ELERONORTE	73,800	39.486,668	-	39.486,668
Jan/20	39	CVBD	VALE S.A.	70631	MINER.AÇO ONSCA RYMA	33592510007980	ORILINDA DO NORTE	ES	EXTR.AÇADO DE MINÉRIAS METÁLICAS	NORTE	01/08/2008	7433	ELERONORTE	125,000	38.665,556	-	38.665,556
Jan/20	39	CVBD	VALE S.A.	71844	VALE BELCOTISMA 8	33592510002106	VITRRIA	ES	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/02/2012	55	ESCELSA	48,000	22.346,319	-	22.346,319
Jan/20	39	CVBD	VALE S.A.	918071	VALE S01 LMS	33592510043473	VIÇOSA	MG	EXTR.AÇADO DE MINÉRIAS METÁLICAS	SUDESTE	01/12/2013	1140	CEMIG CERCADO	100,000	40.720,839	-	40.720,839
Jan/20	100	ALCOA	ALCOA S.A.	918072	ALUMIN. BRAS. ALCOA	33592510043473	SOULIS	MA	EXTR.AÇADO DE MINÉRIAS METÁLICAS	NORTE	01/09/2018	48	ELERONORTE	33,000	10.723,639	-	10.723,639
Jan/20	100	ALCOA	ALCOA S.A.	70162	ALUMIN. BRAS. ALCOA	33592510007016	SOULIS	MA	EXTR.AÇADO DE MINÉRIAS METÁLICAS	NORTE	01/02/2008	48	ELERONORTE	291,210	2.912,106	-	2.912,106
Jan/20	100	ALCOA	ALCOA S.A.	70163	ALUMIN. BRAS. ALCOA	33592510007016	SOULIS	MA	EXTR.AÇADO DE MINÉRIAS METÁLICAS	NORTE	01/02/2008	48	ELERONORTE	820,000	1.421,116	-	1.421,116
Jan/20	104	CEB	INTERCIMENT S.A.	188	CEB-DACIANG	62238884000212	DACI	MG	MINÉRIAS NO-METÁLICOS	SUDESTE	01/09/2002	1141	CEMIG DISTRIB	33,669	17.206,396	-	17.206,396
Jan/20	104	CEB	INTERCIMENT S.A.	188	CEB-DACIANG	62238884000212	DACI	MG	MINÉRIAS NO-METÁLICOS	SUDESTE	01/01/2003	1141	CEMIG DISTRIB	22,700	13.424,437	-	13.424,437
Jan/20	104	CEB	INTERCIMENT S.A.	712	CEB-RODOOJE.NA.MS	62238884000184	RODOQUINA	MS	MINÉRIAS NO-METÁLICOS	SUDESTE	01/01/2003	46	ENERGISA MS	12,000	4.516,849	-	4.516,849
Jan/20	104	CEB	INTERCIMENT S.A.	712	CEB-RODOOJE.NA.MS	62238884000184	RODOQUINA	MS	MINÉRIAS NO-METÁLICOS	SUDESTE	01/01/2004	46	ENERGISA MS	12,200	4.521,430	-	4.521,430
Jan/20	104	CEB	INTERCIMENT S.A.	7843	CEB-SANTANA	62238884001928	SANTANA	RS	MINÉRIAS NO-METÁLICOS	SUDESTE	01/05/2005	3	RGE SUL	7,000	2.006,384	-	2.006,384
Jan/20	104	CEB	INTERCIMENT S.A.	7843	CEB-SANTANA	62238884001928	SANTANA	RS	MINÉRIAS NO-METÁLICOS	SUDESTE	01/05/2005	3	RGE SUL	7,000	2.006,384	-	2.006,384
Jan/20	104	CEB	INTERCIMENT S.A.	71069	CEB-SANTANA DO PAR	62238884002775	SANTANA DO PARAISSO	BA	MINÉRIAS NO-METÁLICOS	NORDESTE	01/01/2012	1141	CEMIG DISTRIB	4,600	1.827,197	-	1.827,197
Jan/20	104	CEB	INTERCIMENT S.A.	71780	CEB-SANTANA DO PAR	62238884004510	CAMPO FORMOSO	BA	MINÉRIAS NO-METÁLICOS	NORDESTE	01/01/2012	10	CELULA	18,500	9.171,414	-	9.171,414
Jan/20	104	CEB	INTERCIMENT S.A.	889255	CEB-SAO MIGUEL DOS	62238884014196	SÃO MIGUEL DOS CAMBOS	AL	MINÉRIAS NO-METÁLICOS	SUL	01/02/2013	1700	CELULA	100,000	4.501,599	-	4.501,599
Jan/20	104	CEB	INTERCIMENT S.A.	889256	CEB-SANTANA	62238884012233	CANAVIA	MS	MINÉRIAS NO-METÁLICOS	SUL	01/02/2013	1700	CELULA	100,000	22.107,041	-	22.107,041