

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

CAIO CAETANO PIMENTEL

**PREVISÃO DE GERAÇÃO HIDRÁULICA NO BRASIL: UM ESTUDO DE CASO
USANDO REDES NEURAS ARTIFICIAIS E REGRESSÃO LINEAR**

MEDIANEIRA

2023

CAIO CAETANO PIMENTEL

**PREVISÃO DE GERAÇÃO DE ENERGIA HIDRÁULICA NO BRASIL: UM ESTUDO
DE CASO USANDO REDES NEURAS ARTIFICIAIS E REGRESSÃO LINEAR**

**Hydraulic power generation forecast in Brazil: a case study using neural
artificial networks and linear regression**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título
de Bacharel em Engenharia Elétrica da Universidade
Tecnológica Federal do Paraná (UTFPR).

Orientador: José Ailton Azevedo dos Santos.

MEDIANEIRA

2023



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

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

CAIO CAETANO PIMENTEL

**PREVISÃO DE GERAÇÃO DE ENERGIA HIDRÁULICA NO BRASIL: UM ESTUDO
DE CASO USANDO REDES NEURAS ARTIFICIAIS E REGRESSÃO LINEAR**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título
de Bacharel em Engenharia Elétrica da Universidade
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 16/novembro/2023

José Airton Azevedo dos Santos
Doutorado
Universidade Tecnológica Federal do Paraná

Filipe Marangoni
Doutorado
Universidade Tecnológica Federal do Paraná

Leandro Antonio Pasa
Doutorado
Universidade Tecnológica Federal do Paraná

MEDIANEIRA

2023

Dedico este trabalho ao meu avô Pedro Paulo Pimentel (in memoriam), maior exemplo de um engenheiro, íntegro e ético, e que sempre me apoiou nos meus projetos.

AGRADECIMENTOS

A Deus, pela minha vida, e por me permitir ultrapassar todos os obstáculos encontrados ao longo da realização deste trabalho.

A esta universidade, seu corpo docente, direção e administração que oportunizaram a janela que hoje vislumbro um horizonte superior, eivado pela acendrada confiança no mérito e ética aqui presentes.

Aos meus avós paternos, pelo amor, incentivo e apoio incondicional.

Não poderia deixar de mencionar a importância da Luana em minha vida e em minha trajetória acadêmica e pessoal. Seu amor, apoio e incentivo foram imprescindíveis para que eu pudesse enfrentar os obstáculos da graduação e seguir em frente com ela.

Ao professor José Airton, pelos ensinamentos e por ter desempenhado tal função com dedicação e amizade.

Por fim, mas não menos importante, aos meus colegas de classe e de estudo, que foram meus companheiros de jornada durante esses últimos anos de graduação e me ajudaram a manter o ânimo e a perseverança em momentos difíceis. Sem a cooperação de vocês, não teria sido possível chegar até aqui.

RESUMO

A energia elétrica é fundamental para o desenvolvimento econômico de um país. É justamente por essa razão que o setor de energia elétrica é um dos setores que mais se expande no país. O Brasil possui uma das matrizes energéticas mais diversificadas do mundo, caracterizada por um sistema hidrotérmico e pela predominância de fontes renováveis de energia. A energia hidráulica (ou hídrica) é uma das opções mais utilizadas no mundo para geração de energia elétrica. Em 2021, no Brasil, a oferta de energia elétrica por fonte correspondeu a 56,8% hidráulica, 12,8% gás natural, 8,2% biomassa, 3% derivados de petróleo, 3,9% de carvão e derivados, 10,6% eólica, 2,2% nuclear e 2,47% solar. Desta forma, a diversidade da matriz energética de cada país implica em diferentes estratégias de planejamento energético, uma vez que cada fonte de energia possui suas limitações. As previsões, capacidade de avaliar o histórico de geração de energia elétrica e projetar uma resposta, podem beneficiar o setor elétrico. São instrumentos de grande importância no auxílio ao processo de tomada de decisão, fornecem informações que proporcionam economia e segurança no fornecimento de energia elétrica. Neste contexto, este trabalho teve como objetivo comparar o desempenho de modelos, de Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais (*RNA*) do tipo *Multilayer Perceptron* (*MLP*) do software *WEKA*, na previsão de geração hidráulica no Brasil. Para realizar esta comparação foi utilizado um conjunto de dados de geração hidráulica. A base de dados apresentou uma série histórica, de geração de energia hidráulica, no período de janeiro/1996 a agosto/2022. A metodologia consistiu em desenvolver modelos de previsão, baseados em algoritmos de Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais (*RNA*) do tipo *Multilayer Perceptron* (*MLP*), utilizando para isso o software *WEKA*. Os resultados mostraram que, para um horizonte de 6 meses, o modelo *MLP* apresentou melhor desempenho, levando em consideração as métricas *MAE*, *MAPE* e *RMSE*. A conclusão foi que, o modelo *MLP* revelou um notável potencial de generalização dentro do horizonte de previsão.

Palavras-chave: previsão; força e energia; estatística matemática.

ABSTRACT

Electricity is fundamental for the economic development of a country. It is precisely for this reason that the electric power sector is one of the fastest-growing sectors in the country. Brazil has one of the most diversified energy matrices in the world, characterized by a hydrothermal system and a predominance of renewable energy sources. Hydroelectric energy is one of the most widely used options globally for electricity generation. In 2021, in Brazil, the energy supply by source corresponded to 56,8% hydro, 12,8% natural gas, 8,2% biomass, 3% petroleum derivatives, 3,9% coal and derivatives, 10,6% wind, 2,2% nuclear, and 2,47% solar. Thus, the diversity of each country's matrix implies different energy planning strategies, as each energy source has its limitations. Forecasts, the ability to assess the history of electricity generation, and project a response can benefit the electricity sector. They are crucial tools in assisting the decision-making process, providing information that promotes efficiency and security in the supply of electric power. In this context, this work aimed to compare the performance of models, Multiple Linear Regression (MLR) and Multilayer Perceptron Artificial Neural Networks (MLP ANN), using the WEKA software, in predicting hydroelectric generation in Brazil. To perform this comparison, a dataset of hydroelectric generation was used. The database presented a historical series of hydroelectric energy generation from January 1996 to August 2022. The methodology consisted of developing prediction models based on Multiple Linear Regression (MLR) and Multilayer Perceptron Artificial Neural Networks (MLP ANN) algorithms using the WEKA software. The results showed that, for a 6-month horizon, the MLP model outperformed MLR, considering the MAE, MAPE and RMSE metrics. The conclusion was that the MLP model demonstrated notable generalization potential within the prediction horizon.

Keywords: forecast; power and energy; mathematical statistics.

LISTA DE ILUSTRAÇÕES

Figura 1 – Capacidade Instalada por estado.....	17
Figura 2 – Exemplo de uma pág. do BEN, contendo um comparativo dos anos - de 2020 e 2021, de geração eólica.....	21
Figura 3 – Consumo de energia por setor.....	21
Figura 4 – Oferta interna de energia elétrica.....	22
Figura 5 – Matriz elétrica brasileira.....	23
Figura 6 – Participação de renováveis na matriz elétrica brasileira.....	24
Figura 7 – Geração de energia elétrica por fonte.....	25
Figura 8 – Capacidade Instalada (TW).....	26
Figura 9 – Mapa dos Principais Intercâmbios de Energia Elétrica.....	28
Figura 10 – Neurônio de McCulloch e Pitts.....	32
Figura 11 – Estrutura básica e componentes principais de um neurônio artificial....	34
Figura 12 – Representação gráfica de uma RNA do tipo <i>MLP</i> com duas cama- das intermediárias	36
Figura 13 – Modelo de um neurônio de uma rede <i>MLP</i>	36
Figura 14 – Exemplos de funções de ativação: (a) função degrau; (b) função li - near; (c) função logística; (d) função hiperbólica.....	38
Figura 15 – Aprendizado supervisionado.....	40
Figura 16 – Superfície de erro para dois pesos e exemplos de mínimo global e - local.....>.....	41
Figura 17 – Exemplos de erros em treinamento de uma rede do tipo <i>MLP</i>	43
Figura 18 – Interface inicial do <i>software WEKA</i> (versão 3.8.6).....	47
Figura 19 – Procedimentos metodológicos.....	49
Figura 20 – Série Histórica de Energia Hidráulica.....	51
Figura 21 – Arquivo <i>ARFF</i>	53
Figura 22 – Área <i>Basic Configuration</i>	55
Figura 23 – Área <i>Advanced Configuration</i>	56
Figura 24 – Algoritmos de previsão.....	57
Figura 25 – Janelas de configuração dos algoritmos.....	58
Figura 26 – Parâmetros para Regressão Linear Múltipla (<i>MLR</i>).....	62
Figura 27 – Resultados – Métricas (<i>MLR</i>).....	62
Figura 28 – Geração – Treinamento e Validação (<i>MLR</i>).....	63
Figura 29 – Parâmetros para <i>Perceptron</i> de Múltiplas Camadas (<i>MLP</i>).....	65
Figura 30 – Resultados – Métricas (<i>MLP</i>).....	65
Figura 31 – Geração – Treinamento e Validação (<i>MLP</i>).....	66
Figura 32 – Previsões do conjunto Teste.....	67
Quadro 1 – Descrição dos parâmetros do modelo <i>MLP</i>	64

LISTA DE TABELAS

Tabela 1 – Empreendimentos em operação no Brasil.....	18
Tabela 2 – Empreendimentos em construção no Brasil.....	19
Tabela 3 – Empreendimentos com construção não iniciada no Brasil.....	19
Tabela 4 – Parte do conjunto de dados de geração hidráulica.....	50
Tabela 5 – Equações: <i>MAE</i> , <i>RMSE</i> e <i>MAPE</i>	59
Tabela 6 – Análise descritiva dos dados.....	60
Tabela 7 – Combinações das configurações para o modelo <i>MLP</i>	64
Tabela 8 – Configuração do modelo <i>MLP</i>	64
Tabela 9 – Dados de Geração Observados e Previstos (<i>MLP</i> e <i>MLR</i>) e a métrica <i>MAE</i>	67
Tabela 10 – Dados de Geração Observados e Previstos (<i>MLP</i> e <i>MLR</i>) e a métrica <i>RMSE</i>	67
Tabela 11 – Dados de Geração Observados e Previstos (<i>MLP</i> e <i>MLR</i>) e a métrica <i>MAPE</i>	68

SUMÁRIO

1	INTRODUÇÃO	13
2	OBJETIVOS	16
2.1	Objetivos Específicos	16
3	REFERENCIAL TEÓRICO.....	17
3.1	Matriz Elétrica do Brasil: Cenário Atual.....	17
3.1.1	Balço de Energia Elétrica	20
3.1.2	Consumo de Energia.....	21
3.1.3	Geraço de Energia Elétrica	22
3.1.4	Intercâmbios de energia elétrica.....	26
3.2	Regressão Linear Múltipla (MLR).....	29
3.3	Redes Neurais Artificiais (RNAs).....	30
3.3.1	Histórico das RNAs	31
3.3.2	Neurônio Artificial	33
3.3.3	Arquitetura de RNAs.....	34
<u>3.3.3.1</u>	<u>Redes <i>Perceptron</i> de Múltiplas Camadas</u>	<u>35</u>
3.3.4	Processos de Aprendizagem.....	39
3.3.5	O algoritmo <i>Backpropagation</i> (Retropropagaço)	44
3.4	<i>Weka</i>	47
4	PROCEDIMENTOS METODOLÓGICOS.....	49
4.1	Tipo de Pesquisa	49
4.2	Séries Históricas	50
4.3	Etapas do Trabalho	51
4.3.1	Coleta, análise e transformação de dados	52
4.3.2	Treinamento e Validaço.....	52
4.3.3	Teste	52
4.4	Arquivo <i>ARFF - Weka</i>.....	52
4.5	<i>Aba Forecast</i>.....	54
4.5.1	Área <i>Basic Configuration</i>	54
4.5.2	Área <i>Advanced Configuration</i>	56
4.6	Métricas	59
5	RESULTADOS.....	60
5.1	Análise Descritiva dos Dados	60
5.2	Treinamento e Validaço.....	61
5.2.1	Regressão Linear Múltipla (MLR).....	61

5.2.2	Redes Neurais <i>Multilayer Perceptron (MLP)</i>	63
5.3	Teste	66
6	CONCLUSÃO	69
	REFERÊNCIAS	70

1 INTRODUÇÃO

A energia elétrica é fundamental para o desenvolvimento econômico de um país. É justamente por essa razão que o setor de energia elétrica é um dos setores que mais se expande no país. O Brasil possui uma das matrizes energéticas mais diversificadas do mundo, caracterizada por um sistema hidrotérmico e pela predominância de fontes renováveis de energia (ROLA, 2017), as quais segundo o relatório do Balanço Energético Nacional (BEN), 2022, (ano base 2021) representam 78,1% da oferta interna de eletricidade no Brasil (soma dos montantes referentes à produção nacional mais as importações), com destaque para a geração hidráulica que responde por 56,8% da oferta interna.

A energia hidráulica (ou hídrica) é uma das opções mais utilizadas no mundo para geração de energia elétrica. É obtida a partir da energia potencial de uma massa de água, utilizando um desnível ou uma queda de água. É convertida em energia mecânica através de turbinas hidráulicas. As turbinas são utilizadas para o acionamento de geradores elétricos, com o objetivo de prover energia elétrica para uma rede de energia (SANTANA, 2019).

Em 2021, no Brasil, a oferta de energia elétrica por fonte correspondeu a 56,8% hidráulica, 12,8% gás natural, 8,2% biomassa (incluindo lenha, bagaço de cana, lixívia e outras recuperações), 3% derivados de petróleo, 3,9% de carvão e derivados, 10,6% eólica, 2,2% nuclear e 2,47% solar (EPE, 2022). Analisando os dados publicados no BEN, 2022, pela Empresa de Pesquisa Energética (EPE), e comparando a matriz energética nacional com outros países, como por exemplo a França, evidencia-se que 74,7% da geração interna francesa é nuclear e não renovável, completamente diferente da realidade brasileira. Desta forma, a diversidade da matriz energética de cada país implica em diferentes estratégias de planejamento energético, uma vez que cada fonte de energia possui suas limitações.

As previsões, capacidade de avaliar o histórico de geração de energia elétrica e projetar uma resposta, podem beneficiar o setor elétrico. São instrumentos de grande importância no auxílio ao processo de tomada de decisão, fornecem informações que proporcionam economia e segurança no fornecimento de energia elétrica. Evitando apagões, cortes, deterioração dos sistemas e até mesmo paralisação de serviços (SANTOS *et al.*, 2022; SOARES, 2022; COSTA, 2022).

Dentro do contexto do setor elétrico, a procura por modelos de previsão de curtíssimo, curto, médio e longo prazo que visam o consumo, demanda, abastecimento e prever a geração de energia elétrica de diferentes fontes, têm estimulado muitos pesquisadores (RAHMAN; HAZIM, 1993; CHARYTONIUK; CHEN, 2000; BECCALI *et al.*, 2004; YAP *et al.*, 2006; ROMERA *et al.*, 2007; MAMLOOK *et al.*, 2009; BASHIR; HAWARY, 2009; MELO, 2012; TIDRE *et al.*, 2013 e GÜNAY, 2016).

Segundo Oliveira e Menciondo, 2014, os primeiros modelos de previsão desenvolvidos baseavam-se em ferramentas estatísticas, como por exemplo, o método de regressão linear, o qual ainda hoje é amplamente utilizado. Devido à necessidade de solucionar problemas cada vez mais complexos e aliados ao avanço da tecnologia computacional, outros métodos foram desenvolvidos e empregados, tais como Redes Neurais Artificiais (RNAs), Algoritmos Genéticos, Lógica *Fuzzy*, entre outros. Considerando horizontes de previsão de curtíssimo, curto, médio e longo prazo, a aplicação de redes neurais mostrou-se uma boa alternativa para estimar comportamentos do setor energético, devido à produção de resultados interessantes para o Sistema Elétrico de Potência (SEP) (ANDRADE, 2010).

As Redes Neurais Artificiais (RNAs) podem ser utilizadas como ferramentas de apoio no processo de tomada de decisão. São ferramentas de grande importância em um sistema de planejamento, pois direcionam a melhor forma de aproveitar todos os recursos disponíveis (BASTIANI *et al.*, 2018). Capazes de atuarem em uma ampla gama de campos científicos, as RNAs, modelos matemáticos inspirados no funcionamento de neurônios biológicos, têm sido utilizadas com sucesso em tarefas de reconhecimento de fala, visão computacional, processamento de linguagem natural, predição e modelagem de séries temporais, classificação de dados, entre outras. Uma série temporal é qualquer conjunto de observações ordenado no tempo. Um modelo de séries temporais fornece previsões de observações futuras por meio da relação que possui com valores passados (MORETIN; TOLOI, 2004; HAYKIN, 2001). Conforme Haykin, 2004, uma RNA é um sistema de processamento massivamente paralelo, composto por unidades simples com capacidade natural de armazenar conhecimento de disponibilizá-lo para uso futuro. Existem diversas técnicas e configurações possíveis para uma RNA e uma arquitetura muito utilizada na literatura é a *Multilayer Perceptron (Perceptrons de Múltiplas Camadas - MLP)*.

Neste contexto, diante do que se discutiu até o presente momento e considerando a importância e relevância de um planejamento elétrico adequado, o qual se concretiza através de modelos e técnicas de previsão, que empregam inúmeras variáveis de influência, o presente trabalho teve como objetivo comparar a eficiência de modelos, de Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais (RNAs) do tipo *MLP* com algoritmo de aprendizado *backpropagation* do software *WEKA* na previsão de geração de energia hidráulica nacional, no horizonte de médio prazo (seis meses), levando em consideração dados do balanço de energia elétrica do país.

Para facilitar a compreensão, o texto foi subdividido em cinco segmentos distintos. Primeiramente, encontra-se a seção denominada “Introdução”, onde é apresentada uma visão geral do tema em questão. Em seguida, no “Capítulo 2”, são delineados os objetivos gerais e específicos desta pesquisa. O “Capítulo 3” aborda o referencial teórico, que tem o propósito de esclarecer os termos considerados mais relevantes para este estudo. O “Capítulo 4” detalha os métodos empregados para alcançar o objetivo estabelecido. No “Capítulo 5”, são apresentados os resultados obtidos por meio dessa abordagem, juntamente com uma comparação entre os métodos utilizados. Por fim, no último capítulo, são expostas as conclusões desta pesquisa.

2 OBJETIVOS

Neste trabalho teve-se como objetivo comparar o desempenho de modelos, de Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais (RNAs) do tipo *MLP*, na previsão da geração de energia hidráulica no Brasil.

2.1 Objetivos Específicos

Os objetivos específicos, descritos abaixo, foram propostos para que o objetivo geral fosse alcançado.

1. Tratar os dados históricos de geração hidráulica mensal (em GWh), desde o ano de 1996;
2. Desenvolver a pesquisa e compreender o funcionamento do software WEKA, explorando suas ferramentas de previsão de algoritmos de aprendizagem;
3. Desenvolvimento de modelos de previsão utilizando regressão linear múltipla (*MLR*) e redes neurais artificiais (RNAs);
4. Realizar a validação dos modelos criados utilizando métricas de erro (*RMSE*, *MAE*, *MAPE*);
5. Realizar uma comparação entre os modelos, a fim de identificar qual deles apresenta o melhor desempenho.

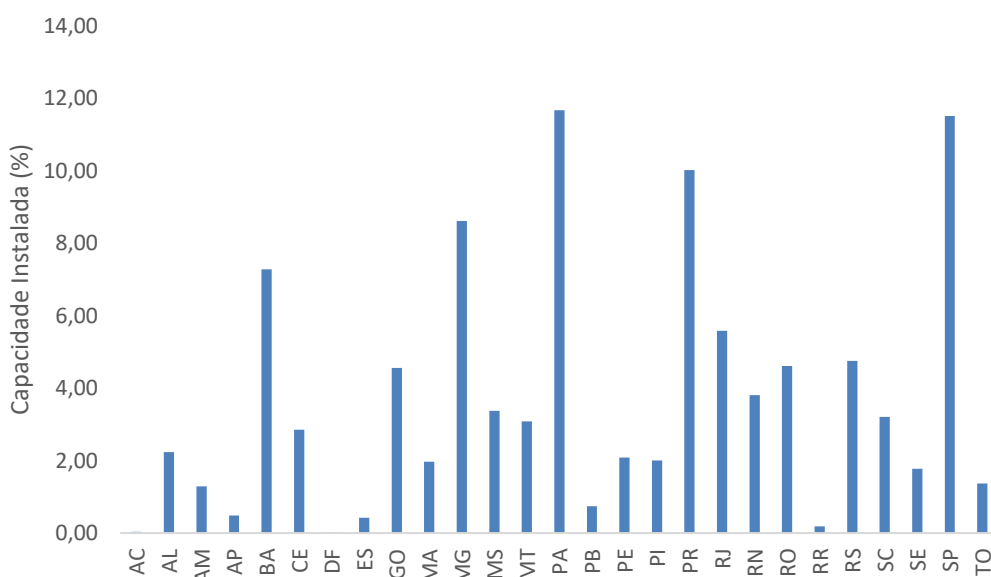
3 REFERENCIAL TEÓRICO

3.1 Matriz Elétrica do Brasil: Cenário Atual

A matriz energética brasileira é predominantemente renovável, com destaque para a energia hidrelétrica, que representa cerca de 60% da capacidade instalada. Em seguida, estão as fontes de energia eólica e solar, que apresentaram crescimento significativo nos últimos anos, representando cerca de 10% da capacidade instalada. Além disso, há empreendimentos de fontes térmicas, como energia nuclear e a termelétrica, que complementam a geração elétrica (EPE, 2022). Segundo dados do Ministério de Minas e Energia (MME), atualmente, a matriz elétrica brasileira tem uma capacidade instalada de aproximadamente 182 GW e está em processo de implementação de novos projetos, como leilões para contratação de energia eólica, solar e hídrica, como objetivo de expandir a capacidade de geração de energia elétrica no país.

A Figura 1 ilustra a capacidade instalada, em porcentagem, no ano de 2022, para todos os estados do território nacional.

Figura 1 – Capacidade instalada por estado.



Fonte: Elaboração própria a partir de dados do BEN (2022).

As principais fontes de energia que compõem a matriz de geração de energia elétrica do Brasil são a hidrelétrica, que corresponde a cerca de 62,5% da

capacidade instalada, seguida da termelétrica 25,6%, e eólica, com 9,1%. Além disso, há a presença de outras fontes, como nuclear, com 1,2% e a solar, com 1,1% (EPE, 2022).

Os empreendimentos derivados da fonte hídrica são classificados conforme a ANEEL em Central Geradora Hidrelétrica (CGH) com potência instalada menor ou igual a 1 MW, Pequena Central Hidrelétrica (PCH) com potência instalada entre 1 e 30 MW e Usina Hidrelétrica (UHE) com potência instalada superior a 30 MW. Já os oriundos da fonte eólica recebem denominação de Central Geradora Eólica (EOL) e os da fonte térmica são chamados de Usina Termelétrica (UTE) (ANEEL, 2022).

Dentre os empreendimentos em operação no Brasil, aproximadamente 46% fazem uso da energia hídrica, com predominância para as UHE (39,64%), seguidos de usinas termelétricas, eólicas e fotovoltaicas, conforme demonstra-se na Tabela 1.

Tabela 1 – Empreendimentos em operação no Brasil.

Fonte de Energia	Quantidade de Empreendimentos	Potência Instalada (MW)	Porcentagem (%)
UHE	168	58.430	39,64
UTE	208	41.670	28,27
EOL	695	24.870	16,87
PCH	525	9.730	6,60
CGH	53	150	0,10
UFV	1.800	12.550	8,51
Total	3.449	147.400	100

Fonte: Elaboração própria a partir de dados da ANEEL (2022).

As projeções, de acordo com a Agência Nacional de Energia Elétrica (ANEEL, 2022), para os próximos anos são otimistas em relação ao crescimento da capacidade instalada de geração de energia elétrica no Brasil, impulsionadas principalmente pela geração distribuída elétrica no país, e que aumente em cerca de 28,5 GW até 2026, com destaque para a energia solar fotovoltaica, que deverá crescer de 82% ao ano. A energia eólica também deverá ter um aumento significativo, com previsão de crescimento de 2,7 GW em 2022 e 6,8 GW em 2026. A geração hidrelétrica, apesar de ainda ser a principal fonte de energia no Brasil, deverá ter um crescimento mais modesto, com previsão de aumento de apenas 1,3 GW até 2026. Além disso, a ANEEL prevê aumento na participação de outras fontes renováveis de energia, como biomassa e biogás, na matriz energética brasileira nos

próximos anos. Na tabela 2, evidenciam-se alguns dos empreendimentos que estavam em construção no Brasil, até o final do ano de 2022.

Tabela 2 – Empreendimentos em construção no Brasil.

Fonte de Energia	Quantidade de Empreendimentos	Potência Instalada (MW)	Porcentagem (%)
UHE	26	4.970	17,37
UTE	7	2.350	8,22
EOL	80	1.890	6,61
PCH	77	1.160	4,06
CGH	4	35	0,12
UFV	1.165	18.200	63,63
Total	1.359	28.605	100

Fonte: Elaboração própria a partir de dados da ANEEL (2022).

Empreendimentos que já possuem registro na ANEEL, porém ainda não tiveram sua construção iniciada, têm perspectiva de uma potência instalada sujeita a mudanças e atualizações ao longo do tempo, devido a fatores como questões regulatórias, viabilidade econômica e condições de mercado. No entanto, nos últimos anos, a energia eólica (EOL) tem se destacado como uma das fontes com maior crescimento e liderança na expansão da capacidade instalada no Brasil. A Tabela 3 descreve que tipos de empreendimentos que não tiveram sua construção iniciada.

Tabela 3 – Empreendimentos com construção não iniciada no Brasil.

Fonte de Energia	Quantidade de Empreendimentos	Potência Instalada (MW)	Porcentagem (%)
UHE	18	5.560	27,61
UTE	2	1.120	5,56
EOL	57	2.850	14,15
PCH	123	3.610	17,92
CGH	8	50	0,25
UFV	415	6.950	34,51
Total	623	20.140	100

Fonte: Elaboração própria a partir de dados da ANEEL (2022).

Os dados apresentados neste subcapítulo, até o presente momento, descrevem registros atualizados (dezembro de 2022) da capacidade instalada dos empreendimentos, separados por tipo, quantidade e potência, além de evidenciar qual a atual situação deles (em operação, construção ou construção não iniciada).

Na sequência apresentar-se-ão informações referentes ao balanço de energia elétrica do país, tais como: consumo, geração, importação e exportação (internacional e do SIN).

3.1.1 Balanço de Energia Elétrica

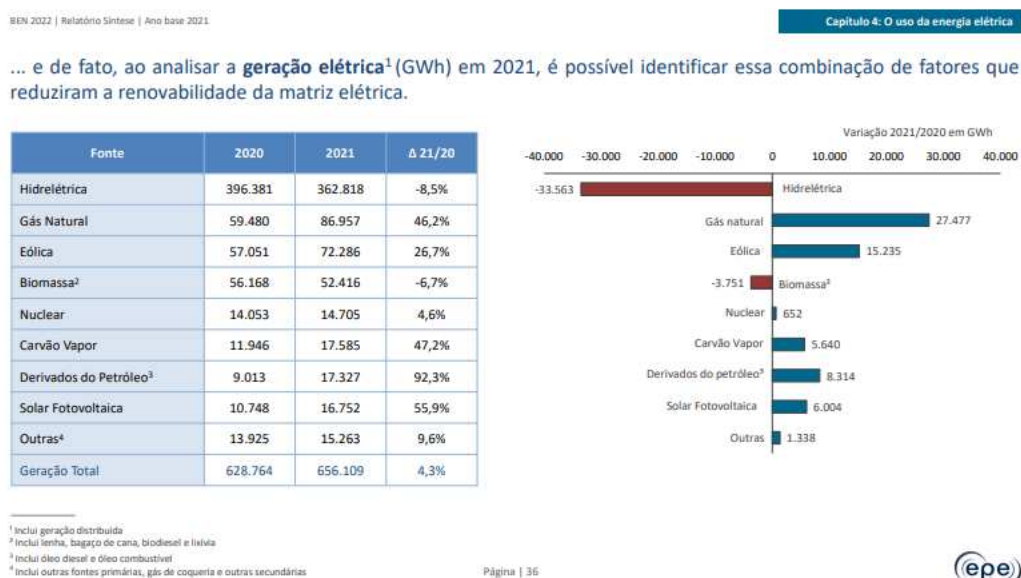
O Balanço Energético Nacional (BEN) pode ser utilizado como uma ferramenta para monitorar e quantificar a oferta de demanda de energia elétrica em um determinado país ou região. Ele permite avaliar a situação energética do país, identificar possíveis deficiências na oferta, acompanhar o crescimento da demanda e planejar estrategicamente a expansão do setor elétrico. Na Figura 2 é possível observar uma página do BEN contendo um comparativo entre os anos de 2020 e 2021, de geração elétrica (GWh) no país.

Um dos cálculos apresentados no BEN envolve a comparação entre a quantidade de energia elétrica gerada e a quantidade de energia elétrica consumida em um determinado período. O processo consiste em três principais componentes:

1. Geração de Energia Elétrica: Nesta etapa, são consideradas todas as fontes de energia utilizadas para a geração de eletricidade, tais como usinas hidrelétricas, termelétricas, eólicas, solares, nucleares, entre outras. Cada fonte é quantificada em termos de sua capacidade instalada e sua produção efetiva de energia elétrica;
2. Transmissão e Distribuição: Após a geração, a energia elétrica é transmitida através de linhas de transmissão de alta tensão, em seguida, distribuída para os consumidores finais. Nesta etapa, são consideradas as perdas de transmissão e distribuição, que ocorrem devido a fatores como resistência dos cabos, transformadores e outros elementos do sistema elétrico;
3. Consumo de Energia Elétrica: Aqui, é estimada a demanda de energia pelos diversos setores da economia, tais como residencial, comercial, industrial, agrícola e serviços públicos. Essa estimativa é

baseada em dados históricos, análises de tendências e projeções de crescimento econômico.

Figura 2 – Exemplo de uma pág. do BEN, contendo um comparativo dos anos de 2020 e 2021, de geração elétrica (GWh).



Fonte: BEN (2022).

3.1.2 Consumo de Energia

Em 2021, o consumo de energia no país foi encabeçado pelo setor de transporte de carga e passageiros, seguido pelos setores industrial, residencial, energético, rural e outros (incluindo serviços e atividades relacionadas) (BEN, 2022). A distribuição do consumo energético por esses setores é visualmente representada na Figura 3.

Figura 3 – Consumo de energia por setor.



Fonte: Adaptada do BEN (2022).

3.1.3 Geração de Energia Elétrica

No contexto energético do Brasil, observou-se um incremento no consumo de energia elétrica, resultando na necessidade do aumento da oferta interna desse recurso para atender à demanda, conforme ilustrado pela Figura 4. Esse cenário reflete a crescente demanda por eletricidade no país, impulsionada pelo desenvolvimento econômico, expansão das atividades industriais, crescimento populacional e avanços tecnológicos. Como resultado desse aumento no consumo, foram realizados investimentos e ampliações na capacidade de geração, transmissão e distribuição de energia elétrica, a fim de garantir o fornecimento contínuo e confiável para os consumidores. Essa expansão da oferta interna de energia elétrica é fundamental para suprir as demandas da sociedade e sustentar o desenvolvimento socioeconômico do país.

Figura 4 – Oferta interna de energia elétrica.



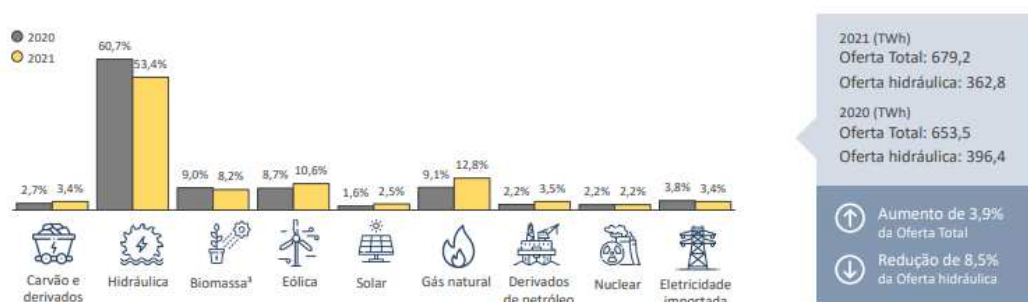
Fonte: Adaptada do BEN (2022).

As perdas de energia, representadas pela soma das perdas comerciais e técnicas, registraram um ligeiro decréscimo de 0,2 ponto percentual em relação ao ano de 2020. Essas perdas são resultado de fatores como falhas técnicas, desgaste de equipamentos, problemas na distribuição, além de questões relacionadas à eficiência energética e à gestão comercial.

A matriz elétrica brasileira em 2021 passou por significativas alterações devido à ocorrência de escassez hídrica ao longo desse ano. A falta de chuvas afetou diretamente o volume de água armazenado nos reservatórios das usinas hidrelétricas, que são a principal fonte de geração de energia no país. Essa situação desafiadora levou a um maior acionamento de termelétricas, que utilizam fontes de energia não-renováveis, como gás natural e óleo combustível, para suprir a demanda energética. Esse acréscimo na participação das termelétricas na matriz

elétrica impactou tanto no aumento da geração de energia proveniente de fontes fósseis quanto no custo operacional do sistema, evidenciado pela Figura 5. Nesse contexto, medidas de gestão da demanda e de estímulo à geração de energia a partir de fontes renováveis se tornaram ainda mais relevantes para garantir o suprimento energético do país.

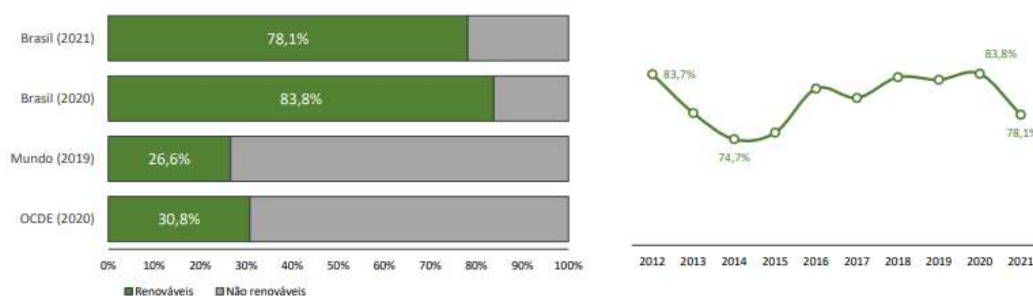
Figura 5 – Matriz elétrica brasileira.



Fonte: Adaptada do BEN (2022).

A escassez de chuvas ao longo do ano de 2021 resultou em uma diminuição significativa do nível dos reservatórios das principais usinas hidrelétricas do Brasil, o que por sua vez impactou negativamente a disponibilidade de energia hidrelétrica. Para compensar essa redução na oferta, houve um aumento substancial na geração de energia proveniente de outras fontes. O carvão vapor registrou um crescimento de 47,2%, o gás natural aumentou em 46,2%, a energia eólica teve um aumento de 26,7% e a energia solar fotovoltaica apresentou um expressivo crescimento de 55,9%. Essas fontes alternativas desempenharam um papel crucial na garantia do abastecimento energético do país diante da escassez hídrica.

Em 2021, a matriz elétrica brasileira registrou um avanço significativo em relação à incorporação de fontes de energia renovável, como evidenciado na Figura 6, alcançando uma participação de 78,1% de renovabilidade. Isso significa que a maior parte da geração de energia elétrica do país foi proveniente de fontes limpas e sustentáveis, como hidrelétricas, eólicas, solares e biomassa.

Figura 6 – Participação de renováveis na matriz elétrica brasileira.

Fonte: Adaptada do BEN (2022).

Essa transição na matriz elétrica brasileira ocorreu como resposta à queda na oferta de energia hidráulica, que foi parcialmente compensada pelo aumento da geração termelétrica, principalmente por meio do uso do gás natural. A escassez hídrica afetou negativamente a disponibilidade de água nos reservatórios das hidrelétricas, reduzindo assim sua capacidade de geração de energia. Diante dessa situação, houve uma maior utilização das usinas termelétricas movidas a gás natural para suprir a demanda energética do país. Essa alternativa contribuiu para garantir a estabilidade no fornecimento de eletricidade e mitigar os impactos causados pela diminuição da geração hidrelétrica devido à falta de chuvas.

E de fato, ao analisar a geração elétrica (GWh) em 2021, como ilustrado pela Figura 7, torna-se evidente a influência desses diferentes fatores que contribuíram para a redução da proporção de fontes renováveis na matriz elétrica. A escassez hídrica e a diminuição da geração hidrelétrica impactaram diretamente a disponibilidade de energia limpa proveniente das usinas hidrelétricas. Como resultado, houve aumento proporcional na geração de energia por meio de fontes não renováveis, como termelétricas movidas a carvão e gás natural. Essa combinação de fatores resultou em uma diminuição da renovabilidade da matriz elétrica brasileira no período analisado.

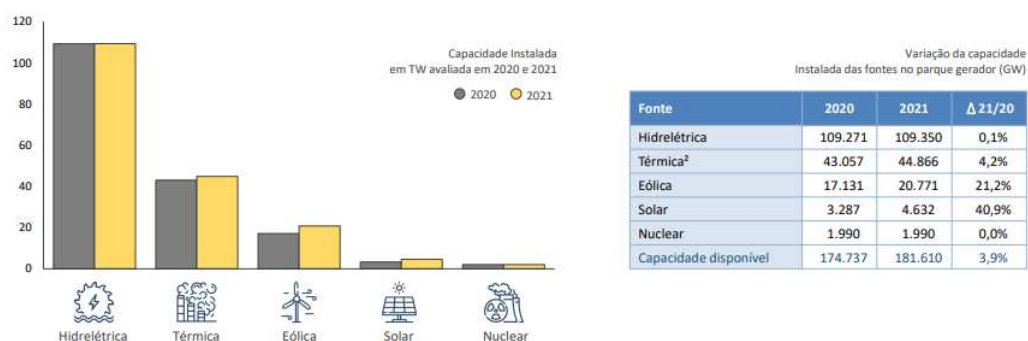
Figura 7 – Geração de energia elétrica por fonte.



Fonte: Adaptada do BEN (2022).

Em 2021, a capacidade instalada de geração de energia elétrica no Brasil registrou um notável crescimento de 3,9% em comparação com o ano anterior, marcando um avanço significativo no setor energético do país. Entre várias fontes de energia, destacam-se a energia eólica e solar como os principais impulsionadores desse aumento substancial. A energia eólica desempenhou um papel essencial nesse cenário, com a expansão de parques eólicos e o aumento do número de aerogeradores em diversas regiões do território nacional. Paralelamente, a energia solar fotovoltaica também exibiu um avanço expressivo, impulsionado pelo aumento da instalação de sistemas de geração distribuída e pelo desenvolvimento de grandes usinas solares.

Esses resultados refletem de maneira inequívoca a crescente importância das fontes renováveis na matriz energética brasileira, contribuindo para a diversificação e sustentabilidade do setor de energia elétrica. A Figura 10, que apresenta uma representação comparativa da capacidade instalada em terawatts (TW) nos anos 2020 e 2021, fornece uma perspectiva visual da notável evolução observada nesse período.

Figura 8 – Capacidade Instalada (TW).

Fonte: Adaptada do BEN (2022).

3.1.4 Intercâmbios de energia elétrica

Segundo dados do Boletim Mensal de Energia, publicado pela Empresa de Pesquisa Energética (EPE) em janeiro de 2022, entre janeiro e novembro de 2021, o Brasil importou energia elétrica principalmente da Argentina, do Uruguai e do Paraguai.

A importação de energia elétrica ocorre principalmente nos períodos de escassez hídrica, quando a geração hidrelétrica nacional é reduzida devido à baixa disponibilidade de água nos reservatórios. Nesses momentos, o Brasil recorre à importação para suprir a demanda interna de energia elétrica. Vale ressaltar que a importação de energia elétrica é uma estratégia utilizada para garantir a segurança energética e o abastecimento elétrico do país em períodos críticos.

Em fevereiro de 2022, ocorreram diversos fluxos de intercâmbio de energia elétrica entre os subsistemas do Brasil e com outros países. Seguem as principais informações:

- Subsistema Norte: O subsistema Norte continuou com perfil exportador, fornecendo 10.569 MW médios de energia elétrica. Esse valor inclui o fluxo nos bipolos do nó Xingu e representa um aumento ligeiro em relações ao mês anterior, que foi de 10.520 MW médios;
- Subsistema Nordeste: O subsistema Nordeste manteve o perfil importador do subsistema Sudeste/Centro-Oeste, recebendo 7.718

MWmédios de energia elétrica. Esse valor é superior ao montante importado no mês anterior que foi de 6.800 MWmédios;

- Bipolos de corrente contínua: Os bipolos de corrente contínua também contribuíram para o intercâmbio de energia. A coletora Porto Velho transmitiu 5.724 MWmédios, o nó de Xingu transmitiu 7.893 MWmédios e os bipolos que escoam a energia de Itaipu (50 Hz) transmitiram 1.172 MWmédios para o subsistema Sudeste/Centro-Oeste;
- Subsistema Sudeste/Centro-Oeste: O subsistema Sudeste/Centro-Oeste manteve o perfil importador em relação aos subsistemas Norte/Nordeste, importando um total de 12.395 MWmédios de energia elétrica. Por outro lado, foi exportador para o subsistema Sul no montante de 7.718 MWmédios resultando e um total de 4.677 MWmédios importados. Além disso, recebeu 14.789 MWmédios por meio dos bipolos de corrente contínua;
- Intercâmbio Internacional: No mês de fevereiro de 2022, também ocorreu intercâmbio de energia elétrica com o Uruguai, sendo que o Brasil importou um montante de 3 MWmédios.

Tanto essas informações quanto o mapa dos subsistemas brasileiros, ilustrado pela Figura 9, foram extraídas a partir do Boletim de Carga Mensal divulgado pelo Operador Nacional do Sistema Elétrico (ONS). Esses dados fornecem *insights* sobre os complexos fluxos de energia que ocorrem entre os diversos subsistemas do país, bem como suas conexões com nações vizinhas. Através da análise dessas informações detalhadas, é possível entender melhor a dinâmica da geração, transmissão e distribuição de energia elétrica no Brasil.

3.2 Regressão Linear Múltipla (MLR)

O modelo de regressão linear, segundo Krajevski, Ritzman e Malhota (2009), é um dos modelos causais mais conhecidos e utilizados, que consiste de uma variável chamada de dependente estar relacionada a uma ou mais variáveis independentes por uma equação linear. Em termos técnicos, pode-se afirmar que a linha de regressão busca minimizar os quadrados dos desvio entre os dados reais e os valores previstos pelo modelo. De acordo com Fávero (2015), o objetivo da análise de regressão é propiciar ao pesquisador condições de avaliar como se comporta uma variável Y com base no comportamento de uma ou mais variáveis X . A equação 1 apresenta o modelo geral para uma regressão linear múltipla:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad 1$$

onde Y refere-se a variável dependente que está sendo prevista, α é o coeficiente de interceptação (ou ponto de corte) da reta de regressão com o eixo Y , $\beta_1, \beta_2, \dots, \beta_k$ são os coeficientes de regressão que representam a relação entre as variáveis independentes, X_1, X_2, \dots, X_{ki} são as variáveis independentes que são usadas para prever a variável dependente Y e ε_i são os termos de erro, que representa a variação não explicada pela relação linear entre as variáveis. Essa equação tem como objetivo estabelecer a relação entre a variável independente, usada para prever uma variável dependente, neste caso, a previsão de geração hidráulica.

3.3 Redes Neurais Artificiais (RNAs)

As RNAs funcionam de forma a representar o funcionamento do cérebro humano de maneira simplificada, sendo capazes de solucionar problemas complexos empregando os conhecimentos adquiridos a partir do ambiente (HAYKIN, 2001). O cérebro é um sistema de processamento de dados bastante complexo, paralelo e não linear. Ele consegue organizar suas estruturas básicas, os neurônios, de modo que consiga processar certos tipos de informação bem mais rapidamente do que qualquer computador já construído (HAYKIN, 2005).

De acordo com Vasilev *et al.*, (2019), existem cinco aspectos importantes a serem considerados no entendimento de redes neurais. Primeiramente, o processamento de informações ocorre em elementos simples conhecidos como neurônios. Os neurônios, um dos mais importantes elementos de uma RNA, são unidades de processamento simples que compõem a mesma e tem por objetivo computar funções matemáticas a partir de entradas apresentadas, sendo que estas funções são na grande maioria, não lineares (HAYKIN, 2004). Esses neurônios, para permitir a troca de informações, possuem ligações de conexão entre si. Segundo Haykin (2004), a disposição paralela dos neurônios constitui camadas, as quais são interligadas com uma camada vizinha através de conexões e são associadas a pesos. Esses pesos recebem o nome de pesos sinápticos em consequência das múltiplas ligações dos neurônios no cérebro humano. As conexões entre unidades de uma rede neural são chamadas sinapses. Esses pesos sinápticos têm a função de armazenar o conhecimento obtido pelo modelo e também servem para avaliar a entrada recebida por cada neurônio da rede (PÁDUA, 2007). A força dessas conexões pode variar, o que é um fator fundamental no processamento das informações pelas redes neurais. Os neurônios também possuem estados internos que são definidos pelas suas conexões diretas. Por fim, as funções de ativação dos neurônios são responsáveis por determinar o valor de saída da rede neural.

Em uma RNA, uma etapa fundamental é o treinamento dos pesos sinápticos, a fim de permitir que a rede represente adequadamente o problema em questão. Essa etapa é denominada de aprendizado. Após a rede ter sido treinada, os pesos sinápticos são fixados e a rede é utilizada como um modelo para estimar as saídas futuras de novos conjuntos de entradas.

Partindo de uma visão matemática, uma rede neural pode ser entendida como um grafo direcionado em que os nós correspondem aos neurônios e as arestas às ligações entre eles. Dessa forma, cada neurônio recebe como valor de entrada a soma ponderada das saídas produzidas pelos neurônios diretamente conectados a ele, e sua saída pode servir de entrada para os próximos neurônios da rede (SHALEV-SHWARTZ; BEM-DAVID, 2014).

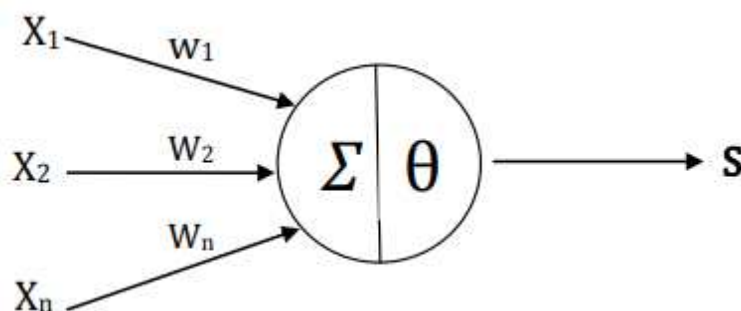
3.3.1 Histórico das RNAs

O conceito de neurônio artificial foi introduzido no trabalho pioneiro de McCulloch e Pitts em 1943, no qual os pesquisadores propuseram um modelo artificial de neurônio biológico, sem entrar no mérito de técnicas de aprendizado. A ideia por trás do modelo era descrever as capacidades computacionais do neurônio biológico, com o objetivo de simular seu comportamento em um ambiente artificial.

O modelo de neurônio de McCulloch e Pitts, também conhecido como neurônio MCP em referência aos nomes dos autores, consiste em um elemento simples que recebe uma ou mais entradas e produz uma saída binária (0 ou 1). O neurônio artificial recebe os sinais de entrada, que são multiplicados por pesos sinápticos, e a soma desses sinais é comparada com um valor limite, chamado de “limiar de ativação”. Se a soma for maior ou igual ao limiar, o neurônio produz uma saída de valor 1.

O neurônio MCP é representado de forma esquemática na Figura 10, que ilustra a estrutura do neurônio artificial e as conexões sinápticas entre os elementos. Embora o modelo seja bastante simples, ele forneceu a base para o desenvolvimento de redes neurais mais complexas, que são capazes de realizar tarefas mais sofisticadas, como reconhecimento de padrões, classificação e previsão.

Figura 10 – Neurônio de McCulloch e Pitts.



Fonte: Adaptada de Guarnieri (2006).

Neste modelo as entradas (x_1, x_2, \dots, x_n) , originários dos dados de entrada da rede ou das saídas de outros neurônios, são ajustadas pelos pesos sinápticos (w_1, w_2, \dots, w_n) das ligações, destacando a limitação (pesos fixos e não ajustáveis). O neurônio faz o somatório (Σ) dos produtos $x_i w_i$, e decide se o neurônio deve ou não disparar, caso este valor seja igual ou superior a um limiar θ , o sinal de saída (S) do neurônio será igual a 1, e diz-se então que o neurônio tem sua saída ativa ou que está excitado, ou se o número for menor que o limiar, a saída (S) será zero (GUARNIERI, 2006; BRAGA, 2007).

Em 1949, Hebb descobriu que a plasticidade de aprendizagem em redes neurais é originada pela alteração das conexões sinápticas dos neurônios de entrada. O autor explicou que o aprendizado ocorre quando células pré e pós-sinápticas são ativadas, ou seja, quando há um reforço na conexão neuronal.

Em 1958, Rosenblatt propôs um novo modelo empregando neurônios de McCulloch e Pitts, mas com pesos sinápticos ajustáveis, que recebeu o nome de *perceptron*. Nesse modelo, os pesos sinápticos eram calibrados de acordo com a saída desejada através de um algoritmo que treinava a rede para executar determinadas funções. O modelo também é conhecido como “*perceptron* de uma camada” porque apenas a camada de saída possui pesos ajustáveis. As primeiras RNAs, assim como a de Rosenblatt, 1958, eram usadas em sua grande maioria para reconhecer e separar grupos de caracteres (SILVA, 2010).

Em 1960, Widrow e Hoff apresentaram uma nova técnica de aprendizado denominada “regra delta”. Esta regra tem por objetivo minimizar o erro de saída de um neurônio com resposta linear através do método do gradiente descendente (HAYKIN, 2001). A regra delta é bastante utilizada em redes do tipo *perceptron* de uma única camada e também pode ser empregada em outras arquiteturas de redes

neurais. O modelo no qual continha esta nova regra de aprendizado recebeu o nome de *ADALINE* (*Adaptive Linear Neuron*) e diferente do *perceptron*, que tem saídas entre 0 e 1, este modelo possui saídas entre -1 e 1. Outra diferença observada é que nesse modelo não existe limiar de ativação, o termo de polarização θ é utilizado para calcular a saída (S) (BRAGA, 2007).

Minsky e Papert, 1969, chamaram atenção para algumas limitações dos modelos de RNA da época, que eram capazes de resolver apenas problemas linearmente separáveis. Eles também destacaram que os algoritmos de aprendizado em modelos *perceptrons* com mais de uma camada não garantiam uma boa convergência em comparação com os modelos de apenas uma camada. Essas considerações desencorajaram muitos pesquisadores, e o assunto só voltou a ser amplamente discutido em 1982, com o trabalho de Hopfield, que evidenciou a relação entre redes recorrentes auto associativas e sistemas físicos (HAYKIN, 2001).

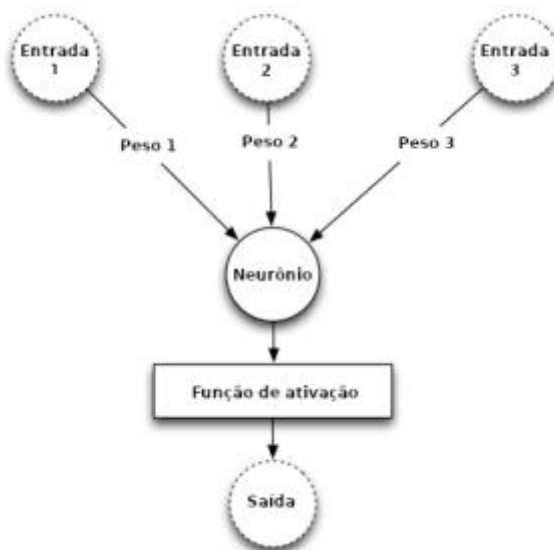
Rumelhart *et al.*, 1986, apresentaram um novo algoritmo de treinamento chamado *backpropagation*, que contradisse as críticas de Minsky e Papert, demonstrando a viabilidade de solucionar problemas complexos por meio do treinamento de RNAs de múltiplas camadas.

Braga, 2007, descreve que, desde então, houve inúmeras pesquisas propondo a variação do algoritmo *backpropagation* para obter uma convergência mais rápida e que o ressurgimento do interesse em RNA continua até hoje.

3.3.2 Neurônio Artificial

Um neurônio artificial é considerado a unidade fundamental de processamento em uma rede neural (HAYKIN, 2005). Ele realiza uma operação simples, que envolve a multiplicação de cada valor de entrada pelo seu respectivo peso sináptico. Esses pesos são responsáveis por atribuir diferentes níveis de importância aos atributos de entrada. Em seguida, os valores resultantes são somados e passam por uma função de ativação, que tem como objetivo restringir a amplitude de saída do neurônio (HEATON, 2015). A Figura 11 apresenta a estrutura básica de um neurônio artificial juntamente com seus componentes principais.

Figura 11 – Estrutura básica e componentes principais de um neurônio artificial.



Fonte: Adaptada de Heaton (2015).

A equação 2 representa matematicamente um neurônio artificial, onde x_i representa os valores de entrada do neurônio, ω_i denota os pesos das conexões entre esses valores de entrada e b é um tipo de peso especial chamado *bias*. O resultado obtido é então passado por uma etapa de transformação f , geralmente não linear, conhecida como função de ativação ou função de transferência (VASILEV *et al.*, 2019).

$$y = f\left(\sum_i x_i \cdot \omega_i + b\right) \quad 2$$

3.3.3 Arquitetura de RNAs

A escolha de arquiteturas de RNAs é baseada no tipo de problema a ser abordado. Existem três parâmetros fundamentais para a definição da arquitetura de uma rede: o número de camadas da rede, número de neurônios em cada camada e um tipo de conexão entre os neurônios.

No que diz respeito ao número de camadas, conforme evidenciado na subseção 3.3.1 “Histórico das RNAs”, existem redes de uma única camada, que possuem apenas um neurônio entre as camadas de entrada e saída da rede, e são adequadas para resolver problemas linearmente separáveis. Por outro lado, as

RNAs de múltiplas camadas possuem mais de um neurônio entre a camada e a saída da rede e são utilizadas para solucionar problemas que envolvem funções linearmente separáveis ou não (CYBENKO, 1988).

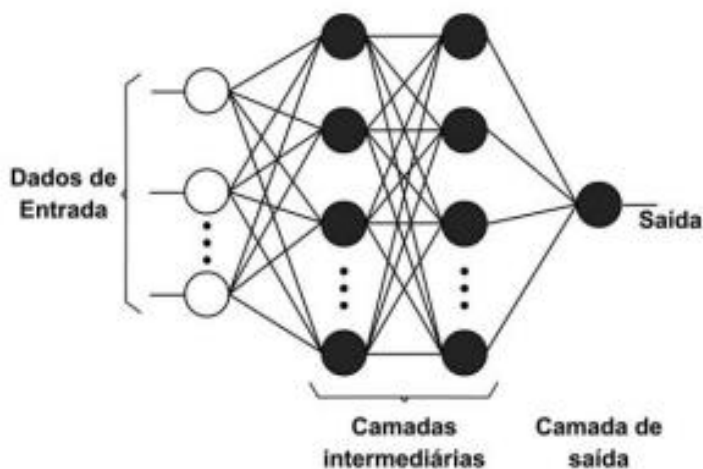
Em relação ao número de neurônios a serem empregados, que desempenha um papel importante no desempenho da rede, não existe uma regra geral na literatura para determinar a quantidade adequada para um determinado problema. De acordo com Braga, 2007, uma abordagem é buscar uma estrutura mínima que atenda aos requisitos de minimização do erro quadrático no conjunto de treinamento. Conforme Boser *et al.*, 1992, essa abordagem empírica de solucionar problemas pode ser caracterizada como minimização de risco estrutural.

Existem duas conexões básicas na arquitetura das RNAs: as não-recorrentes (ou *feedforward*) e as conexões recorrentes (ou *feedback*). As topologias que não possuem realimentação das saídas para as entradas são chamadas de não-recorrentes, enquanto as recorrentes possuem realimentação das saídas para as entradas (HAYKIN, 2001). Neste trabalho, foi enfatizada a utilização de uma RNA do tipo não-recorrente, devido à sua aplicabilidade.

3.3.3.1 Redes *Perceptron* de Múltiplas Camadas

Como demonstrado na subseção 3.3.3, o *perceptron* é um modelo de rede neural composto por uma camada de neurônios de entrada e outra de saída, o que limita sua capacidade de resolver problemas linearmente separáveis. Para contornar essa limitação, foram adicionadas uma ou mais camadas intermediárias, também conhecidas como camadas escondidas. Assim, surgiu a rede *perceptron* de múltiplas camadas, conhecida como *Multilayer Perceptron (MLP)*, que é composta por um conjunto de entradas, uma ou mais camadas intermediárias e uma camada de saída. A Figura 12 ilustra um exemplo de uma *MLP* com duas camadas intermediárias.

Figura 12 – Representação gráfica de uma RNA do tipo *MLP* com duas camadas intermediárias

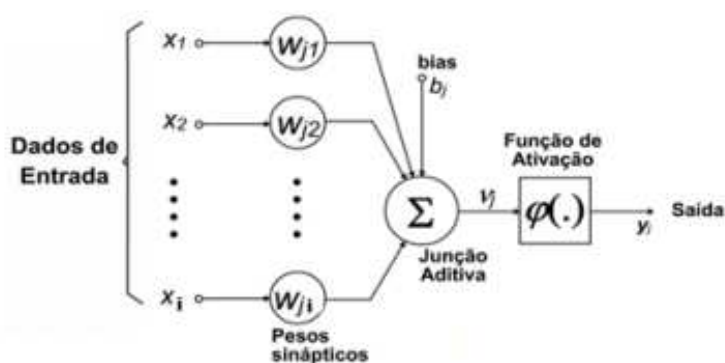


Fonte: Adaptada de Fiorin *et al.*, (2011).

Diversos autores desconsideram o termo “camada de entrada”, na arquitetura de rede neurais, pois essa camada apenas recebe e distribui os dados, sem realizar nenhuma operação matemática como ponderação, somatório ou comparação com limiares. Portanto, assume-se apenas o termo “entrada” (WASSERMAN, 1989).

A diferença entre os neurônios do modelo *perceptron* de uma única camada e os neurônios do *MLP* de múltiplas camadas é que os neurônios do *MLP* permitem a saída de um valor contínuo, ao contrário do *perceptron* que produz uma saída binária (0 ou 1) através de um somatório ponderado das entradas comparado a um limiar. Para contornar essa limitação, novos modelos de neurônios artificiais foram criados, aplicando funções de ativação nos valores ponderados das entradas. A Figura 13 ilustra um neurônio genérico em uma estrutura de rede do tipo *MLP*.

Figura 13 – Modelo de um neurônio de uma rede *MLP*.



Fonte: Adaptada de Haykin (2001).

Na Figura 13, v_j representa o nível de atividade interna do neurônio, que é obtido através da combinação do somatório das entradas ponderadas ($x_i \omega_{ji}$) com um valor ajustável adicional chamado *bias* (b_j). Ao aplicar o nível de atividade interna a uma função de ativação ($\varphi(\cdot)$), obtêm-se a saída do neurônio (y_j). A função de ativação do neurônio atua sobre a superfície de mapeamento definida pelos pesos sinápticos, e seu principal objetivo é introduzir não linearidade no modelo. Sem a função de ativação, o poder de computação do neurônio, e a rede neural como um todo seria equivalente a uma transformação linear, o que não permitiria a solução de problemas não-lineares (SILVA *et al.*, 2010).

O valor ajustável adicional, *bias*, desempenha um papel semelhante ao limiar θ do neurônio de McCulloch e Pitts, 1943. Ele tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se o *bias* é positivo ou negativo, respectivamente (HAYKIN, 2004; FIORIN, 2011). As equações 3 e 4 descrevem todas essas considerações para o neurônio j .

$$v_j = \sum_{i=1}^m x_i \omega_{ji} + b_j \quad 3$$

$$y_j = \varphi(v_j) \quad 4$$

A camada intermediária é responsável pela extração das características do problema, e geralmente são utilizadas funções de ativação logarítmicas ou sigmóides (REIDMILLER, 1994). Algumas das funções de ativação comumente empregadas incluem a função degrau (utilizada por McCulloch e Pitts, 1943), a função linear e as funções sigmóides logística e tangente hiperbólica, representadas pelas Equações 5, 6, 7 e 8 respectivamente (BRAGA *et al.*, 1998; HAYKIN, 2001).

Função Degrau

$$\varphi(v_j) = \begin{cases} +k & \text{se } v_j \geq 0 \\ -k & \text{se } v_j < 0 \end{cases} \quad 5$$

Função Linear

$$\varphi(v_j) = \alpha v_j \quad 6$$

Função Sigmóide

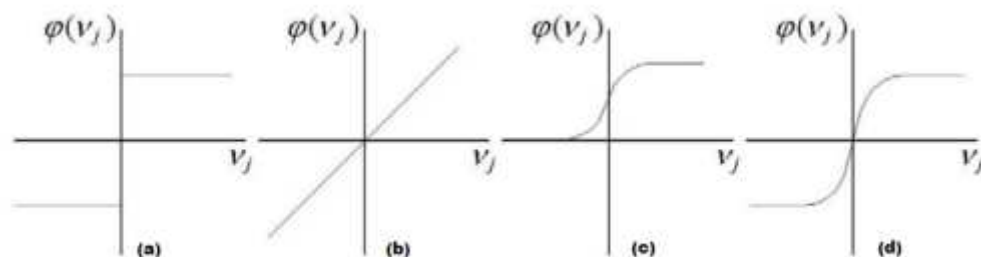
$$\varphi(v_j) = \frac{1}{1 + e^{-\alpha v_j}} \quad 7$$

Função Sigmóide Tangente Hiperbólica

$$\varphi(v_j) = \tanh\left(\frac{v_j}{2}\right) = \frac{1 - e^{-v_j}}{1 + e^{-v_j}} \quad 8$$

De acordo com Haykin, 2001, as funções de ativação (φ) são configuradas para operar em uma faixa de valores específica. A função degrau geralmente é definida com valor $k = 1$, embora possa ser redefinida para que saídas estejam dentro do intervalo de 0 a 1, conforme proposto por McCulloch e Pitts, 1943, em seu modelo de neurônio. No caso da função linear, a definição de $\alpha = 1$ permite que a saída do neurônio seja igual ao nível de atividade interna. As funções sigmóides são amplamente utilizadas em aplicações de RNAs devido à sua propriedade de serem estritamente crescentes, diferenciáveis, contínuas e não lineares (BRAGA *et al.*, 2007). As funções sigmóides logística e tangente produzem saídas entre 0 e 1, e entre -1 e 1, respectivamente. A Figura 14 evidencia os exemplos de funções de ativação.

Figura 14 – Exemplos de funções de ativação: (a) função degrau; (b) função linear; (c) função logística; (d) função tangente hiperbólica.



Fonte: Adaptada de Haykin (2001).

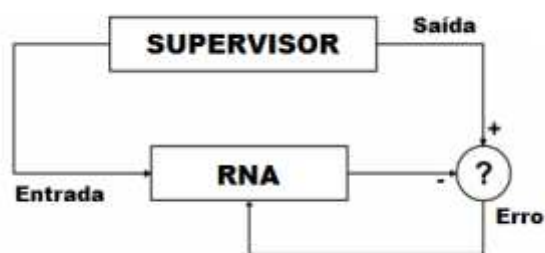
3.3.4 Processos de Aprendizagem

De acordo com Osório *et al.*, 2000, o aprendizado é a capacidade de adaptar, modificar e aprimorar as respostas e comportamentos, sendo um atributo fundamental de seres considerados inteligentes, sejam eles humanos ou não. A habilidade de generalização refere-se à capacidade de um sistema de “Aprendizado de Máquina” aprender por meio de exemplos e/ou dados apresentados (“adquirir experiência”), e ser capaz de fornecer as saídas corretas para entradas novas e não vistas anteriormente (BISHOP, 2006).

Os processos de aprendizagem ou treinamento baseia-se na apresentação de um conjunto de exemplos à RNA, a qual ajusta os seus pesos sinápticos por meio de iterações, buscando extrair características para representar as informações fornecidas (SILVA, 2010). Esses pesos sinápticos são armazenados em uma memória e são utilizados quando novos valores de entrada (novas amostras) são apresentados à rede, gerando resultados.

Os processos de aprendizagem podem ser classificados em supervisionados e não supervisionados. O aprendizado não supervisionado ocorre quando não há um “supervisor externo”, ou seja, apenas os padrões de entrada estão disponíveis para rede. No treinamento supervisionado, há um supervisor externo que fornece à rede tanto os valores de entrada disponíveis quanto os valores de saída esperados. A cada vez que uma entrada é apresentada à rede, verifica-se se a saída obtida após a simulação é igual ou muito próxima da saída desejada para aquela entrada (BARRETO, 1999). Caso haja uma diferença, a rede ajusta os pesos novamente até que o erro fique próximo do valor desejado, armazenando o conhecimento adquirido. Esse método de aprendizagem é amplamente utilizado na literatura, (BRAGA *et al.*, 2007) e foi utilizado neste estudo. A Figura 15 ilustra um exemplo de aprendizado supervisionado.

Figura 15 – Aprendizado supervisionado.



Fonte: Adaptada de Haykin (2001).

O aprendizado supervisionado pode ser dividido em treinamento dinâmico e treinamento estático. No treinamento dinâmico, acontecem alterações na estrutura da rede, como a redução ou aumento do número de camadas, o número de neurônios em cada camada e o número de conexões (BARRETO, 1999). Por outro lado, no treinamento estático, a única alteração que ocorre são os pesos sinápticos, e a arquitetura da rede permanece inalterada. Este tipo de treinamento foi empregado neste trabalho. Em resumo, considerando um neurônio de saída j , a cada iteração t , é gerado um sinal de erro $e_j(t)$, calculado como a diferença entre o valor de saída do neurônio $y_j(t)$ e o valor de saída desejado $d_j(t)$ (ou valor alvo), conforme demonstrado na Equação 9.

$$e_j(t) = d_j(t) - y_j(t) \quad 9$$

Dessa forma, a cada iteração, são feitas pequenas modificações nos pesos sinápticos, resultando em uma diminuição incremental do erro. O objetivo é convergir para o valor desejado e obter um erro menor na próxima iteração ($t + 1$):

$$w_{ji}(t + 1) = w_{ji}(t) + \Delta w_{ji}(t) \quad 10$$

Na próxima iteração, o peso sináptico $w_{ji}(t)$ da conexão entre o neurônio j e o neurônio i da camada anterior é ajustado através do valor $\Delta w_{ji}(t)$, resultando em um novo valor $w_{ji}(t + 1)$.

De acordo com Braga, 2007, os algoritmos de aprendizagem diferem principalmente na fórmula utilizada para calcular $w_{ji}(t)$. No entanto, uma forma

genérica de atualização dos pesos por correção de erros é apresentada na Equação 11:

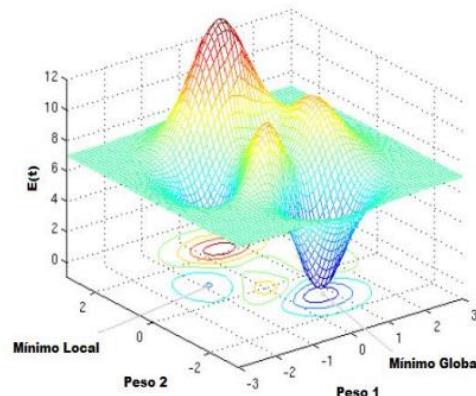
$$w_{ji}(t + 1) = w_{ji}(t) + \eta e_j(t) x_i(t) \quad 11$$

onde, $x_i(t)$ é a entrada para o neurônio i no tempo t e η é a taxa de aprendizado, que define a velocidade com que os pesos serão ajustados buscando o menor erro $e_j(t)$. A obtenção da equação de ajuste dos pesos em diferentes modelos de RNAs utiliza a soma dos erros quadráticos de todas as saídas como parâmetro de desempenho da rede e também como função de custo (Equação 12), que tem por objetivo ser minimizada pelo algoritmo de treinamento:

$$E(t) = \frac{1}{2} \sum_j (d_j(t) - y_j(t))^2 = \frac{1}{2} \sum_j e_j^2(t) \quad 12$$

Durante a fase de aprendizado, a função custo é uma medida de desempenho do sistema que depende dos pesos e *bias* da rede. Essa função pode ser representada como uma superfície de erro, onde os pesos e *bias* atuam como coordenadas (HAYKIN, 2001). Segundo Braga, 1998, quando são utilizadas funções de ativação lineares, a superfície de erro pode ter apenas um ponto mínimo, enquanto que, no caso de funções não-lineares, podem existir vários mínimos locais além do mínimo global. A Figura 16 ilustra uma superfície de erro com dois pesos, destacando tanto o mínimo local quanto o mínimo global, para uma melhor compreensão.

Figura 16 – Superfície de erro para dois pesos e exemplos de mínimo global e local.



Fonte: Adaptada de Raizer *et al.*, (2009).

Durante o processo de aprendizado, os pesos e *bias* da rede, que também são chamados de parâmetros livres, são ajustados de forma a se aproximarem de um ponto mínimo da superfície de erro. Essa direção de ajuste é determinada pelo gradiente da superfície de erro no ponto em questão (HAYKIN, 2001). Portanto, a minimização da função de custo (Equação 13), que é realizada através da regra delta, é conhecida como o método do gradiente para a minimização do erro. Nesse método, os pesos são ajustados no sentido oposto ao gradiente, conforme mostrado a seguir:

$$\Delta \vec{w} = -\eta \cdot \nabla E(\vec{w}) \quad 13$$

onde, conforme Haykin, 2001, \vec{w} é um vetor que possui todos os pesos e o *bias* do neurônio, $\nabla E(\vec{w})$ corresponde ao gradiente da função de custo (Equação 12) do vetor \vec{w} (com a mesma dimensão do vetor \vec{w}), e onde cada componente é a derivada parcial de $E(\vec{w})$ com relação a cada peso em \vec{w} .

Durante a etapa de aprendizado ou treinamento existem dois principais critérios de parada: o erro de saída da rede (método mais utilizado) ou o número de épocas (apresentação de todos os padrões de treinamento disponível correspondente a uma época). Segundo Haykin, 2001, um valor de η pode aumentar a velocidade do treinamento, porém também pode fazer com que um mínimo global não seja alcançado. Por outro lado, um valor baixo de η diminui a velocidade de treinamento, porém aumenta a probabilidade do mesmo em convergir para um mínimo global. Ainda conforme o autor, para determinar o momento de parada e avaliar o desempenho da rede são definidos três grupos com funções objetivos distintos sobre os dados empregados, denominados de conjunto de treinamento, validação e teste. Na sequência serão explicados os três grupos através das considerações do autor.

O conjunto de treinamento refere-se à parcela de dados que será utilizado para o aprendizado da RNA, no qual a partir deles o modelo fará o ajuste dos pesos. O recomendado é que este conjunto contenha a maior parcela de dados comparado aos outros dois grupos, pois definirá o funcionamento da *MLP* e desta forma o modelo reconhecerá a maior quantidade de padrões existente entre os dados.

O conjunto de validação é usado para avaliar o treinamento e determinar a parada dele, porém não é empregado na aprendizagem e utiliza-se dados fora da amostra de treinamento. Este conjunto tem o objetivo de evitar problemas de *Overfitting* (excesso de ajuste, neste caso a rede não extrai padrões e somente memoriza os dados).

O conjunto de teste é muito semelhante ao grupo de validação, entretanto ele somente é empregado após toda etapa de treinamento ter sido finalizada. O conjunto tem como objetivo avaliar a performance, através da medida de erro escolhido, do sistema já concluído.

Outro problema que pode ocorrer durante o aprendizado, além do *Overfitting*, é o *Underfitting*, definido pela incapacidade do modelo de aprendizado em generalizar o problema que lhe foi exposto (BRAGA *et al.*, 1998). Resumidamente, o modelo de aprendizado não consegue encontrar a relação (padrões) entre os valores de entradas e as saídas desejadas (valor alvo) durante o tempo de treinamento. A Figura 17 apresenta um exemplo para a determinação de parada de treinamento de uma RNA do tipo *MLP* usando os erros do conjunto de treinamento e validação (ANDRADE FILHO, 2016).

Figura 17 – Exemplos de erros em treinamento de uma rede do tipo MLP.



Fonte: Adaptada de Andrade Filho (2016).

Conforme demonstrado na Figura 17 e destacado pelo autor, o ponto de cor verde indica a situação de *Underfitting*, onde ambos os erros (treinamento e

validação) estão altos e ainda decaindo. O ponto em vermelho representa a situação de *Overfitting*, na qual o erro para o conjunto de treinamento é baixo, porém o erro do conjunto de validação é alto, tendo este crescido depois de determinado ponto. O ponto indicado na cor azul representa a localização na qual o treinamento deve ser finalizado, pois a partir deste ponto o erro de validação começa a aumentar, demonstrando que o modelo está perdendo capacidade de generalização do problema.

Haykin, 2001, destaca sobre a importância da divisão dos dados (treinamento, validação e teste) e o problema envolvido, pois cada amostra pode conter representatividades diferentes o que influenciará diretamente no erro.

Para definir a capacidade de generalização de uma *MLP* um fator muito importante é o método pelo qual é realizado o ajuste ou calibração dos pesos (REIDMILER, 1994). Para esta tarefa foi considerado neste estudo o método *backpropagation* (retropropagação), sendo este um algoritmo de treinamento supervisionado para redes do tipo MLP, que segundo diversos autores (BARRETO, 1999; HAYKIN, 2001; BRAGA *et al.*, 2007; SILVA *et al.*, 2010), é um dos mais empregados e difundidos na literatura.

3.3.5 O algoritmo *Backpropagation* (Retropropagação)

O algoritmo *backpropagation* (*BP*) foi descrito de uma forma clara por Rumelhart *et al.*, 1986 para ser utilizado no treinamento supervisionado de redes *feedforward* com múltiplas camadas (WASSERMAN, 1989). Este algoritmo divide-se em duas fases, na primeira fase, denominada *forward*, uma entrada é aplicada à rede neural e seu sinal funcional é propagado pela rede, camada a camada até a saída, mantendo seus pesos fixos. Na segunda fase, denominada *backward*, a saída é comparada com o alvo gerando um sinal de erro e por sua vez este sinal é propagado da saída para a entrada, ajustando os pesos com o objetivo de minimizar o erro (RUMELHART *et al.*, 1986; HAYKIN, 2001; BRAGA, 2007). O algoritmo *backpropagation* é baseado na regra delta (WIDROW e HOFF, 1960), porém ele não ajusta somente pesos da camada de saída e sim de todas as camadas da rede.

Conforme descrito por Braga *et al.*, 2007, a diferença entre os algoritmos de aprendizado encontra-se no modo pelo qual é calculado o ajuste dos pesos ($\Delta w_{ji}(t)$). Apresentam-se neste estudo as principais equações do algoritmo *backpropagation*, pois a dedução completa do modelo foge do objetivo deste trabalho, sendo possível verificar a dedução completa em Rumelhart *et al.*, 1986, Haykin, 2001 ou Braga *et al.*, 2007.

Neste algoritmo o ajuste dos pesos baseia-se no método do gradiente para minimização de erro (Equação 13), isto é, a correção aplicada a cada peso sináptico do neurônio j de saída se dá através da derivada parcial da função de custo $E(t)$, (Equação 12), em relação ao peso analisado, conforme demonstrado nas Equações 14 e 15:

$$\frac{\partial E(t)}{\partial w_{ji}} = -e_j(t) \cdot \varphi'_j(v_j(t)) \cdot x_i(t) \quad 14$$

$$\Delta w_{ji}(t) = -\eta \frac{\partial E(t)}{\partial w_{ji}} = \eta \cdot e_j(t) \cdot \varphi'_j(v_j(t)) \cdot x_i(t) \quad 15$$

onde, $e'_j(v_j(t))$ é a derivada da função de ativação aplicada ao nível de atividade interna do neurônio. Agrupando os termos de sinal de erro, $e_j(t)$, correspondente para aquele neurônio e a derivada da função de ativação associada, $e'_j(v_j(t))$, obtêm-se o chamado gradiente local, δ_j , que aponta para as modificações necessárias nos pesos sinápticos (HAYKIN, 2001):

$$\delta_j(t) = e_j(t) \cdot \varphi'_j(v_j(t)) \quad 16$$

Desta forma, reescrevendo a Equação 15, obtêm-se o ajuste dos pesos de um neurônio j qualquer da camada de saída:

$$\Delta w_{ji}(t) = -\eta \frac{\partial E(t)}{\partial w_{ji}} = \eta \cdot \delta_j \cdot x_i(t) \quad 17$$

Após o cálculo do erro propagado até a saída (fase *forward*) os neurônios da camada intermediária têm sua parcela de responsabilidade sobre o ocorrido, desta forma dá-se início a fase *backward*, para que os mesmos sejam calibrados e minimizados através da retropropagação (BRAGA *et al.*, 1998). Considerando agora que neurônio intermediário corresponde ao neurônio j , conectado a neurônios k da camada de saída, o gradiente local é calculado através da Equação 18:

$$\delta_j(t) = \varphi'_j(v_j(t)) \cdot \sum_k \delta_k(t) w_{kj}(t) \quad 18$$

onde, $\sum_k \delta_k(t)$ é o somatório dos gradientes locais de cada neurônio k da camada seguinte, $w_{kj}(t)$ é a ponderação das conexões sinápticas que ligam cada neurônio k ao neurônio j e $\varphi'_j(v_j(t))$ é a derivada da função de ativação aplicada ao nível de atividade interna do neurônio. Desta forma, o ajuste dos pesos da camada oculta é realizado através da substituição da Equação 18 na Equação 17:

$$\Delta w_{ji}(t) = \eta \cdot \varphi'_j(v_j(t)) \cdot \sum_k \delta_k(t) w_{kj}(t) \cdot x_i(t) \quad 19$$

Com o objetivo de aumentar a velocidade de aprendizagem sem tornar a rede instável e sem aumentar a taxa de aprendizagem, η , emprega-se um método que modifica a Equação 13 através da adição do um termo de *momentum* (β), proposto por (RUMELHART *et al.*, 1986) em seu modelo:

$$\Delta w_{ji}(t) = -\eta \frac{\partial E(t)}{\partial w_{ji}} + \beta \Delta w_{ji}(t - 1) \quad 20$$

onde, o termo *momentum* relaciona as alterações dos pesos na iteração t com as realizadas na iteração anterior ($t - 1$). Segundo o autor, a adição deste termo dificulta a mudança de tendência na atualização dos pesos, além de ser empregada para minimizar a probabilidade da ocorrência de mínimos locais.

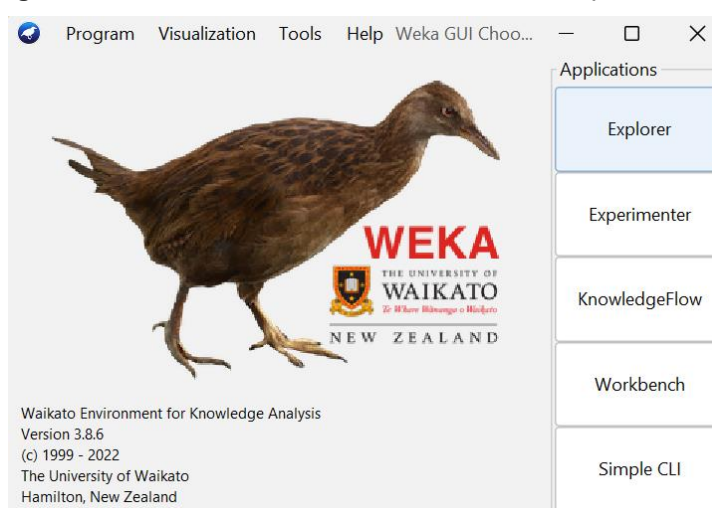
3.4 Weka

O *WEKA* é uma abrangente coleção de algoritmos de aprendizado de máquina projetados para funções de mineração de dados. Ele engloba diversas ferramentas essenciais para o processo de mineração de dados, incluindo a preparação dos dados de entrada (MENOTTI, 2017).

Desenvolvido na Universidade de Waikato, na Nova Zelândia, o nome *WEKA* é a abreviação de “*Waikato Environment for Knowledge Analysis*”, que traduzido para o português significa “Ambiente Waikato para Análise de Conhecimento”. O *software* é construído em Java e é compatível com praticamente todas as plataformas (WITTEN; FRANK; HALL, 2016).

A interface inicial do programa, na versão 3.8.6, oferece quatro opções para os usuários escolherem: *Explorer*, *Knowledge Flow*, *Experimenter* e *Workbench*. A Figura 18 ilustra essa interface inicial.

Figura 18 – Interface inicial do software WEKA (versão 3.8.6).



Fonte: Adaptado de WEKA (2023).

A interface *Experimenter* facilita a seleção dos melhores métodos e parâmetros para resolver um problema específico. Ela possibilita a automatização do

processo de classificação com diversas configurações de parâmetros para um conjunto de dados, além de coletar estatísticas de desempenho e realizar testes de significância (WITTEN; FRANK; HALL, 2016). Essa interface é especialmente útil para estudos em larga escala e permite que as análises de desempenho coletadas sejam revisadas posteriormente (FRANK *et al.*, 2016).

Existem três principais formas de utilização do *WEKA* (WITTEN; FRANK; HALL, 2016):

1. Comparação de desempenho de diversos métodos de aprendizagem;
2. Utilização dos modelos de aprendizagem para realizar previsões em novos casos;
3. Adoção de métodos de aprendizagem para inspecionar as saídas e coletar informações relevantes.

Com sua vasta gama de algoritmos e ferramentas, o *WEKA* tem se mostrado uma poderosa ferramenta para a análise e mineração de dados, oferecendo suporte a pesquisadores, cientistas de dados e profissionais que trabalham com aprendizado de máquina e tarefas de mineração de dados em geral.

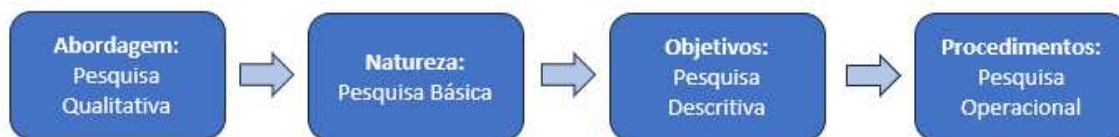
4 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, serão descritos e analisados os procedimentos metodológicos adotados para a execução deste trabalho. Será apresentado um panorama abrangente das estratégias e etapas utilizadas na condução da pesquisa, com intuito de assegurar uma compreensão aprofundada das abordagens empregadas.

4.1 Tipo de Pesquisa

Os procedimentos metodológicos desenvolvidos neste estudo são descritos detalhadamente e apresentados no fluxograma ilustrativo exibido na Figura 19 (GIL, 2010). Este fluxograma representa a sequência de etapas que foram seguidas no estudo, fornecendo uma visão geral do planejamento e execução do estudo.

Figura 19 – Procedimentos metodológicos.



Fonte: Adaptada de Gil (2010).

Onde:

- Pesquisa Quantitativa: caracteriza-se pelo uso de dados numéricos históricos. Nessa abordagem, são aplicadas estatísticas, tais como cálculo da média, desvio padrão, coeficiente de variação, entre outras, para analisar os dados coletados;
- Pesquisa Básica: tem como propósito analisar conhecimentos que contribuam para o avanço da ciência. Ou seja, busca-se ampliar o entendimento sobre determinado assunto, explorando conceitos, teorias e princípios que possam fornecer bases sólidas para futuras investigações e aplicações práticas;
- Pesquisa Descritiva: destaca-se por analisar as relações entre os dados reais coletados. Por meio de técnicas estatísticas e análise exploratória, procura-se

descrever, sumarizar e interpretar os dados de forma a identificar padrões, tendências e associações existentes entre as variáveis estudadas;

- Pesquisa Operacional: busca soluções e problemas específicos utilizando modelos matemáticos ou estatísticos. Essa abordagem busca otimizar recursos, tomar decisões eficientes e encontrar soluções viáveis para problemas complexos por meio de técnicas como programação linear, simulação e análise de sensibilidade.

4.2 Séries Históricas

Para o estudo de caso, foi utilizado uma série histórica de geração de energia elétrica hidráulica, no período de janeiro/1996 a agosto/2022, referentes ao Sistema Nacional Integrado (SIN) (IPEA, 2023).

A Tabela 4 apresenta, como exemplo, uma parte do conjunto de dados de geração hidráulica (IPEA, 2023).

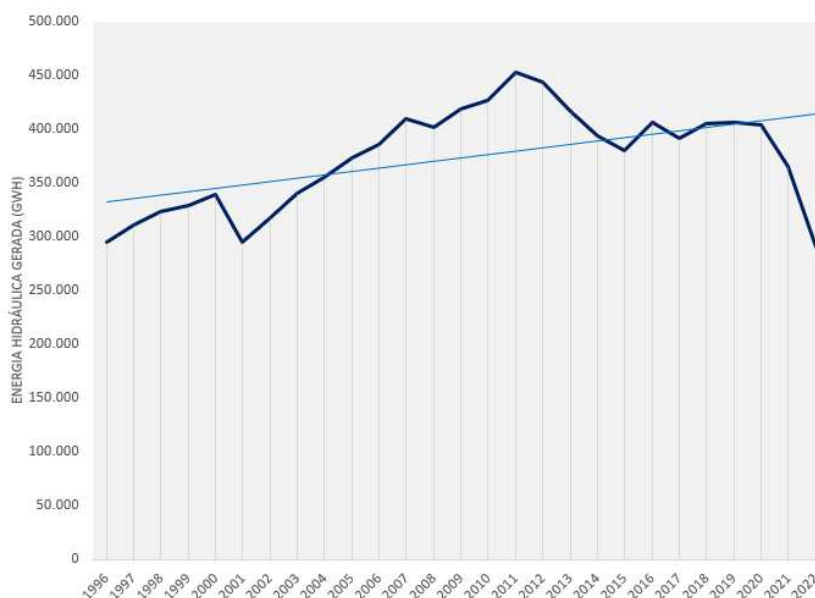
Tabela 4 – Parte do conjunto de dados de geração hidráulica.

	2016	2017	2018	2019	2020	2021	2022
Janeiro	35075	39402	38112	42374	38037	34279	38280
Fevereiro	35975	37493	36002	36668	38356	34986	37278
Março	38397	39726	41040	39601	40683	39892	42154
Abril	35852	33076	36409	37848	32338	34467	38221
Mai	34009	32965	33331	36401	31509	32985	36145
Junho	31929	31580	28925	31435	30007	27222	33534
Julho	32431	28771	28400	29619	32468	24889	32880
Agosto	32072	27744	28616	27553	31901	23901	31780
Setembro	30438	27634	27147	28853	33755	24943	-
Outubro	31798	28955	33104	31310	32916	25426	-
Novembro	31817	29401	32253	30701	29210	28733	-
Dezembro	36225	34153	39066	33743	32450	33773	-

Fonte: Autoria própria (2023).

A representação gráfica da série histórica, de geração de energia hidráulica total por ano, pode ser observada na Figura 20.

Figura 20 – Série Histórica de Energia Hidráulica.



Fonte: Autoria própria (2023).

É importante destacar que no gráfico da série histórica (Figura 20) também é apresentada uma linha de tendência. Além disso, é relevante observar que os dados de geração de energia hidráulica para o ano de 2022 não estão disponíveis para os meses de setembro a dezembro, o que resulta em um montante gerado menor em comparação aos demais anos.

4.3 Etapas do Trabalho

Para desenvolver os modelos, os modelos de Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais do tipo *MLP*, foram executadas, neste trabalho, três etapas:

4.3.1 Coleta, análise e transformação de dados

A primeira etapa, deste estudo, consistiu na análise de dados. Nesta etapa inicial, foram realizados procedimentos de pré-processamento dos dados, com objetivo de prepará-los para serem utilizados nos algoritmos de modelagem. Em seguida, foi realizada uma análise exploratória dos dados obtidos do IPEA, visando compreender suas características e identificar possíveis padrões ou tendências.

Na sequência, na etapa de transformação dos dados, estes, após processados, foram exportados para um arquivo em um formato aceito pelo sistema de mineração de dados escolhido.

4.3.2 Treinamento e Validação

Na segunda etapa deste estudo, denominada treinamento e validação, foram selecionados modelos de Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais do tipo *MLP*. Para isso, foram utilizadas métricas de avaliação que permitissem comparar o desempenho dos modelos.

4.3.3 Teste

Na última etapa deste estudo, a etapa de teste, os modelos outrora citados foram submetidos a um conjunto de dados separado, que não foi utilizado durante as etapas de treinamento e validação. O objetivo dessa etapa foi o de testar o desempenho dos modelos em dados completamente novos, verificando sua capacidade de generalização.

4.4 Arquivo *ARFF* – *WEKA*

Neste contexto de pesquisa, os modelos de Regressão Linear Múltipla e Redes Neurais Artificiais, ambos de relevância crítica para o estudo em questão,

foram cuidadosamente selecionados e implementados com o auxílio do *software WEKA*.

É importante enfatizar que o *WEKA* opera utilizando arquivos no formato *ARFF (Attribute-Relation File Format)*, que é uma estrutura padrão amplamente aceita para representar conjuntos de dados e seus atributos. Essa formatação é de suma importância, pois simplifica a manipulação dos dados de entrada e estabelece a base necessária para garantir a interação precisa e eficaz entre os modelos escolhidos e o conjunto subjacente.

A Figura 21 ilustra de forma clara a relação intrínseca entre o *software WEKA* e o formato *ARFF*. Esta representação visual destaca a importância fundamental da formatação adequada dos dados, fornecendo uma visão abrangente da interação harmoniosa entre o *WEKA* e os arquivos *ARFF* durante a fase de análise e modelagem. Essa integração é de importância crítica para assegurar o sucesso da análise e o correto funcionamento dos modelos de Regressão Linear Múltipla e Redes Neurais Artificiais no contexto deste estudo. Portanto, a escolha meticulosa da ferramenta e a formatação adequada dos dados são elementos essenciais para garantir a validade e a eficácia da pesquisa.

Figura 21 – Arquivo ARFF.

```
@relation 'Energia'
@attribute Data date "yyyy-MM"
@attribute Geracao numeric

@data

1996-01,24.203
1996-02,23.745
1996-03,25.426
1996-04,24.426
1996-05,24.627
1996-06,23.569
1996-07,24.823
1996-08,24.984
1996-09,23.973
1996-10,25.077
1996-11,24.391
1996-12,25.222
1997-01,25.917
1997-02,23.853
1997-03,26.139
1997-04,25.140
1997-05,25.353
1997-06,24.530
1997-07,26.180
1997-08,26.252
```

Fonte: Autoria própria (2023).

Observa-se na Figura 21, que o arquivo, do *software WEKA*, é formado por algumas seções. A seção *@relation* contém o nome do arquivo, a segunda seção é a *@attribute*, onde os atributos são declarados, podendo ser do tipo *numeric*, *date* ou *string*, e por último *@data*, que marca o início dos dados que serão analisados pelos algoritmos de mineração de dados.

4.5 Aba *Forecast*

Neste estudo, foi empregada a funcionalidade “*Forecast*” do *software WEKA*, uma seção dedicada à realização de previsões em séries temporais. Essa ferramenta específica é projetada para facilitar a análise e a modelagem de dados temporais, permitindo a previsão de valores futuros com base em padrões e tendências identificadas no histórico da série.

Dentro da aba “*Forecast*”, os usuários têm acesso a duas áreas distintas: “*Basic Configuration*” e “*Advanced Configuration*”. A seção “*Basic Configuration*” é destinada a oferecer opções de configuração e ajustes simples, permitindo aos usuários realizarem previsões iniciais com facilidade e rapidez, adequadas para análises iniciais de dados temporais.

Por outro lado, a seção “*Advanced Configuration*” proporciona opções mais sofisticadas e detalhadas para ajustes finos do modelo de previsão. Essas configurações avançadas permitem maior flexibilidade na modelagem, podendo ser úteis quando se deseja explorar nuances mais complexas dos dados ou quando é necessário ajustar o modelo de previsão com maior precisão.

4.5.1 Área *Basic Configuration*

Em “*Basic Configuration*”, foram realizadas configurações relevantes para a análise de previsões de séries temporais, conforme ilustra a Figura 22.

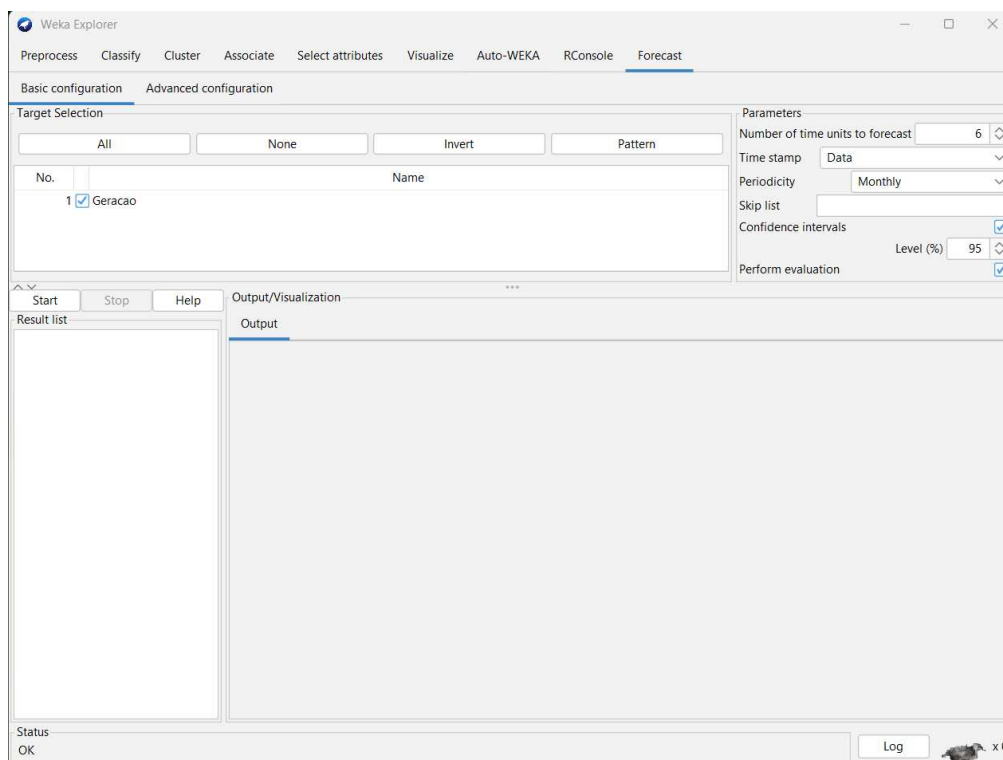
Na seção “*Parameters*”, foram efetuadas definições cruciais para o processo de previsão. O horizonte de previsão, estabelecido em 6 meses para este trabalho, foi configurado na área denominada “*Number of times units*”. Isso implica que as previsões visam antecipar eventos e valores futuros em um período de seis meses a partir da data presente.

Os campos “*Time stamp*” e “*Periodicity*” foram utilizados para registrar informações fundamentais sobre a estrutura temporal dos dados. A entrada “*Data*” foi empregada no campo “*Time stamp*” para indicar a data associada a cada observação, enquanto “*Monthly*” foi registrado no campo “*Periodicity*” para indicar a frequência mensal dos dados temporais.

Ademais, duas caixas de seleção foram ativadas para cumprir com determinados requisitos da análise. A opção “*Confidence intervals*” foi habilitada e seu valor mantido em 95%, permitindo a avaliação da incerteza associada às previsões, por meio da definição de um intervalo de confiança de 95%. Isso possibilita estimar a margem de erro em torno das previsões obtidas.

Da mesma forma, a caixa “*Perform Evaluation*” foi assinalada, o que indica a necessidade de realizar a avaliação dos resultados da previsão. Essa avaliação é essencial para medir a qualidade e precisão das previsões geradas pelos modelos, possibilitando a análise comparativa do desempenho entre diferentes abordagens de previsão.

Figura 22 – Área *Basic Configuration*.



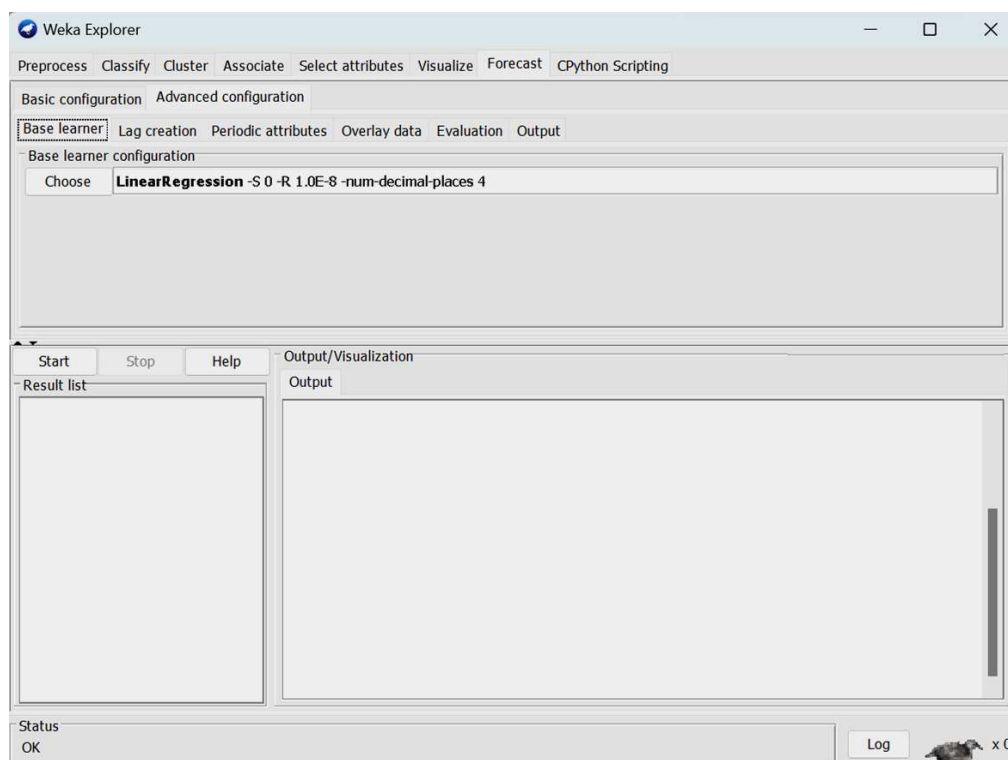
Fonte: Autoria própria (2023).

Após a conclusão das configurações na seção “*Basic Configuration*”, o próximo passo é a transição para a área denominada “*Advanced Configuration*”. Nessa etapa, têm-se acesso a opções mais sofisticadas e detalhadas para o ajuste refinado do modelo de previsão em séries temporais.

4.5.2 Área *Advanced Configuration*

Em “*Advanced Configuration*” (Figura 23), foram realizadas configurações específicas para este trabalho, nas quais quatro das seis abas disponíveis foram selecionadas e ajustadas. As abas configuradas são descritas detalhadamente a seguir:

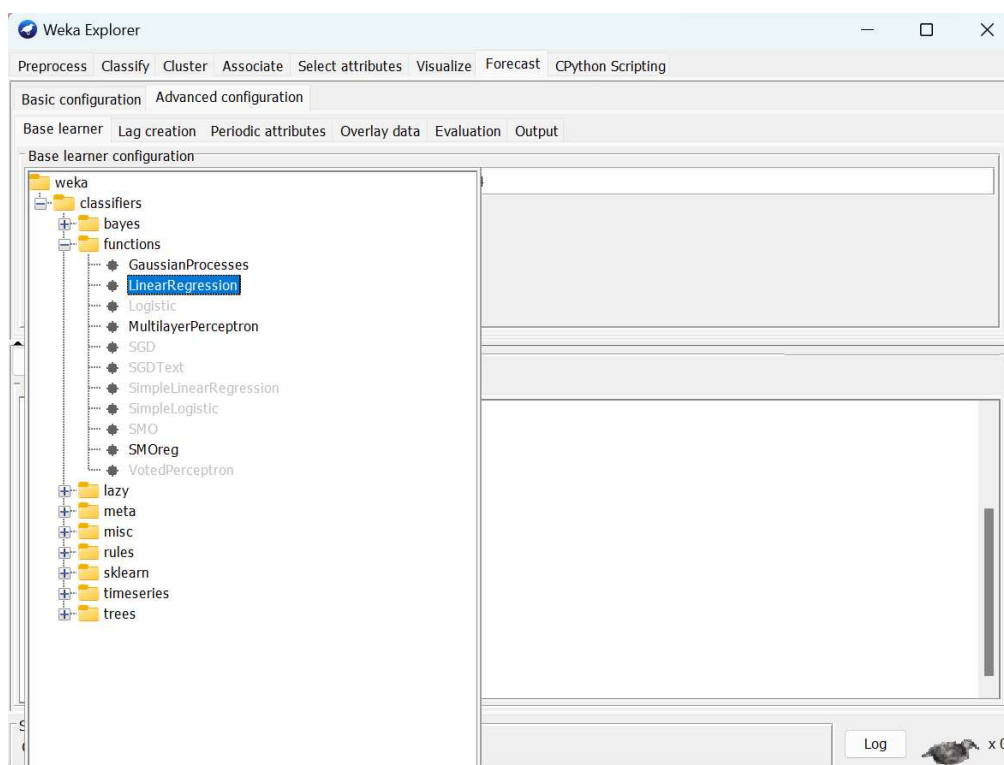
Figura 23 – Área *Advanced Configuration*.



Fonte: Autoria própria (2023).

Base Learner: Nesta aba é fornecida a possibilidade de selecionar o método de previsão a ser adotado no estudo em questão. Neste trabalho específico, foram escolhidos e aplicados dois métodos distintos de previsão: Regressão Linear Múltipla (*MLR*) e Redes Neurais Artificiais do tipo *Multilayer Perceptron (MLP)* (Figura 24).

Figura 24 – Algoritmos de previsão.



Fonte: Autoria própria (2023).

Lag Creation: Nesta aba foram definidas as configurações de atraso (*lag*) para a análise das séries temporais. No âmbito deste estudo, optou-se por manter as configurações originais do *WEKA*, que, neste caso, consistiram em um atraso de 7 dias.

Ao adotar um atraso de 7 dias, a análise leva em consideração a influência das observações ocorridas na semana anterior para prever os valores futuros. Esse atraso pode ajudar a capturar padrões semanais ou sazonalidades que ocorram regularmente em intervalos de 7 dias.

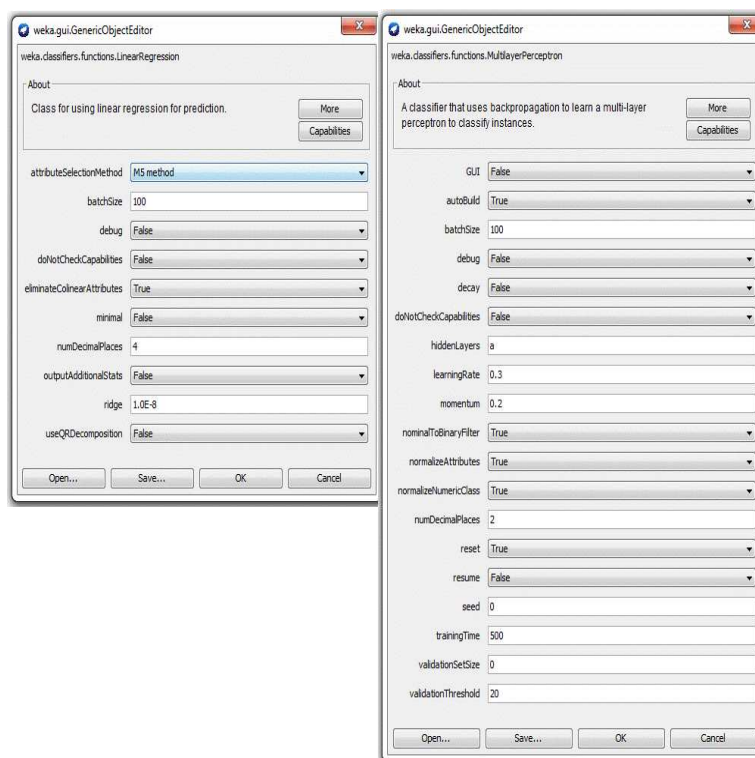
Evaluation: Nesta aba são definidas as métricas utilizadas para avaliar o desempenho dos modelos de previsão. Neste estudo, três métricas específicas foram selecionadas: *MAE* (*Mean Absolute Error*), *RMSE* (*Root Mean Square Error*) e *MAPE* (*Mean Absolute Percent Error*). Essas métricas são amplamente empregadas na avaliação de modelos de previsão e fornecem informações valiosas sobre a acurácia das previsões realizadas.

Além disso, nesta mesma aba, é configurada a divisão dos dados entre o conjunto de treinamento e o conjunto de validação. Para este trabalho específico, a proporção selecionada foi de 70% para o conjunto de treinamento e 30% para o

conjunto de validação. Essa divisão permite o treinamento do modelo em uma parte dos dados e a avaliação do desempenho do modelo em outra parte, garantindo uma análise mais confiável e imparcial do seu desempenho em dados não utilizados durante o treinamento.

As telas de configurações dos dois algoritmos escolhidos são apresentadas na Figura 25.

Figura 25 – Janelas de configuração dos algoritmos.



Fonte: Autoria própria (2023).

Nestas janelas (Figura 25), é possível definir os hiperparâmetros específicos para cada um dos algoritmos utilizados na análise e previsão das séries temporais. Os hiperparâmetros são parâmetros que não são aprendidos diretamente pelo modelo, mas que afetam seu desempenho e comportamento durante o processo de treinamento e previsão.

Ao ajustar os hiperparâmetros de forma criteriosa, é possível aprimorar a performance dos algoritmos e otimizar as previsões realizadas. Essas configurações podem impactar significativamente os resultados obtidos, tornando-as uma etapa essencial para alcançar previsões mais precisas e confiáveis.

4.6 Métricas

As métricas que foram utilizadas neste trabalho foram: o Erro Médio Absoluto Percentual (*MAPE*), a Raiz Quadrada do Erro Médio Quadrático (*RMSE*) e o Erro Médio Absoluto (*MAE*). O *RMSE* mede a magnitude média do erro. Uma vez que os erros são elevados ao quadrado antes de serem calculados, o *RMSE* dá um peso relativamente alto a erros grandes, o *MAE* representa o desvio médio entre os valores observados e previstos pelo modelo. No *MAE* todas as diferenças individuais são ponderadas igualmente na média. Já o *MAPE* mede o tamanho do erro em termos percentuais. (CANKURT; SUBASI, 2015; PINHEIRO, 2020).

As equações, das métricas *MAE*, *RMSE* e *MAPE*, são apresentadas na Tabela 5.

Tabela 5 – Equações: *MAE*, *RMSE* e *MAPE*.

Sigla	Equação
<i>MAE</i>	$\frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $
<i>RMSE</i>	$\sqrt{\frac{1}{n} \sum_{i=1}^n ((y_i - \hat{y}_i)^2)}$
<i>MAPE</i>	$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)/y_i \times 100$

Fonte: Autoria própria (2023).

Onde: y_i é o valor do período i , \hat{y}_i é a previsão para o período i e o n é o número de observações.

Para cada etapa do processo de previsão, estas métricas foram utilizadas para se saber qual o nível de confiança de cada um dos modelos.

5 RESULTADOS

Inicialmente, neste capítulo, faz-se uma breve análise dos dados obtidos do Instituto de Pesquisa Econômica Aplicada (IPEA). Posteriormente, apresentam-se os desfechos concernentes aos processos de Treinamento e Validação dos modelos de Regressão Linear Múltipla (*MLR*) e de Redes Neurais Artificiais do tipo *Multilayer Perceptron* (*MLP*). Na fase subsequente, a etapa de Teste, ocorre uma comparação entre os dados provenientes das previsões e os dados reais, empregando métricas de avaliação.

5.1 Análise Descritiva dos Dados

Na Tabela 6 apresenta-se a análise descritiva dos dados adquiridos do Operador Nacional do Sistema Elétrico (ONS).

Tabela 6 – Análise descritiva dos dados.

Resumo descritivo	Geração
Instância(s)	314
Média (GWh)	31,385
Mínimo (GWh)	20,708
Máximo (GWh)	42,775
Desvio Padrão (GWh)	4,45
Coeficiente de Variação (%)	14,179

Fonte: Elaboração própria a partir de dados do ONS (2023).

A análise dos dados, apresentados na Tabela 6, revela que a geração média durante o período de estudo foi de 31,385 GWh. No mesmo intervalo temporal, observa-se tanto uma geração mínima quanto uma máxima de 20,708 GWh e 42,775 GWh, respectivamente. Além disso, pode-se inferir, a partir da Tabela 6, um coeficiente de variação de 14,179%, o qual é classificado como moderado, de acordo com a classificação de Pimentel (2009). Vale mencionar que um coeficiente de variação mais elevado corresponde a uma maior amplitude de variação nos dados analisados.

5.2 Treinamento e Validação

No âmbito deste estudo, com finalidade de alcançar uma modelagem otimizada da geração de energia hidráulica, diversas configurações arquiteturais e hiperparâmetros foram submetidos à avaliação. Os algoritmos empregados passaram por um processo de treinamento empregando 220 amostras (correspondentes a 70% do conjunto total) e uma fase de validação com 94 amostras restantes (representando 30% do total de dados).

Destaca-se que todos os algoritmos adotados neste trabalho empregaram a técnica de ajuste de variância. Essa técnica se revela essencial para mitigar as flutuações presentes nos dados de entrada, conferindo maior robustez e estabilidade aos modelos desenvolvidos.

5.2.1 Regressão Linear Múltipla (MLR)

No que tange ao algoritmo de Regressão Linear Múltipla, procedeu-se a uma abordagem na qual as configurações padrão foram mantidas, enquanto três parâmetros específicos foram alvo de modificações: *batchSize*, *numDecimalPlaces* e *ridge*.

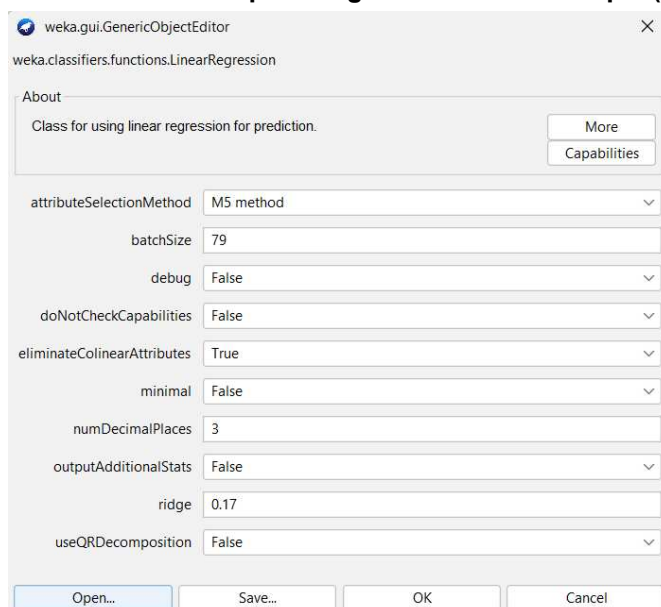
O parâmetro *batchSize*, que inicialmente possuía um valor 100, foi ajustado. Este parâmetro, o qual denota o número de instâncias processadas em uma única iteração, foi recalibrado para 79. Este valor foi obtido pela divisão do total de 314 instâncias, presentes no arquivo *ARFF*, por 4, visando uma adequada avaliação do modelo de Regressão Linear.

Paralelamente, o parâmetro *numDecimalPlaces*, inicialmente estabelecido em 4, que representa o número de casas decimais após o ponto decimal, foi sincronizado com a precisão dos dados de geração hidráulica, limitando-se a 3 casas decimais.

Adicionalmente, o parâmetro *ridge*, que inicialmente adotava o valor padrão de 1×10^{-8} , um valor que auxilia na mitigação da complexidade do modelo de aprendizado, foi recalibrado para $1,7 \times 10^{-1}$. Este novo valor foi alicerçado em uma

série de simulações meticulosas. Com essas considerações, o modelo de Regressão Linear foi parametrizado como delineado na Figura 26.

Figura 26 – Parâmetros para Regressão Linear Múltipla (MLR).



Fonte: Adaptado de WEKA (2023).

Os resultados das métricas para esta específica combinação de parâmetros são expostos na Figura 27. As métricas em questão são vinculadas à fase de Validação.

Figura 27 – Resultados – Métricas (MLR).

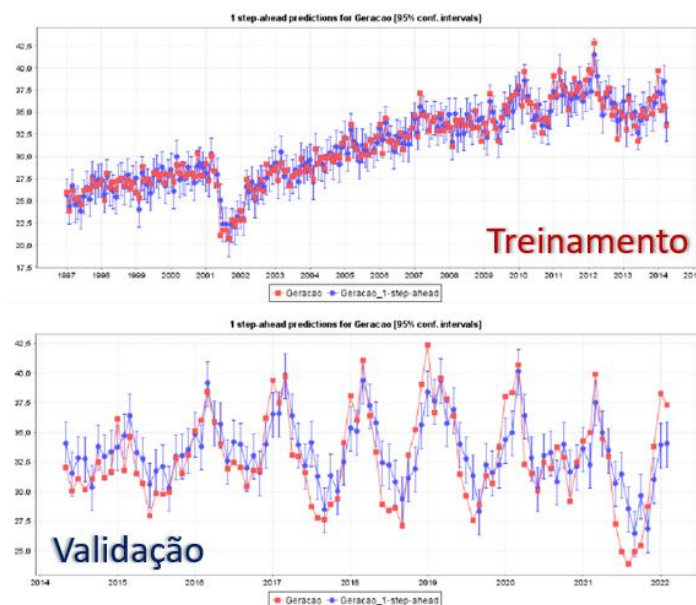
=== Evaluation on test data ===						
Target	1-step-ahead	2-steps-ahead	3-steps-ahead	4-steps-ahead	5-steps-ahead	6-steps-ahead

Geracao						
N	94	93	92	91	90	89
Mean absolute error	1.9455	2.7834	3.2377	3.4078	3.4367	3.4335
Mean absolute percentage error	6.1495	8.8419	10.3304	10.9008	11.0031	11.0059
Root mean squared error	2.3416	3.3356	3.8529	4.0233	4.0499	4.0452
Total number of instances: 94						

Fonte: Adaptado de WEKA (2023).

A Figura 28 exibe os resultados alcançados por meio da aplicação da Regressão Linear Múltipla nos conjuntos de Treinamento e Validação. Essa avaliação é conduzida com consideração a um intervalo de confiança de 95%.

Figura 28 – Geração – Treinamento e Validação (MLR).



Fonte: Adaptado de WEKA (2023).

Ao analisar a Figura 28, torna-se evidente que os resultados exibiram uma discrepância que, embora não seja marcante, permanece perceptível em comparação com os valores reais. Essa disparidade justifica a obtenção dos valores correspondentes às métricas em análise.

5.2.2 Redes Neurais *Multilayer Perceptron* (MLP)

A Tabela 7 fornece uma descrição abrangente das várias combinações que foram cuidadosamente selecionadas e empregadas durante o processo de busca dos melhores parâmetros destinados a otimizar o desempenho do modelo de *Perceptron* de Múltiplas Camadas (MLP). Essas combinações de parâmetros foram meticulosamente escolhidas para abranger uma ampla gama de configurações possíveis, permitindo uma análise abrangente do impacto de diferentes valores nos resultados obtidos. Através dessa abordagem estruturada, procurou-se identificar configurações que mais se alinham com os dados e o problema em questão, visando obter um modelo MLP altamente eficaz em termos de previsão.

Tabela 7 – Combinações das configurações para o modelo *MLP*.

Resumo descritivo	Geração
<i>Neurons</i>	[1,3,5,7,9]
<i>Hidden Layers</i>	[1,2,3,4]
<i>Activation</i>	Sigmoid
<i>Learning Rate</i>	[0.1,0.01,0.001]
<i>Momentum</i>	[0.1,0.3,0.5,0.7,0.9]
<i>Epochs</i>	[2000,1500,1000,500,100]

Fonte: Autoria própria (2023).

No Quadro 1 estão disponíveis as definições correspondentes aos termos empregados na Tabela 7.

Quadro 1 – Descrição dos parâmetros do modelo *MLP*.

Termos	Descrição
<i>Neurons</i>	Trazem mais ou menos complexidade para o modelo à medida em que são adicionados ou removidos.
<i>Hidden Layers</i>	Define o número de camadas ocultas da rede neural.
<i>Activation</i>	Função de ativação utilizada pelos neurônios.
<i>Learning Rate</i>	Indica o ritmo com que os pesos são atualizados.
<i>Momentum</i>	Utilizado para ajustes nos pesos da rede. Seu valor varia entre 0 e 1.
<i>Epochs</i>	Quantidade de vezes em que o conjunto de treinamento é acessado pelo algoritmo de aprendizado para atualizar o peso dos neurônios.

Fonte: Autoria própria (2023).

A Tabela 8 exibe as combinações mais eficazes de parâmetros identificadas para o modelo de *Perceptron* de Múltiplas Camadas (*MLP*).

Tabela 8 – Configuração do modelo *MLP*.

Parâmetros	Valores Definidos
<i>Neurons</i>	3,3,4,8
<i>Hidden Layers</i>	4
<i>Activation</i>	Sigmoid
<i>Learning Rate</i>	0.07
<i>Momentum</i>	0.8
<i>Epochs</i>	550

Fonte: Autoria própria (2023).

A Figura 29 exibe os parâmetros de entrada que se revelaram mais eficazes para a configuração do modelo de *Perceptron* de Múltiplas Camadas (*MLP*).

Figura 29 – Parâmetros para *Perceptron* de Múltiplas Camadas (*MLP*).



Fonte: Adaptado de WEKA (2023).

De maneira similar à determinação dos parâmetros ótimos para o modelo de Regressão Linear Múltipla (MLR), alguns parâmetros iniciais do modelo de *Perceptron* de Múltiplas Camadas (*MLP*) foram mantidos. Os parâmetros que foram alvo de modificações incluem: *batchSize*, *hiddenLayers*, *learningRate*, *momentum*, *numDecimalPlaces* e *trainingTime* (*Epochs*).

Os resultados das métricas, obtidos a partir dessa configuração particular de parâmetros, estão expostos na Figura 30. Essas métricas são correspondentes à fase de Validação.

Figura 30 – Resultados – Métricas (*MLP*).

```

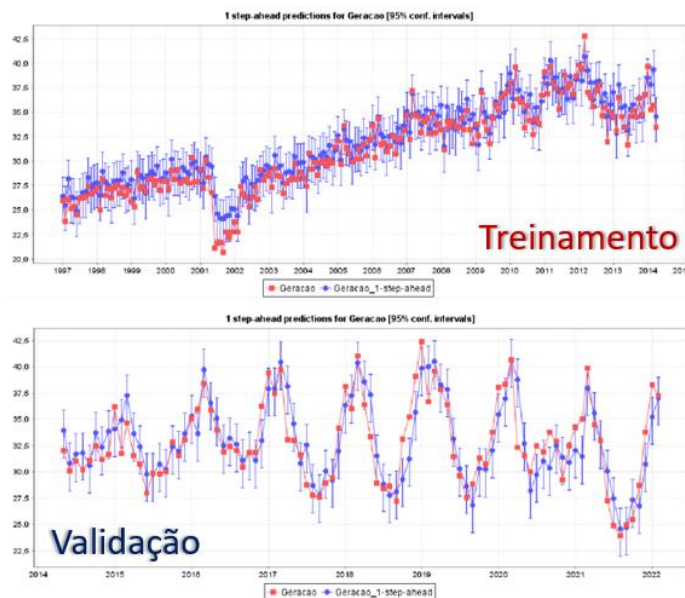
=== Evaluation on test data ===
Target          1-step-ahead  2-steps-ahead  3-steps-ahead  4-steps-ahead  5-steps-ahead  6-steps-ahead
=====
Geracao
N                94             93             92             91             90             89
Mean absolute error  1.5968        2.4085        3.2046        3.6733        3.7958        3.783
Mean absolute percentage error  4.8785        7.3446        9.8217        11.2894       11.7219       11.6825
Root mean squared error  2.0091        3.0623        3.9516        4.3814        4.506         4.5171
Total number of instances: 94

```

Fonte: Adaptado de WEKA (2023).

A Figura 31 ilustra os resultados alcançados por meio da aplicação da rede *Perceptron* de Múltiplas Camadas (*MLP*) nos conjuntos de Treinamento e Validação, levando em consideração um intervalo de confiança de 95%.

Figura 31 – Geração – Treinamento e Validação (MLP)



Fonte: Adaptado de *WEKA* (2023).

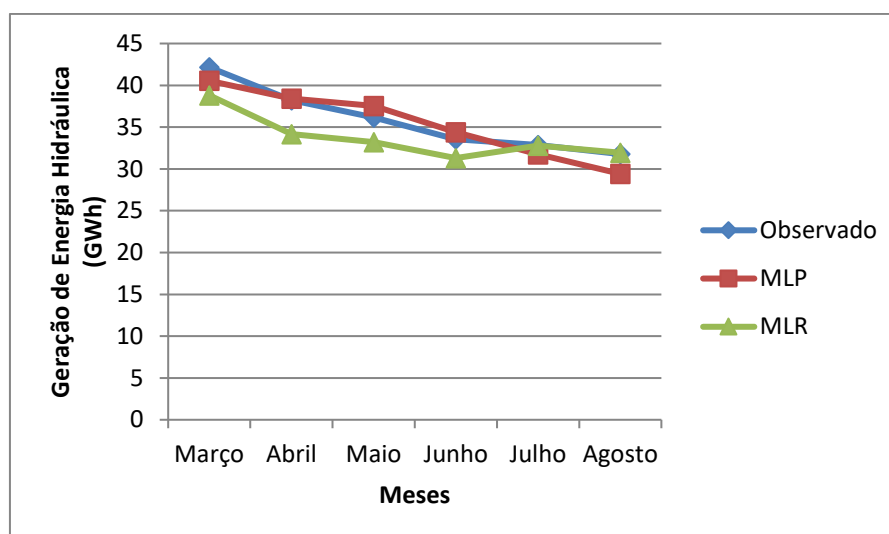
Em contraste com o modelo de Regressão Linear Múltipla (*MLR*), o modelo baseado em *Perceptron* de Múltiplas Camadas (*MLP*) demonstrou uma variabilidade mais reduzida, resultando em desempenhos mais substanciais e notavelmente superiores em sua fase de validação.

5.3 Teste

Na fase de Teste, foi conduzida uma análise comparativa entre os modelos *MLR* e *MLP*. Na Tabela 9, são exibidos os resultados concretizados e previstos pelos algoritmos, juntamente com as métricas *MAE*, *RMSE* e *MAPE*, abarcando o período de seis meses (Março, Abril, Maio, Junho, Julho e Agosto de 2022) que não foi utilizado nas etapas de Treinamento e Validação.

A Figura 32, por sua vez, visualiza de maneira gráfica os desfechos das previsões referentes aos mencionados seis meses (Março, Abril, Maio, Junho, Julho e Agosto).

Figura 32 – Previsões do conjunto Teste.



Fonte: Autoria Própria (2023).

Na Tabela 9, é factível a observação dos desfechos provenientes dos dois métodos, no âmbito da fase de Teste, mediante a utilização da métrica *MAE*.

Tabela 9 – Dados de Geração Observados e Previstos (*MLP* e *MLR*) e a métrica *MAE*.

Mês	Observado	<i>MLP</i>	$ y_i - \hat{y}_i $	<i>MLR</i>	$ y_i - \hat{y}_i $
Março	42,154	40,5436	1,61	38,8078	3,35
Abril	38,221	38,4078	0,19	34,1579	4,06
Maio	36,145	37,5193	1,37	33,2060	2,94
Junho	33,534	34,3840	0,85	31,3002	2,23
Julho	32,880	31,7519	1,13	32,7909	0,09
Agosto	31,780	29,4165	2,36	31,9527	0,17
MAE			1,25		2,14

Fonte: Autoria própria (2023).

Na Tabela 10, é possível efetuar a visualização das saídas resultantes dos dois métodos e sua confiabilidade é avaliada em consonância com a métrica *RMSE*.

Tabela 10 – Dados de Geração Observados e Previstos (*MLP* e *MLR*) e a métrica *RMSE*.

Dia	Observado	<i>MLP</i>	$((y_i - \hat{y}_i)^2)$	<i>MLR</i>	$((y_i - \hat{y}_i)^2)$
Março	42,154	40,5436	2,59	38,8078	11,20
Abril	38,221	38,4078	0,03	34,1579	16,51
Maio	36,145	37,5193	1,89	33,2060	8,64
Junho	33,534	34,3840	0,72	31,3002	4,99
Julho	32,880	31,7519	1,27	32,7909	0,01
Agosto	31,780	29,4165	5,59	31,9527	0,03
RMSE			1,42		2,63

Fonte: Autoria própria (2023).

Na Tabela 11, é evidenciado o emprego da métrica *MAPE* para a avaliação dos resultados obtidos pelos dois modelos.

Tabela 11 – Dados de Geração Observados e Previstos (*MLP* e *MLR*) e a métrica *MAPE*.

Dia	Observado	MLP	$\frac{ y_i - \hat{y}_i }{y_i} * 100$	MLR	$\frac{ y_i - \hat{y}_i }{y_i} * 100$
Março	42,154	40,5436	3,82	38,8078	7,94
Abril	38,221	38,4078	0,49	34,1579	10,63
Maio	36,145	37,5193	3,80	33,2060	8,13
Junho	33,534	34,3840	2,53	31,3002	6,66
Julho	32,880	31,7519	3,43	32,7909	0,27
Agosto	31,780	29,4165	7,44	31,9527	0,54
MAPE			3,59		5,70

Fonte: Autoria própria (2023).

Com base nas análises dos resultados registrados nas Tabelas 9, 10 e 11, pode-se deduzir que o modelo empregando o *Perceptron* de Múltiplas Camadas (*MLP*) apresentou um desempenho superior em comparação ao modelo baseado na Regressão Linear Múltipla (*MLR*), no que concerne às métricas de Erro Médio Absoluto (*MAE*), Raiz Quadrada do Erro Médio (*RMSE*) e Erro Percentual Médio Absoluto (*MAPE*).

6 CONCLUSÃO

No âmbito desta pesquisa, foi efetuada uma comparação detalhada entre dois algoritmos de previsão de séries temporais incorporados na estrutura do *software WEKA*. Estes algoritmos em análise são a Regressão Linear Múltipla e as Redes Neurais Artificiais do tipo *Multilayer Perceptron*. Ambos os algoritmos foram submetidos a um processo sequencial, compreendendo etapas de preparação de dados, estruturação de modelos, estimativa, avaliação de resultados e validação.

Iniciou-se pelo critério da Raiz Quadrada do Erro Médio (*RMSE*), Erro Médio Absoluto (*MAE*) e Erro Médio Percentual Absoluto (*MAPE*) para seleção dos modelos mais apropriados no Conjunto de Validação. Posteriormente, realizou-se a estimativa dos valores para os próximos seis meses subsequente ao mês de fevereiro (Conjunto de Teste). A comparação destes valores estimados com os valores reais foi realizada através das métricas estabelecidas. Ao longo de um horizonte de previsão de seis meses, observou-se que os modelos analisados demonstraram valores de métricas bastantes semelhantes entre si. Entretanto, ressalta-se que o modelo baseado em *Perceptron* de Múltiplas Camadas (*MLP*) exibiu resultados substancialmente superiores caracterizados por valores mais baixos de métricas, em relação ao modelo de Regressão Linear Múltipla (*MLR*), mantendo-se consistentemente aproximado dos valores reais. Como resultado, pode-se inferir que o modelo *MLP* revelou um notável potencial de generalização dentro deste horizonte de previsão.

Como perspectiva para futuras pesquisas, sugere-se a aplicação desses mesmos métodos para conjuntos de dados subsequentes aos abordados neste estudo. Além disso, uma comparação entre o desempenho do modelo de Rede *MLP* com algoritmos de aprendizado profundo, como *Long Short-Term Memory (LSTM)*, *Bidirectional LSTM (BILSTM)*, *Gated Recurrent Unit (GRU)* e *Bidirectional GRU (BGRU)*, poderia enriquecer ainda mais a compreensão da capacidade preditiva do modelo em questão.

REFERÊNCIAS

ANDRADE FILHO, R. T. **Aplicação de Redes Neurais no Controle de Tuberculose Bovina**. – Rio de Janeiro: UFRJ/COPPE, 2016.

ANDRADE, L. M. **Abordagem Neurofuzzy Para Previsão de Demanda de Energia Elétrica no Curtíssimo Prazo**. Tese de Doutorado. Universidade de São Paulo, 2010.

ANEEL. Agência Nacional de Energia Elétrica. **Quantidade de empreendimentos de geração de energia em operação**. Disponível em: <<https://dadosabertos.aneel.gov.br/dataset/empreendimentos-em-operacao>>. Acesso em: 19 abr. 2023.

ANTONIOLLI, E. A. **Previsão do consumo de energia elétrica em um frigorífico: um estudo de caso utilizando regressão linear, redes neurais e máquinas de vetores de suporte**. Dissertação de Mestrado. Universidade Tecnológica Federal do Paraná. Medianeira, 2021.

BEN (Balanço Energético Nacional) 2022: Ano base 2021 / **Empresa de Pesquisa Energética**. Rio de Janeiro : EPE, 2022. Disponível em: <<https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2022>>. Acesso em: 18 abr. 2023.

BARRETO, J. M. Inteligência Artificial no limiar do século XXI. **Florianópolis: PPP edições**, v. 97, 1999.

BASHIR, Z. A.; EL-HAWARY, M. E. Applying Wavelets to Short-Term Load Forecasting Using PSO-Based Neural Networks, **IEEE Transactions On Power Systems**, v. 24, p. 20-27, 2009.

BASTIANI, M.; SANTOS, J. A. A.; SCHMIDT, C. A P.; SEPULVEDA, G. P. L. Application of data mining algorithms in the management of the broiler production. **Geintec**, v. 8, 2018.

BECCALI, M.; CELLURA, M.; LO BRANO, V.; MARVUGLIA, A. Forecasting Daily Urban Electric Load Profiles using Artificial Neural Networks, **Energy Conversion and Management**, v. 45, p. 2879-2900, 2004.

BISHOP, C. M. Pattern recognition. **Machine Learning**, v. 128, p. 1-58, 2006.

BOSE, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. **Proceedings of the fifth annual workshop on Computational learning theory**. p. 144-152. ACM, 1992.

BRAGA, A. P.; CARVALHO A. C. P. F.; LURDEMIR, T. B. **Fundamentos de redes neurais artificiais**. Rio de Janeiro: 11ª Escola de Computação/Universidade Federal do Rio de Janeiro. 246p., 1998.

BRAGA, A. P.; DE CARVALHO, A. P. L. F.; LUDERMIR, T.B. **Redes neurais artificiais: teoria e aplicações**. LTC Editora, 2007.

CANKURT, S.; SUBASI, A. Comparasion of linear regression and neural network models forecasting tourist arrivals to turkey. **Eurasian Journal of Science & Engineering**, 2015.

CHARYTONIUK, W.; CHEN, M. S. Very short-term load forecasting using artificial neural networks. **IEEE transactions on Power Systems**, v. 15, n. 1, p. 263-268, 2000.

CYBENKO, G. **Continuous valued neural networks with two hidden layers are sufficient**. Medford, MA: Department of Computer Science, Tufts University, 1988. Technical Report.

EPE. (Empresa de Pesquisa Energética). **Matriz Energética e Elétrica**. Disponível em: <<https://www.epe.gov.br/pt/abcdenergia/matriz-energetica-e-eletrica>>. Acesso em: 19 abr. 2023.

FÁVERO, L. P. **Análise de Dados**. Rio de Janeiro: Elsevier, 2015. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788595153226/>>. Acesso em: 22 mai. 2023.

FIORIN, D. V.; SCHUCH, N. J. F. R.; Pereira, E. B. Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares. **Revista Brasileira de Ensino de Física**, v. 33, n. 1, p. 1309, 2011.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. São Paulo: Atlas, 2010.

Guarnieri, R. A. Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do modelo Eta. **Instituto Nacional de Pesquisas Espaciais**, 171pp, 2006.

GUNAY, M. E. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. **Energy Policy**, v. 90, p. 92-101, 2016.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, U. H. **The WEKA Data Mining Software: An Update**. SIGKDD Explorations, 2009. 11(1), 10-18.

HAYKIN, S. A comprehensive foundation. **Neural Networks**, v. 2, n. 2004, p. 41, 2004.

HAYKIN, S. **Neural networks: a comprehensive foundation**. New Delhi: Pearson Prentice Hall, 2001.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Hamilton, Ontario: Bookman, 2005. ISBN 0-13-273350-1.

HEATON, J. **Artificial intelligence for humans**. 1. ed. Clarkston, Georgia: Heaton Research, 2015. v. 3. ISBN 978-1505714340.

HEBB, D. O. **The organization of behavior**. Wiley: New York, NY, USA, 1949.

IPEA. **Geração de energia elétrica hidráulica: quantidade**. Disponível em: <[https:// http://www.ipeadata.gov.br/Default.aspx](https://http://www.ipeadata.gov.br/Default.aspx)>. Acesso em: 19 abr. 2023.

KRAJEWSKI, L.; RITZMAN, L.; MALHOTRA, M. **Administração de produção e operações**. 8a. Ed. São Paulo. Pearson Prentice Hall. 2009.

MAMLOOK, R.; BADRAN, O.; ABDULHADI, E. A Fuzzy Inference Model for ShortTerm Load Forecasting, **Energy Policy**, v. 37, p. 1239-1248, 2009.

MCCULLOCH, W. S.; PITTTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.

MELO, A. P. **Desenvolvimento de um método para estimar o consumo de energia de edificações comerciais através da aplicação de redes neurais**. Tese de Doutorado. Universidade Federal de Santa Catarina, 2012.

MME (Ministério de Minas e Energia). Boletim Mensal de Monitoramento do Sistema Elétrico Brasileiro. Disponível em: <https://www.gov.br/mme/pt-br/assuntos/noticias/BoletimdeMonitoramentodoSistemaEltricoFev2022_v15.pdf>. Acesso em: 25 abr. 2023.

MINSKY, M.; PAPERT, S. **Perceptrons: an introduction to computational geometry**. Massachusetts: MIT Press, 1969.

MORETTIN, P. A.; TOLOI, C. **Análise de séries temporais**. Blucher, 2006.

OLIVEIRA, P. S. S. ; Mendiondo, R. Métodos de Previsão de Consumo de Energia Elétrica: Regressão e Redes Neurais Artificiais. **Congresso Nacional de Excelência em Gestão**, Rio de Janeiro, 2014.

OSÓRIO, F. S.; BITTENCOURT, J. R.; OSÓRIO, J. R. Sistemas inteligentes baseados em redes neurais artificiais aplicados ao processamento de imagens. I **Workshop de inteligência artificial**. 2000.

PÁDUA, A. B.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. LTC Editora, 2007.

PINHEIRO, T. C., SANTOS, J. A. A., PASA, L. A. Gestão da produção de frangos de corte por meio de redes neurais artificiais, **Revista Holos**, 2020.

RAHMAN, S.; HAZIM, O. A Generalized Knowledge-Based Short-Term Load Forecasting Technique. **IEEE Transactions on Power Systems**, v. 8, p. 508-514, 1993.

RAIZER, K.; IDAGAWA, H. S.; de OLIVEIRA NÓBREGA, E. G.; FERREIRA, L. O. S. Training and Applying a Feedforward Multilayer Neural Network in GPU. In **Proceedings of the 30th Iberian-Latin-American Congress on Computational Methods in Engineering**. 2009.

REIDMILLER, M. Advanced supervised learning in multi-layer perceptrons - from backpropagation to adaptive learning algorithms, **Computer Standards & Interfaces**, v. 16, n. 3, p. 265–278, 1994.

ROLA, M. C. **Previsão da geração de energia elétrica no médio prazo para o estado do Rio Grande do Sul empregando redes neurais artificiais**. [Dissertação de Mestrado], Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.

ROMERA, G. E.; MORAN, M. A. J.; FERNANDEZ, D. C. Forecasting of the Electric Energy Demand Trend and Monthly Fluctuation with Neural Networks. **Computers & Industrial Engineering**, v. 52. p. 336-343, 2007.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychol. Rev.**, v. 65, p. 386-408, 1958.

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

SANTANA, E. **Energia hidroelétrica**. (2019) Disponível em: <<https://www.educamaisbrasil.com.br/enem/fisica/energia-hidreletrica>>. Acesso em: 10 mai. 2023.

SANTOS, J. A. A. S.; ANTONIOLLI, E. A.; GUEDES, A. L.; PASA, L. A. Previsão do consumo de energia elétrica em um frigorífico: um estudo de caso usando regressão linear, redes neurais e máquinas de vetor de suporte. **Tekhne e Logos**, v.13, n.2, 2022.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding Machine Learning**. Cambridge, Massachusetts: Cambridge University, 2014. ISBN 978-1-107-05713-5.

Silva, I. D.; Spatti, D. H.; Flauzino, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, p. 33-111, 2010.

SOARES, M. A.; COSTA H. K. M. The Brazil's Power distribution utilities: an assessment about crisis in 2001 e 2021. **Conjecturas**, v.22, n.2, 2022.

SOUZA, B. C. S. **Mineração de dados aplicada a previsão do preço de ações utilizando o WEKA**. Trabalho de Conclusão de Curso. Pontifícia Universidade Católica de Goiás. Goiânia, 2021.

TIDRE, P. V. V.; BIASE, N. G. G.; SILVA, M. I. S. Utilização dos modelos de séries temporais na previsão do consumo mensal de energia elétrica da região Norte do Brasil. **Matemática e Estatística em Foco**, v. 1, n. 1, p. 57-66, 2013.

VASILEV, I. et al. **Python deep learning: exploring deep learning techniques and neural network architectures with PyTorch, Keras, and TensorFlow**. 2. ed. Birmingham: Packt Publishing, 2019. ISBN 978-1-78934-846-0.

WASSEMAN, P. D. **Advanced methods in neural computing**. New York: Van Nostrand Reinhold, 1989.

WIDROW, B.; LEHR, M. A. 30 Years of Adaptive Neural Networks: Perceptron, Madaline and Backpropagation. **Proceedings of the IEEE**, v. 78, n. 9, 1990.

YAP, K. S.; ABIDIN, I. Z.; LIM, C. P.; SHAH, M. S. Short Term Load Forecasting Using a Hybrid Neural Network, **First International Power and Energy Conference (PECon)**, p. 123-128, 2006.