

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"

FACULDADE DE ENGENHARIA - CÂMPUS ILHA SOLTEIRA

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

BACHARELADO EM ENGENHARIA ELÉTRICA

VINÍCIUS RODRIGUES SILVA

**PREVISÃO DE CARGA EM SISTEMAS ELÉTRICOS DE
POTÊNCIA: UM ESTUDO COM ENFOQUE BASEADO NO USO
DE MÉTODOS DE MACHINE LEARNING**

ILHA SOLTEIRA (SP), 2022

VINÍCIUS RODRIGUES SILVA

**PREVISÃO DE CARGA EM SISTEMAS ELÉTRICOS DE
POTÊNCIA: UM ESTUDO COM ENFOQUE BASEADO NO USO
DE MÉTODOS DE MACHINE LEARNING**

Trabalho de graduação apresentado à Faculdade de Engenharia de Ilha Solteira – UNESP – Câmpus de Ilha Solteira, como parte dos requisitos para obtenção do título de Bacharel em Engenharia Elétrica.

Orientador: Carlos Roberto Minussi

FICHA CATALOGRÁFICA

Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

S586p Silva, Vinícius Rodrigues.
Previsão de carga em sistemas elétricos de potência: um estudo com enfoque baseado no uso de métodos de machine learning / Vinícius Rodrigues Silva. -- Ilha Solteira: [s.n.], 2022
65 f. : il.

Trabalho de conclusão de curso (Graduação em Engenharia Elétrica) -
Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira, 2022

Orientador: Carlos Roberto Minussi
Inclui bibliografia

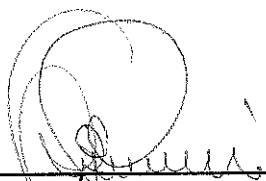
1. Previsão de carga. 2. Curto-prazo. 3. Machine learning. 4. Rede neural.

Raiane da Silva Santos
Raiane da Silva Santos

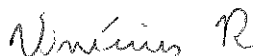
Supervisora Técnica de Seção
Seção Técnica de Referência, Atendimento ao usuário e Documentação
Diretoria Técnica de Biblioteca e Documentação
CRB/8 - 9999

ATA DE DEFESA DE TRABALHO DE GRADUAÇÃO

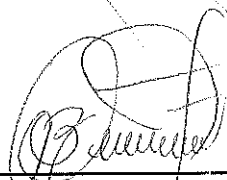
Aos doze dias do mês de agosto do ano de dois mil e vinte e dois, o discente **Vinícius Rodrigues Silva**, matriculado sob o nº 162053959, tendo como banca examinadora o seu orientador, o Prof. Dr. Carlos Roberto Minussi, o Doutorando Leonardo Brain Garcia Fernández e o Doutorando Salvador Falcon Canillas, apresentou o Trabalho de Graduação intitulado "**Previsão de carga em sistemas elétricos de potência: um estudo com enfoque baseado no uso de métodos de machine learning**", obtendo a nota 9,0 (NOVE VÍRGULA ZERO) e conceito APROVADO.



Prof. Dr. Carlos Roberto Minussi
- Orientador -



Vinícius Rodrigues Silva
- Discente -



Doutorando Leonardo Brain Garcia Fernández
- Membro da Banca -



Doutorando Salvador Falcon Canillas
- Membro da Banca -

Agradecimentos

Agradeço a Deus pela oportunidade e resiliência para cumprir as etapas desta jornada.

Agradeço aos meus pais Angelita e Volney por todo o carinho, apoio e aconselhamentos.

Agradeço ao meu orientador, Prof. Carlos Roberto Minussi, por ter acreditado em mim e aceitado me orientar.

Agradeço aos amigos pela parceria de longa data e por terem me ajudado quando eu precisei.

Agradeço à UNESP por ter me possibilitado um acesso a uma licença acadêmica do MATLAB, ferramenta fundamental para o desenvolvimento das atividades deste Trabalho de Graduação.

“A experiência não é o que acontece com um homem. É o que um homem faz com o que acontece com ele.”

Aldous Huxley

Resumo

Com o intuito de atender o gerenciamento da demanda de energia, a previsão de carga é um componente vital para os sistemas de energia. O principal papel de uma organização de energia elétrica é fornecer energia elétrica de forma segura e econômica, mantendo a qualidade. Em sistemas de energia, a previsão de carga é uma questão essencial e tem sido amplamente estudada para obter resultados de previsão de carga mais precisos. Como em sistemas de energia, a geração de energia do dia seguinte deve ser programada todos os dias, a Previsão de Carga de Curto Prazo (PCCP) é uma tarefa diária importante para as empresas de energia. A PCCP representa a previsão de carga elétrica para um período de algumas horas a vários dias. Neste trabalho de graduação será usado um método no contexto das Redes Neurais Artificiais (RNA). RNA é um método que imita a forma como o cérebro humano processa informações. Vários tipos de RNAs foram apresentados nesta pesquisa, dentre eles a Rede Neural Feedforward, com treinamento via Algoritmo Levenberg-Marquardt Retropropagado, foi utilizada para criar um modelo PCCP. O erro percentual médio absoluto (MAPE) foi usado como métrica para avaliar o desempenho dos resultados da previsão. Por meio de avaliações detalhadas de desempenho, esta pesquisa evidencia que o modelo PCCP apresentado é capaz de prever com precisão a carga elétrica do dia seguinte atendendo às exigências do setor elétrico (MAPE igual, ou inferior a 5%).

Palavras-chave: Previsão de Carga, Curto Prazo, Rede Neural Artificial, Aprendizado de máquina.

Abstract

To meet energy demand management, load forecasting is a vital component of energy systems. The primary role of an electrical energy organization is to provide electrical energy safely and economically while maintaining quality. In power systems, load forecasting is an essential issue and has been extensively studied to obtain more accurate load forecasting results. As in power systems, day-ahead power generation must be scheduled every day. Short-term load forecasting (PCCP) is an essential daily task for power companies. PCCP represents the electrical load forecast for a few hours to several days. In this undergraduate work (TG), a method in the context of Artificial Neural Networks (ANN) is used to create a PCCP model. RNA is a method that mimics the way the human brain processes information. Several types of ANNs were presented in this research. The Feedforward Neural Network, training via the Retropropagated Levenberg-Marquardt Algorithm, was used to create a PCCP model. The Mean Absolute Percentage Error (MAPE) was used to evaluate the forecast results performance. Through detailed performance evaluations, this research shows that the PCCP model presented is capable of predicting the day ahead electric load accurately (MAPE equal to, or less than 5%).

Keywords: Load Forecast, Short Term, Artificial Neural Network, Machine Learning.

Lista de figuras

Figura 1 – Modelagem de Série Temporal de Carga	21
Figura 2 – Modelo da Função de Transferência (FT)	23
Figura 3 – Diagrama Esquemático Básico de um Neurônio Biológico	26
Figura 4 – Modelo do Neurônio de McCulloch-Pitts	26
Figura 5 – Função Relé	27
Figura 6 – Função Sigmoides 1	28
Figura 7 – Função Sigmoides 2	28
Figura 8 – Rede Neural Artificial	29
Figura 9 – Rede Neural Feedforward	30
Figura 10 – Rede Neural Recursiva	31
Figura 11 – Rede Neural Perceptron Multicamadas Com Duas Camadas Ocultas.	32
Figura 12 – Rede Neural Recorrente	32
Figura 13 – Fluxograma do Processo de Treinamento de Rede Neural	33
Figura 14 – Fluxograma Treinamento Retropagação	38
Figura 15 – Ferramenta Neural Network Training (nntraintool)	44
Figura 16 – Rede Neural Feedforward com 2 Camadas	45
Figura 17 – Rede Neural Feedforward com 3 Camadas	45
Figura 18 – Previsão de Carga Elétrica Global - Aplicação 1: Período 1 e Configuração 1	46
Figura 19 – Previsão de Carga Elétrica Global - Aplicação 2: Período 1 e Configuração 2	47
Figura 20 – Previsão de Carga Elétrica Global - Aplicação 3: Período 1 e Configuração 3	47
Figura 21 – Previsão de Carga Elétrica Global - Aplicação 4: Período 1 e Configuração 4	48
Figura 22 – Previsão de Carga Elétrica Global - Aplicação 5: Período 1 e Configuração 5	48
Figura 23 – Previsão de Carga Elétrica Global - Aplicação 6: Período 1 e Configuração 6	49
Figura 24 – Previsão de Carga Elétrica Global - Aplicação 7: Período 1 e Configuração 7	49
Figura 25 – Previsão de Carga Elétrica Global - Aplicação 8: Período 1 e Configuração 8	50
Figura 26 – Previsão de Carga Elétrica Global - Aplicação 9: Período 2 e Configuração 1	50
Figura 27 – Previsão de Carga Elétrica Global - Aplicação 10: Período 2 e Configuração 2	51
Figura 28 – Previsão de Carga Elétrica Global - Aplicação 11: Período 2 e Configuração 3	51
Figura 29 – Previsão de Carga Elétrica Global - Aplicação 12: Período 2 e Configuração 4	52
Figura 30 – Previsão de Carga Elétrica Global - Aplicação 13: Período 2 e Configuração 5	52
Figura 31 – Previsão de Carga Elétrica Global - Aplicação 14: Período 2 e Configuração 6	53
Figura 32 – Previsão de Carga Elétrica Global - Aplicação 15: Período 2 e Configuração 7	53
Figura 33 – Previsão de Carga Elétrica Global - Aplicação 16: Período 2 e Configuração 8	54
Figura 34 – Previsão de Carga Elétrica Global - Aplicação 17: Período 3 e Configuração 1	54
Figura 35 – Previsão de Carga Elétrica Global - Aplicação 18: Período 3 e Configuração 2	55
Figura 36 – Previsão de Carga Elétrica Global - Aplicação 19: Período 3 e Configuração 3	55

Figura 37 – Previsão de Carga Elétrica Global - Aplicação 20: Período 3 e Configuração 4 . .	56
Figura 38 – Previsão de Carga Elétrica Global - Aplicação 21: Período 3 e Configuração 5 . .	56
Figura 39 – Previsão de Carga Elétrica Global - Aplicação 22: Período 3 e Configuração 6 . .	57
Figura 40 – Previsão de Carga Elétrica Global - Aplicação 23: Período 3 e Configuração 7 . .	57
Figura 41 – Previsão de Carga Elétrica Global - Aplicação 24: Período 3 e Configuração 8 . .	58

Lista de tabelas

Tabela 1 – Períodos utilizados para treinamento da RNA	43
Tabela 2 – Configurações de Camada	45
Tabela 3 – Parâmetros de Treinamento da Rede Neural	46
Tabela 4 – Resultados Comparativos Médios Para Previsões do Período 1	58
Tabela 5 – Resultados Comparativos Médios Para Previsões do Período 2	59
Tabela 6 – Resultados Comparativos Médios Para Previsões do Período 3	59

Lista de abreviaturas e siglas

AR	Autoregressive
ARIMA	Auto Regressive Integrated Moving-Average
ARMA	Autoregressive Moving-Average
E _{max}	Erro Máximo Percentual
LF	Lógica Fuzzy
MA	Moving-Average
MAPE	Erro Percentual Médio Absoluto
PCCP	Previsão de Carga à Curto Prazo
RL	Regressão Linear
RLM	Regressão Linear Multinível
RMSE	Erro Quadrático Médio
RNA	Rede Neural Artificial
SE	Sistema Especialista
SEP	Sistema Elétrico de Potência

Sumário

1	INTRODUÇÃO	15
1.1	Objetivo do trabalho	15
1.2	Organização do trabalho	16
2	PREVISÃO DE CARGAS ELÉTRICAS	17
2.1	Introdução	17
2.2	Importância da previsão	17
2.3	Fatores que influenciam a previsão de carga elétrica	18
2.3.1	Fatores climáticos	18
2.3.2	Fatores de tempo	18
2.3.3	Fatores econômicos	18
2.3.4	Fatores aleatórios	19
2.4	Classificação da previsão de cargas elétricas	19
2.5	Metodologias de previsão de cargas elétricas	20
2.5.1	Métodos de regressão	20
2.5.2	Séries temporais	21
2.5.3	Filtro Kalman	22
2.5.4	Sistemas especialistas	23
2.5.5	Lógica fuzzy	24
2.5.6	Redes neurais	24
3	REDES NEURAIS ARTIFICIAIS	25
3.1	Introdução	25
3.2	Modelo neural biológico	25
3.3	Modelo neural artificial	26
3.4	Estrutura da rede neural	29
3.5	Arquiteturas de redes neurais	29
3.5.1	Rede neural feedforward	30
3.5.2	Rede neural recursiva	30
3.5.3	Rede neural perceptron multicamadas	31
3.5.4	Rede neural convolucional	31
3.5.5	Rede neural recorrente	31
4	PROJETO DE REDES NEURAIS PARA PREVISÃO	33
4.1	Introdução	33
4.2	Coleta de dados e pré-processamento	33

4.3	Tipo de rede e arquitetura	35
4.3.1	Número de camadas ocultas	35
4.3.2	Número de neurônios nas camadas ocultas	35
4.4	Treinamento da rede neural	36
4.4.1	Pesos iniciais	36
4.4.2	Classificação de algoritmos de treinamento	36
4.4.3	Algoritmo de treinamento	37
4.4.4	Crítérios para interromper o treinamento	40
4.5	Análise de desempenho da rede	40
4.6	Erro de previsão	41
5	METODOLOGIA E RESULTADOS	43
5.1	Introdução	43
5.2	Metodologia	43
5.3	Resultados	46
5.3.1	Curvas de Carga Global	46
5.3.2	Análise comparativa de treinamento e desempenho	58
6	CONCLUSÃO	60
	REFERÊNCIAS	62

1 Introdução

O conhecimento do perfil da carga é fundamental para o planejamento da expansão, fluxo de potência, operação econômica, análise, e controle de segurança sistemas de energia elétrica. (O'DONOVAN, 1983) Na literatura clássica existem vários métodos estatísticos como, por exemplo, o Auto Regressive Integrated Moving Average (ARIMA) de Box e Jenkins (BOX; JENKINS, 1976), filtro de Kalman, alisamento exponencial, estimação de estado, modelos de regressão linear (simples ou múltiplo), etc. (MOGHRAM, 1989) Todavia para utilizar estes modelos é necessária a modelagem prévia da carga por meio de variáveis que afetam a demanda de energia como condições meteorológicas (nebulosidade, velocidade do vento, estações do ano, variações de temperatura), efeito de dias não convencionais (feriados, greves, eventos esportivos) e entre outras variáveis exógenas tornando a previsão da carga complexa.

Um método alternativo para a previsão de cargas elétricas é o uso de Rede Neural Artificial (RNA). (HAYKIN, 2009). O uso da inteligência artificial está se destacando por causa a sua eficiência em aplicações reais, grande capacidade de processamentos de dados e independência de modelos matemáticos complexos.

As redes neurais artificiais, modelos computacionais baseados no cérebro humano, possuem capacidade de aquisição e manutenção de conhecimento adquirido por dados passados. Uma de suas vantagens é a fase de aprendizado realizada em modo “off-line” refletindo em um tempo computacional relativamente baixo. (HAYKIN, 2009) Além disso, por não ser necessário realizar a modelagem prévia dos dados, as rede neurais possuem a capacidade de prever a carga sem estabelecer a correlação entre as grandezas que influenciam o seu comportamento.

A neurocomputação surgiu em 1943, em especial, a partir da publicação de McCulloch e Pitts. Neste artigo os autores propuseram um modelo artificial inspirado no neurônio biológico. As redes neurais artificiais começaram a ter destaque apenas na década de 70 com a proposição da rede de Hopfield (HAYKIN, 2009), advento do algoritmo retropapagação (backpropagation) (WERBOS, 1974) e avanços da microeletrônica.

O algoritmo retropapagação visa o aprendizado supervisionado de RNA e fundamenta-se no método gradiente descendente do erro quadrático para a adaptação dos pesos sinápticos. Este algoritmo é considerado padrão na literatura especializada, pois com ele é possível executar com precisão tarefas como diagnósticos complexos, previsão e reconhecimento de padrões.

1.1 Objetivo do trabalho

A proposta deste trabalho é realizar uma rotina de previsão de carga elétrica global a curto prazo utilizando a rede neural feedfoward (PARK et al., 1991) via algoritmo

retropapagação.

1.2 Organização do trabalho

O trabalho está organizado em seis capítulos. O primeiro capítulo corresponde a presente introdução. O segundo capítulo aborda informações pertinentes à previsão de cargas elétricas, como os fatores que influenciam a previsão da carga elétrica e as principais metodologias de previsão. O terceiro capítulo enfoca nos conceitos mais importantes de redes neurais artificiais. O quarto capítulo aborda como dimensionar os principais parâmetros de uma rede neural bem como o procedimento de treinamento de uma rede neural. O quinto capítulo apresenta a metodologia empregada e os resultados obtidos com a rotina de previsão proposta. Por fim o capítulo 6 mostra as conclusões sobre previsão realizada.

2 Previsão de cargas elétricas

2.1 Introdução

Um sistema elétrico de potência (SEP) é um sistema de entrega de energia em tempo real. Difere de um sistema de gás ou água que são sistemas de armazenamento. Um SEP é considerado um sistema de tempo real por a energia ser gerada, transportada e fornecida no momento que ela é utilizada. Em sistemas elétricos de potência, existem três estágios para fornecer a energia elétrica para o cliente, são eles: geração, transmissão e distribuição. (KYRIAKIDES; POLYCARPOU, 2007)

Na operação de um SEP, a capacidade de prover carga para os clientes é o aspecto mais desafiador, pois isso significa ser sempre necessário garantir os requerimentos da carga a todo instante. O gerador precisa ter uma carga extra que pode ser utilizada a qualquer momento. Em razão das flutuações significativas da carga durante o dia, o sistema operante deve conseguir prever a demanda de carga a curto prazo para que um planejamento adequado possa ser realizado.

Para ter conhecimento prévio dos requerimentos da carga, é necessário a previsão de carga. A capacidade de prever carga é um dos aspectos efetivos mais importantes de gerenciamento de sistemas de potência. A previsão de carga é essencial para planejamento e tomada de decisão operacional. (HAHN; PICKL, 2009) Por isso é imprescindível o conhecimento do perfil futuro da carga. Desse modo, neste capítulo será apresentado conceitos sobre o problema de previsão de cargas elétricas com enfoque nos principais métodos e técnicas utilizadas para solucionar este problema.

2.2 Importância da previsão

Segundo SCHMIDT C. A. J.; LIMA, o setor elétrico é essencial para o desenvolvimento econômico de um país. Para evitar crises futuras neste setor, existe uma necessidade de expansão por investimentos na área de pesquisa e desenvolvimento. Neste cenário, por afetar os preços das tarifas elétricas, a previsão de cargas elétricas tem um papel fundamental para os sistemas de potência. (VEHVILÄINEN I; KEPPO, 2003)

Hoje em dia, a previsão de carga a curto prazo (PCCP) tornou-se um assunto urgente. O erro na previsão pode levar a um aumento do custo operacional e diminuição da receita. Realizada uma previsão precisa, o setor de energia elétrica consegue otimizar a sua distribuição, atendendo de maneira adequada as demandas, minimizando as perdas e maximizando os lucros. (LOPES, 2005)

2.3 Fatores que influenciam a previsão de carga elétrica

A carga de um sistema é a somatório da carga de todos os consumidores em simultâneo. O objetivo da PCCP é prever a carga futura de um sistema. O entendimento adequado das características de um sistema ajuda a projetar um modelo de previsão razoável e selecionar modelos de operação adequados em diferentes situações. Vários são os fatores que influenciam o comportamento da carga de um sistema, estes podem ser classificados nas seguintes categorias: (GROSS; GALIANA, 1987)

- Fatores climáticos
- Fatores de tempo
- Fatores econômicos
- Fatores aleatórios

2.3.1 Fatores climáticos

Os fatores climáticos possuem um efeito significativo na PCCP de um sistema de potência. Estes incluem temperatura, umidade, precipitação, velocidade do vento, cobertura de nuvens, intensidade da luz, etc. De acordo com uma pesquisa de previsão de carga elétrica a temperatura do ar é o principal fator a considerar para um modelo de PCCP. (EUGENE; DORA, 2004) A alteração do clima pode causar desconforto nos consumidores e, no que lhe concerne, o uso de alguns aparelhos como aquecedor, ventilador e ar-condicionado. Estes aparelhos afetam significativamente um sistema de potência dado que as cargas elétricas dos aparelhos mencionados tendem a ser as maiores do sistema.

2.3.2 Fatores de tempo

Os fatores de tempo possuem um papel importante ao influenciar padrões de carga. Estes incluem efeitos sazonais e efeitos cíclicos semanais-diários. Em relação aos efeitos sazonais ocorrem variações graduais no padrão da carga em resposta a diferentes agentes como a quantidade de horas de iluminação diária, mudanças na estrutura tarifária, início do ano escolar, período de férias, etc. (GROSS; GALIANA, 1987) Sobre efeitos cíclicos semanais-diários é possível observar o comportamento da carga em dias especiais da semana como redução da mesma nos finais de semana e pequena redução nas sextas-feiras e nas segundas-feiras. (MURTO, 1998)

2.3.3 Fatores econômicos

A eletricidade é uma categoria de mercadoria. A situação econômica regional também influencia na utilização desta mercadoria. Fatores econômicos como nível de industrialização,

preço da eletricidade e política de gerenciamento de carga possuem impactos significativos na tendência de crescimento ou declínio da carga do sistema. Com o desenvolvimento de mercados modernos de energia, a relação entre o preço da energia e o perfil da carga é ainda maior. (KHATOON et al., 2014) Contudo como fatores econômicos estão associados a longa escala de tempo, estes não são explicitamente considerados em modelos de previsão de carga a curto prazo. (GROSS; GALIANA, 1987)

2.3.4 Fatores aleatórios

O sistema de potência moderno é composto por um elevado número de usuários. Embora não seja possível prever o comportamento individual de consumo de energia elétrica, a quantidade total de carga de todos os usuários obedecem a um comportamento estatístico (estocástico) que tende a uma curva de carga suave. A inicialização ou interrupção de grandes cargas como metalúrgicas, síncrotrons e túneis de ventos levam a um impulso óbvio na curva de carga. Este impulso na curva de carga é um distúrbio aleatório, pois para os fornecedores o tempo de inicialização e interrupção de cargas pelos consumidores é aleatório, ou seja, não existe uma regra óbvia de como e quando eles consomem energia da rede. Quando os dados desta curva são consideradas na previsão de carga, o componente impulso da carga aumenta a dificuldade da previsão. Outra fonte de distúrbios aleatórios são eventos especiais conhecidos antecipadamente, no entanto, os efeitos na carga por causa deles não são certos. Um típico evento é a copa do mundo de futebol, em que os fornecedores sabem que ocorrerá acréscimo no consumo de televisão, contudo não conseguem dimensionar a quantidade de uso. Outros eventos típicos incluem greves, paralisação de indústrias e o gerenciamento compulsório do governo devido à escassez de eletricidade prevista.

2.4 Classificação da previsão de cargas elétricas

A previsão de carga é essencial para o planejamento e tomada de decisão operacional. Com base no intervalo de tempo, a previsão de carga pode ser categorizada em quatro grupos: (AL-KANDARI; SOLIMAN; EL-HAWARY, 2004)

- Longo prazo: previsão com intervalo de tempo anual. É utilizada principalmente na operação de planejamento de sistemas de potência.
- Médio prazo: previsão com intervalo de tempo mensal. É empregada para o planejamento de suprimentos de combustível, programas de manutenção e planejamento de produção.
- Curto prazo: previsão com intervalo de tempo semanal-diária. É aplicado principalmente para previsões operacionais, econômicas e de segurança diária do sistema elétrico de potência.

- Curtíssimo prazo: previsão com intervalo de tempo entre segundos e minutos. É aproveitado no planejamento de operação “on-line” e controle.

As diferenças nas escalas de tempo têm consequências para os modelos e métodos aplicados e para os dados de entrada disponíveis e selecionados. O responsável pela tomada de decisão deve considerar não apenas encontrar a categoria de modelo apropriado, mas também determinar os fatores externos importantes necessários para obter a previsão mais precisa. (HAHN; PICKL, 2009) No trabalho desenvolvido será focado somente o problema de previsão de carga a curto prazo.

2.5 Metodologias de previsão de cargas elétricas

As abordagens de pesquisa de PCCP podem ser divididas em duas categorias: métodos estatísticos e métodos de inteligência artificial. (GROSS; GALIANA, 1987) Nos métodos estatísticos, as equações podem ser obtidas mostrando a relação entre a carga e seus fatores relativos após treinar os dados históricos, enquanto os métodos de inteligência artificial tentam imitar a maneira de pensar e raciocinar dos seres humanos para obter conhecimento da experiência e prever a carga futura.

Os métodos estatísticos incluem suavização exponencial geral (CHRISTIANSE, 1971), regressão linear múltipla (PAPALEXOPOULOS; HESTERBERG, 1990), espaço de estado (VILLALBA; BEL, 2000), séries temporais estocásticas (AMJADY, 2001), etc. Recentemente, um método de aprendizado estatístico muito promissor chamado suporte à regressão vetorial (SRV) (YANG; CHENG, 2004; LI; FANG, 2002) também tem sido aplicado à PCCP e tem mostrado resultados positivos. Normalmente, os métodos estatísticos podem prever muito bem a curva de carga de dias comuns, mas não conseguem analisar a propriedade de carga de feriados e outros dias anômalos, devido à inflexibilidade de sua estrutura.

Sistemas especialistas (HWAN; KIM, 2001), lógica nebulosa (AL-KANDARI; SOLIMAN; EL-HAWARY, 2004), rede neural artificial (RNA) (DESOUKY; ELKATEB, 2000), inferência difusa (KIM; YOUN; KANG, 2000) e algoritmo evolucionário pertencem à categoria de inteligência computacional. Em comparação aos métodos estatísticos, as técnicas de inteligência computacional são mais flexíveis em encontrar a relação entre a carga e seus fatores relativos, especialmente para a previsão de carga anômala.

Alguns dos principais métodos de PCCP são apresentados com mais detalhes a seguir.

2.5.1 Métodos de regressão

O modelo de regressão linear (RL) representa a relação linear entre a carga e outros fatores como clima, cliente e dia. RL utiliza a técnica de estimação de mínimos quadrados

ponderados a partir de dados históricos. Nesse procedimento a temperatura é a informação mais importante para a previsão de carga elétrica aliado a outros fatores climáticos.

A regressão linear multinível (RLM) é mais utilizada do que outras categorias de regressões lineares. (HONG et al., 2010) A RLM descreve como uma variável, linearmente influenciada por vários fatores preditores, pode prever a carga elétrica em um tempo definido a partir do clima e outras variáveis. A análise de correlação é usada para selecionar estes fatores, que possuem mais influência na curva de carga. (AMRAL; OZVEREN; KING, 2007)

Os métodos de regressão são relativamente simples de implementar. A sua vantagem é que a relação entre as variáveis de entrada e de saída é fácil de ser compreendida. A sua desvantagem está no fato da não linearidade e complexidade da relação entre a carga de demanda e as variáveis influenciadoras. Além disso, para conseguir resultados razoavelmente precisos é necessário um poder computacional elevado.

2.5.2 Séries temporais

As séries temporais podem ser definidas como um conjunto sequencial de dados medidos temporalmente, como a carga de pico horária, diária ou semanal. Elas utilizam de dados históricos da carga para prever a carga futura (MANOHAR; VEERA, 2008). A carga elétrica $y(t)$ é modelada como a saída de um filtro linear com uma entrada em série aleatória $a(t)$ também chamada ruído branco. Esta entrada aleatória tem uma média zero com variância desconhecida fixa $\sigma_a^2(t)$. Diferentes modelos de PCCP podem ser classificados dependendo da característica do filtro linear. (ASHOUR; ABU; EL-FATTAH, 1991)

Figura 1 – Modelagem de Série Temporal de Carga



Fonte: Adaptado de (ASHOUR; ABU; EL-FATTAH, 1991)

O modelo “Autoregressive” (AR) define a carga elétrica prevista $y(t)$, em termos dos valores de carga anteriores e um sinal de ruído aleatório $a(t)$ (MOGHARAM, 1989).

$$y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + a(t) \quad (2.1)$$

sendo,

$\phi_1, \dots, \phi_p$ = coeficientes autorregressivos;
 p = ordem do modelo.

O modelo “Moving-Average” (MA) define a carga elétrica prevista $y(t)$, em termos dos valores atuais e anteriores de séries de sinais de ruído aleatório. Este ruído é construído via os erros de previsão ou resíduos quando as observações de carga se tornam disponíveis. (MOGHRAM, 1989)

$$y(t) = a(t) - \Theta_1 a(t-1) - \Theta_2 a(t-2) - \dots - \Theta_q a(t-q) \quad (2.2)$$

em que,

$\Theta_1, \dots, \Theta_q$ = coeficientes de média móvel;
 q = ordem do modelo.

O modelo “Autoregressive Moving-Average” (ARMA) define a carga elétrica prevista $y(t)$, com a combinação dos modelos AR e MA.

$$y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + a(t) - \Theta_1 a(t-1) - \Theta_2 a(t-2) - \dots - \Theta_q a(t-q) \quad (2.3)$$

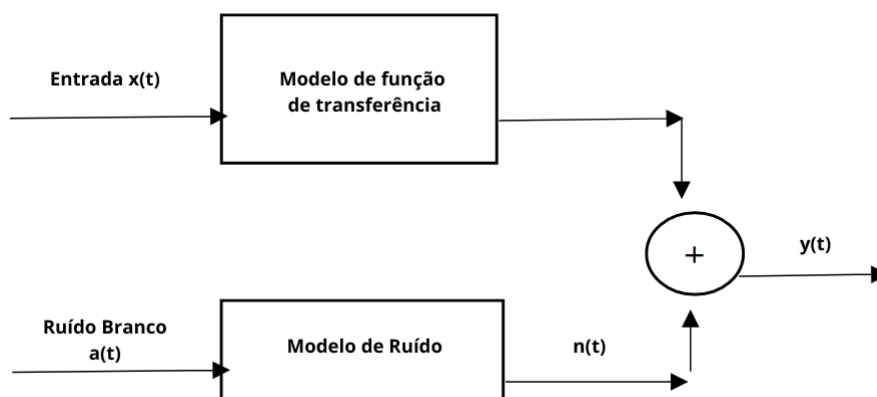
O AR, MA e ARMA são processos conhecidos como processos estacionários. Se o processo é não estacionário, é necessária uma transformação alcançada por diferenças no modelo ARMA. O modelo de transformação é chamado “Autoregressive Integrated Moving-Average” (ARIMA). O modelo função de transferência (FT) utiliza dos modelos AR, MA, ARMA ou ARIMA para representar o ruído branco e os dados de carga com uma ou mais variáveis como temperatura do ar, umidade, velocidade do vento, etc. (MOGHRAM, 1989) A figura 2 mostra como o modelo da FT pode ser utilizado.

2.5.3 Filtro Kalman

O filtro Kalman é um algoritmo utilizado para estimar adaptativamente o estado do modelo. Na previsão da carga, o comportamento de entrada e de saída de um sistema é representado por um modelo de estado e espaço com um filtro de Kalman utilizado para estimar o estado desconhecido do modelo. Assim, o filtro de Kalman utiliza o erro atual da previsão e os vigentes dados meteorológicos para estimar o próximo vetor de estado.

PARK et al. desenvolveu um modelo de estado para a carga nominal que consiste em três componentes: carga nominal, tipo de carga e carga residual. A carga nominal é modelada de forma que o filtro de Kalman possa ser utilizado, e os parâmetros do modelo são adaptados pelo método dos mínimos quadrados recursivos ponderados exponencialmente. O efeito dos dias da semana são representados pelo tipo de carga, que são acrescentados à carga nominal

Figura 2 – Modelo da Função de Transferência (FT)



Fonte: Adaptado de (ASHOUR; ABU; EL-FATTAH, 1991)

estimada pelo filtro de Kalman. O tipo de carga é determinado por suavização exponencial. Por fim, para contabilizar o erro de modelagem, a carga residual também é calculada.

Embora o filtro de Kalman seja utilizado em PCCP, este contém algumas limitações. Uma delas é a dificuldade de identificar os parâmetros do modelo de estado e espaço.

2.5.4 Sistemas especialistas

Um sistema especialista (SE) é um modelo computacional que compreende quatro partes principais: uma base de conhecimento, um banco de dados, um mecanismo de inferência e uma “interface” de usuário. A base de conhecimento é um conjunto de regras derivadas da experiência de especialistas humanos. A base de dados é uma coleção de fatos obtidos de especialistas humanos e informações saídas por meio de regras do mecanismo de interferência do sistema. O mecanismo de inferência é a parte do SE que “pensa”. Na previsão de carga, o SE usa algumas regras que normalmente são de natureza heurística para obter uma previsão precisa. O SE transformará as regras e procedimentos utilizados pelos especialistas humanos no software que consegue prever automaticamente sem assistência humana. Este processo traz vantagens, pois os sistemas especialista podem tomar decisões quando os especialistas humanos não estão disponíveis. Eles podem reduzir a carga de trabalho de especialistas humanos e tomar decisões rápidas em caso de emergência.

Para obter o melhor resultado ao utilizar sistemas especialistas, é necessária uma colaboração entre a disponibilidade do especialista humano e os desenvolvedores de software. Isso ocorre porque o tempo de transmissão do conhecimento do especialista para o software do SE deve ser considerado na elaboração do software do sistema.

2.5.5 Lógica fuzzy

A lógica booleana é o conceito central da lógica fuzzy (LF). Ela possui três partes principais: variáveis linguísticas, regras e conjuntos. As variáveis linguísticas são entradas do sistema. Estas entradas são transformadas em saída através de regras de conjuntos. A LF não precisa de equações matemáticas complicadas, pois consiste em variáveis linguísticas, então elas são compreensíveis. LF pode adaptar para novas condições ao adicionar ou eliminar alguma regra.

Entre as vantagens do uso da lógica fuzzy estão a ausência da necessidade de um modelo matemático mapeando as entradas para as saídas e a ausência da necessidade de entradas precisas. Com essas regras genéricas de condicionamento, sistemas de LF adequadamente projetados podem ser muito robustos quando aproveitados para previsão. [MIRANDA; MONTEIRO](#) e [SKARMAN; GEORGIPOULOUS](#) descrevem aplicações da lógica fuzzy para previsão de carga.

2.5.6 Redes neurais

Redes neurais artificiais (RNA) são implementadas em várias aplicações devido a sua habilidade de aprendizagem. RNAs são baseadas no neurônio biológico e, frequentemente utilizadas na previsão de carga. A ideia de usar redes neurais para previsão é a suposição de que existe uma função não linear que relaciona alguma variável externa a valores futuros da série temporal. [KYRIAKIDES; POLYCARPOU](#) comenta que existem três passos necessários que devem ser considerados ao utilizar modelos de RNA para previsão de séries temporais:

1. Projeto do modelo de rede neural. Por exemplo, escolha da rede neural que será empregada, números de camadas e número de parâmetros ou pesos ajustáveis.
2. Treino da rede neural. Este processo inclui selecionar o algoritmo de treino, os dados que serão utilizados para treinamento e o pré-processamento de dados.
3. Validação da rede neural, ou seja, utilizar a rede treinada em um conjunto de dados que não foram utilizados durante o estágio de treinamento.

A rede feedforward é frequentemente aplicada na previsão de cargas. Ela consiste em várias camadas sucessivas de neurônios com uma camada de entrada, várias camadas ocultas e uma camada de saída. O procedimento básico de aprendizado ou ajuste de peso é a retropropagação (backpropagation) que propaga o erro para trás e ajusta o peso de acordo.

3 Redes neurais artificiais

3.1 Introdução

As redes neurais artificiais (RNA) são modelos computacionais fundamentados no comportamento do sistema nervoso humano. A partir de dados armazenados, estes modelos são treinados para estabelecer correlações e interações entre as diferentes variáveis de determinado problema. O cérebro humano é um sistema de processamento paralelo, complexo e não linear. Assim, como no cérebro humano, nas redes neurais também existem componentes de processamento chamados neurônios artificiais, assemelhando-se ao cérebro em dois aspectos: (HAYKIN, 2009)

1. O conhecimento é adquirido pela rede neural de seu ambiente por meio de um processo de aprendizagem.
2. Fortes conexões interneurais, conhecidas como pesos sinápticos, são utilizados para armazenar o conhecimento adquirido.

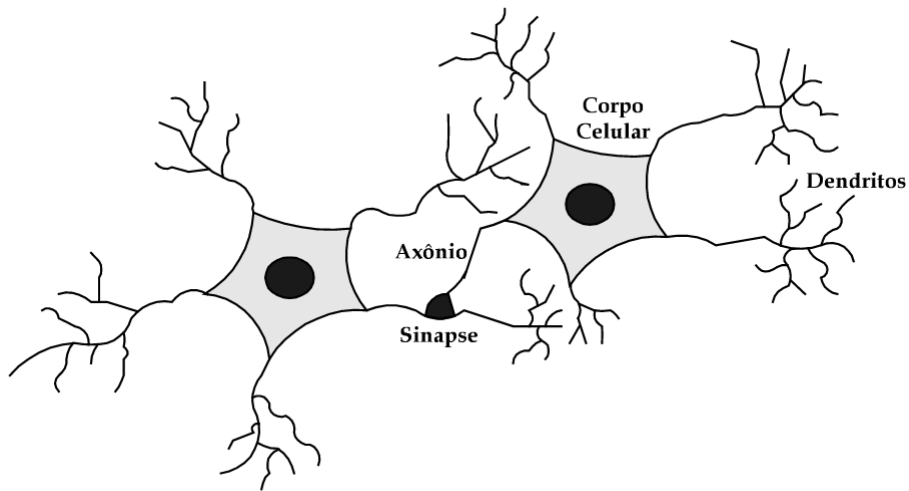
O processo de aprendizagem é conhecido como algoritmo de aprendizado, cuja função é modificar ordenadamente os pesos sinápticos da RNA de modo a atingir um objetivo do projeto desejado.

3.2 Modelo neural biológico

Existem aproximadamente 10 bilhões de neurônios no córtex humano (dos 86 bilhões estimados de neurônios do cérebro humano). Cada neurônio é conectado a vários milhares de outros neurônios sendo a velocidade de operação típica medida em milissegundos.

A maioria dos neurônios codifica suas ativações ou saídas como uma série de breves pulsos elétricos. O corpo celular do neurônio processa as ativações de entrada e as converte em ativações de saída. O núcleo dos neurônios contém o material genético na forma de DNA. Os dendritos são fibras que emanam do corpo celular e fornecem as zonas receptivas que recebem ativação de outros neurônios. Os axônios são fibras que atuam como linhas de transmissão que enviam ativação para outros neurônios. As junções que permitem a transmissão de sinal entre axônios e dendritos são chamadas sinapses. O processo de transmissão é por difusão de substâncias químicas chamadas neurotransmissores via a fenda sináptica. Na figura 3 são apresentados os componentes do neurônio biológico mais comum.

Figura 3 – Diagrama Esquemático Básico de um Neurônio Biológico

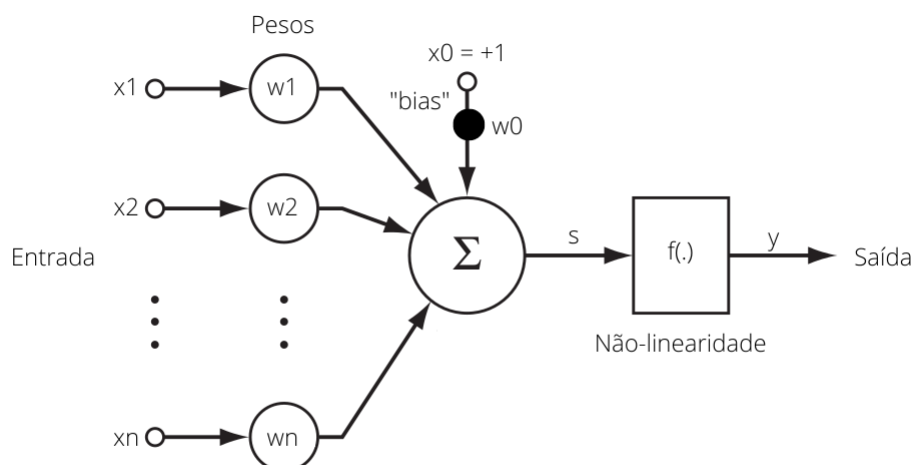


Fonte: (FILHO; CARVALHO, 1997)

3.3 Modelo neural artificial

O neurônio artificial é uma unidade de processamento de informação fundamental para a operação de uma rede neural. O modelo de McCulloch-Pitts (MCCULLOCH; PITTS, 1943) apresentado na figura 4 foi desenvolvido em 1943. Este modelo é o mais empregado na resolução de problemas como o reconhecimento de padrões.

Figura 4 – Modelo do Neurônio de McCulloch-Pitts



Fonte: Adaptado de (HAYKIN, 2009)

O modelo de McCulloch-Pitts descreve que a atividade s de um neurônio é a soma de n entradas ponderadas por seus respectivos pesos. O sinal y que será enviado aos demais

neurônios é resultado da interação da atividade s com uma função de ativação $f(.)$. As funções de ativação são importantes, haja visto que elas permitem a resolução de problemas não lineares. As funções de ativação mais frequentemente utilizadas são:

1. Relé

$$f(s) = \begin{cases} +1 & : \text{se } s \geq 0 \\ -1 & : \text{se } s < 0 \end{cases}$$

2. Sigmoide 1

$$f(s) = \frac{1 - e^{-\lambda s}}{1 + e^{-\lambda s}}$$

sendo:

λ = inclinação da curva.

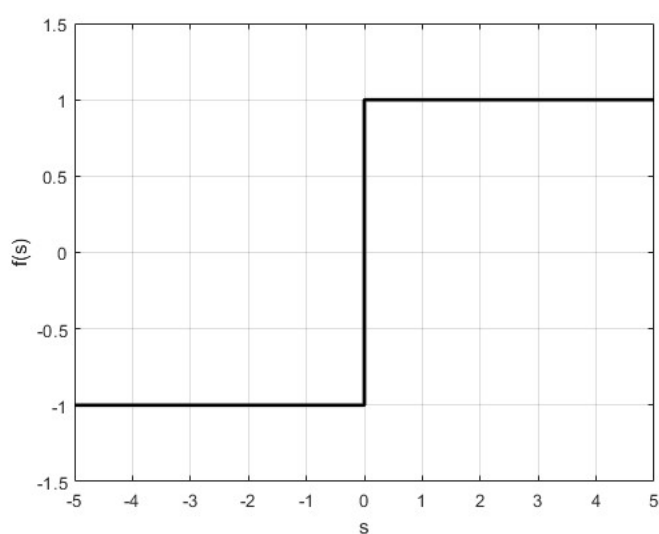
3. Sigmoide 2

$$f(s) = \frac{1}{1 + e^{-\lambda s}}$$

sendo:

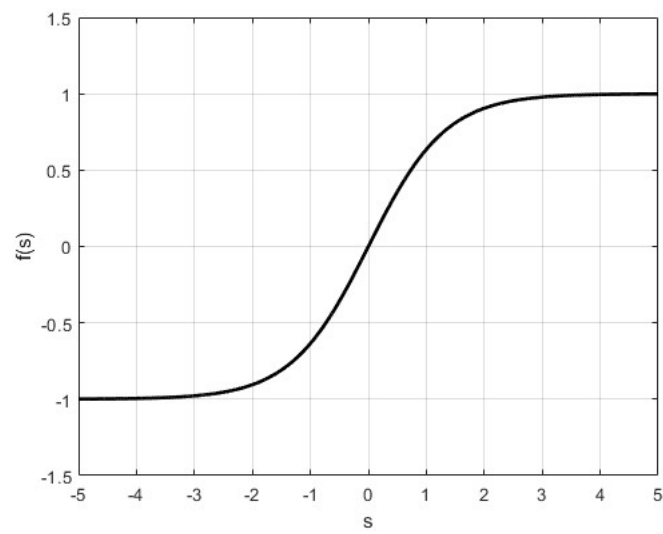
λ = inclinação da curva.

Figura 5 – Função Relé



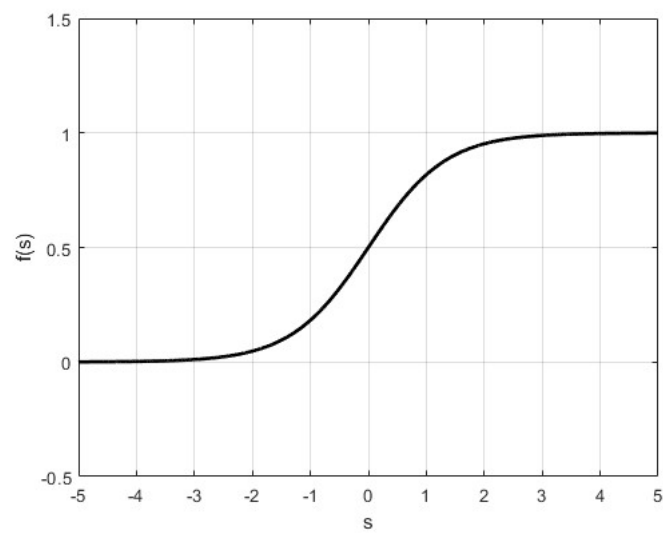
Fonte: Elaborado pelo autor.

Figura 6 – Função Sigmoide 1



Fonte: Elaborado pelo autor.

Figura 7 – Função Sigmoide 2

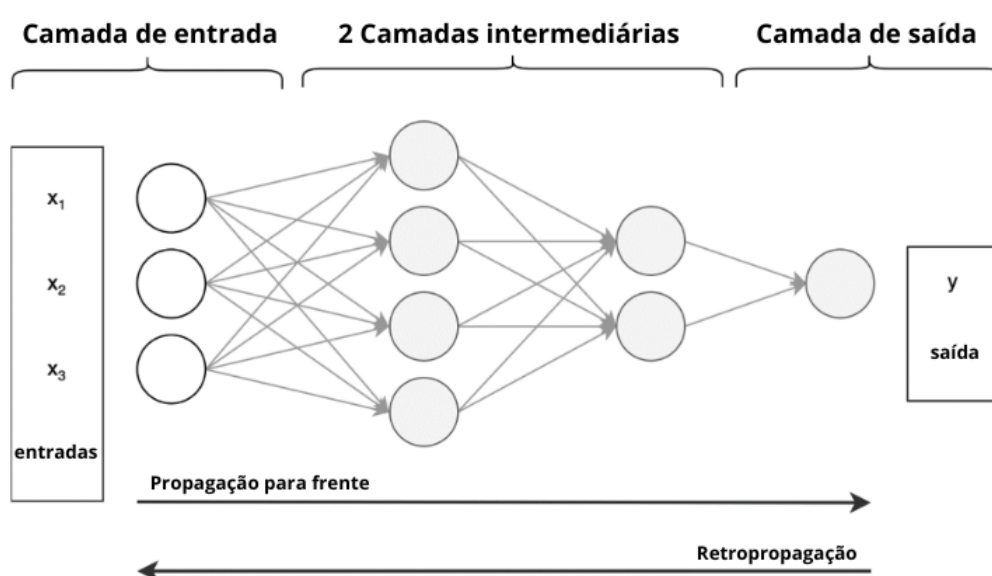


Fonte: Elaborado pelo autor.

3.4 Estrutura da rede neural

Neurônios em uma RNA podem ser divididos em três camadas: camada de entrada, camadas intermediárias (ou ocultas) e camada de saída. O número de camadas de uma RNA é a soma das camadas intermediária com a camada de saída. Como a camada de entrada é apenas um vetor de entrada, o número de neurônios pode ser calculado ao omitir a camada de entrada, ou seja, apenas neurônios da camada intermediária e de saída são considerados. (KARPATHY, n.d.)

Figura 8 – Rede Neural Artificial



Fonte: Adaptado de (KARPATHY, n.d.)

O trabalho de uma RNA pode ser descrito em ciclos, consistindo de duas fases: propagação para frente (forward propagation) e retropropagação (backpropagation). Quando a saída final da rede neural é calculada, é fundamental calcular o quão errado é a saída

O treinamento real de uma RNA consiste em ajustar os parâmetros: pesos e desvios. No início, uma RNA gera esses parâmetros aleatoriamente. Após o cálculo da função de custo, os parâmetros de uma rede serão ajustados durante a fase de retropropagação de acordo com um algoritmo de otimização. Normalmente, as instâncias de treinamento são processadas em lotes, tornando o procedimento de treinamento mais rápido, pois o cálculo do gradiente descendente e a atualização dos parâmetros da rede são demorados. (RUDER, 2016)

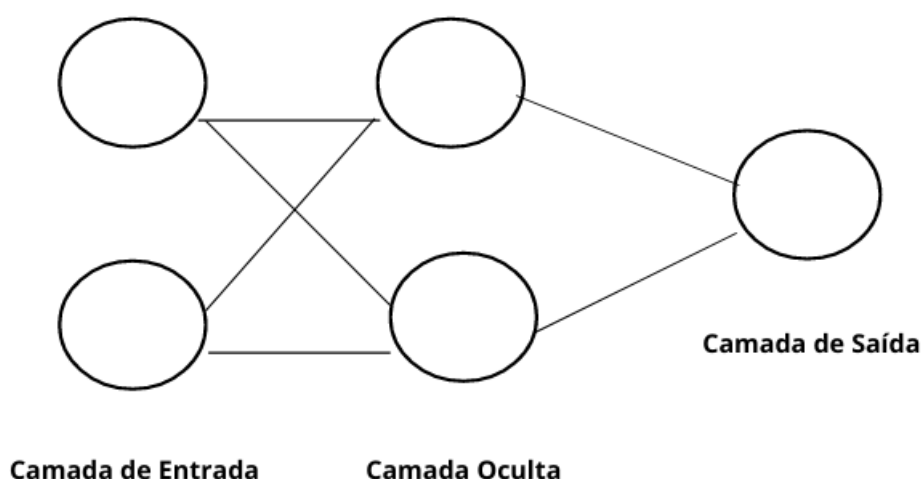
3.5 Arquiteturas de redes neurais

Existem várias arquiteturas de redes neurais, cada uma baseada em uma arquitetura diferente. A seguir serão apresentadas as principais classes de arquitetura de redes neurais.

3.5.1 Rede neural feedforward

Na rede neural feed-forward todos os nós estão totalmente conectados e os dados fluem em apenas uma direção, da camada de entrada para a camada de saída, sem nenhum loop de retorno. Esta abordagem tornou-se funcional na década de 1950. (HIPPERT; PERDREIRA; SOUZA, 2001) Na figura 9 a seguir é apresentado uma rede feedforward com uma camada oculta.

Figura 9 – Rede Neural Feedforward



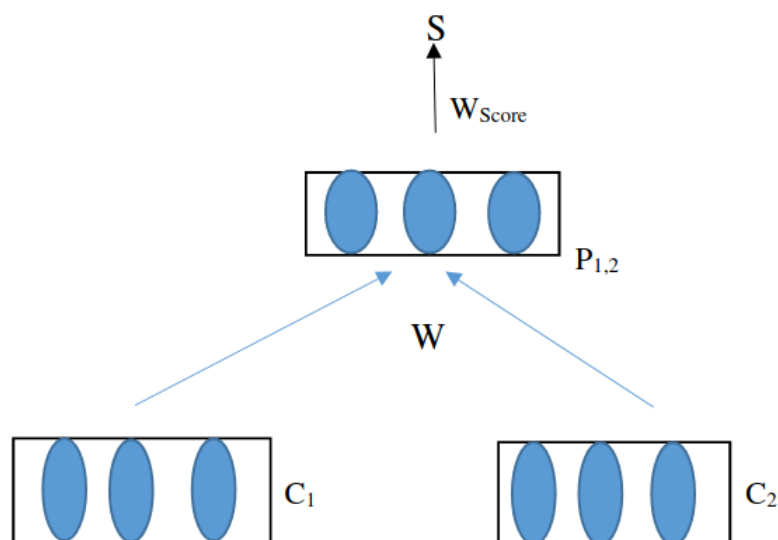
Fonte: Adaptado de (CHEN; CANIZARES; SINGH, 2001)

3.5.2 Rede neural recursiva

Redes neurais recursivas são modelos adaptativos não lineares com a capacidade de aprender informações estruturadas em profundidade. Elas foram introduzidos no final da década de 90. (FRASCONI; GORI; SPERDUTI, 1998) Estes modelos computacionais são adequados tanto para problemas de regressão quanto de classificação conseguindo realizar tarefas de treinamento não supervisionado e supervisionado.

A rede é feita aplicando o mesmo conjunto de pesos recursivamente sobre uma estrutura para realizar uma previsão sobre estruturas de entrada de tamanho variável, percorrendo a estrutura fornecida em ordem topológica. As redes neurais recursivas são muito poderosas no aprendizado de estruturas hierárquicas em forma de árvore. Esses modelos não foram amplamente aceitos, por causa da complexidade inerente e, também, em razão ao caso da fase de treinamento computacionalmente cara. (C.; A., 1996)

Figura 10 – Rede Neural Recursiva



Fonte: (FRASCONI; GORI; SPERDUTI, 1998)

3.5.3 Rede neural perceptron multicamadas

As redes neurais perceptron multicamadas têm mais de duas camadas totalmente conectadas e, são utilizadas quando os dados não são separados linearmente. Nesta RNA as funções de ativação não são lineares. A estrutura perceptron multicamada é aproveitada principalmente em programas de tradução e reconhecimento de fala. A figura 11 ilustra uma rede neural perceptron multicamadas.

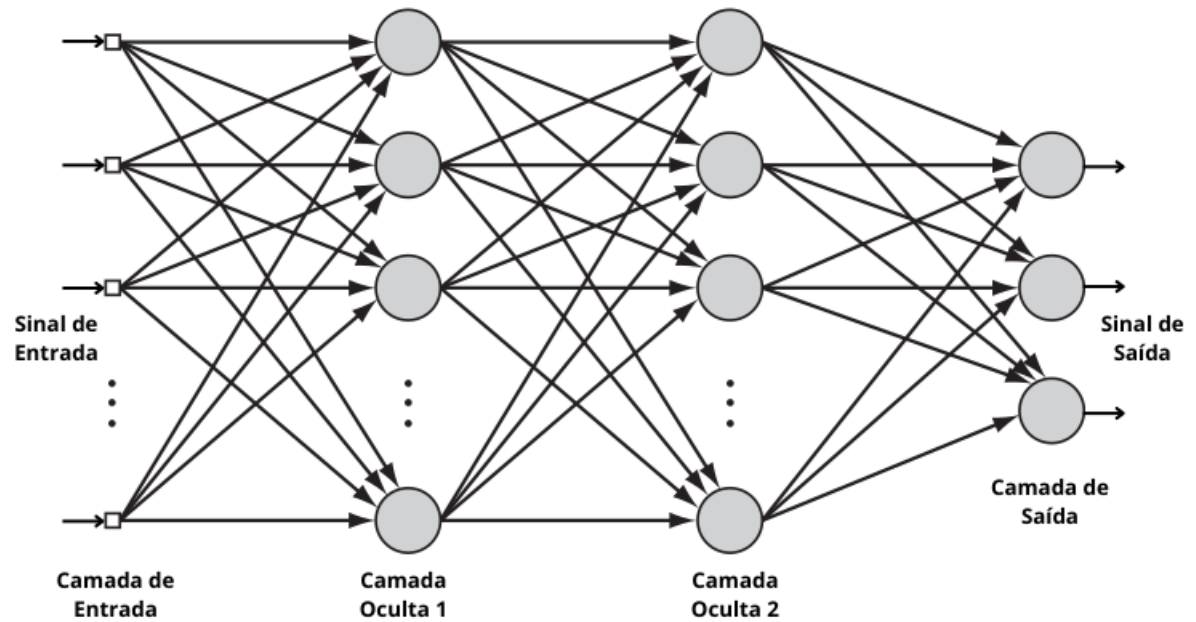
3.5.4 Rede neural convolucional

As redes neurais convolucionais possuem pelo menos uma camada convolucional. Em alguns casos, as camadas convolucionais são completamente interconectadas, mas também podem ser agrupadas. Redes neurais convolucionais são utilizadas principalmente em programas de processamento de linguagem e reconhecimento de vídeo ou foto pois a camada convolucional ajuda a rede a resolver problemas mais complexos com menos parâmetros. A rede convolucional também é muito útil no processamento de sinais e imagens. (BENGIO; DUCHARME; VINCENT, 2003)

3.5.5 Rede neural recorrente

Uma rede neural recorrente possui pelo menos um loop de feedback. Ela pode consistir em uma única camada de neurônios com cada neurônio alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios. Autofeedback refere-se a uma situação em que a saída de um neurônio é realimentada em sua própria entrada. A presença de loops de feedback

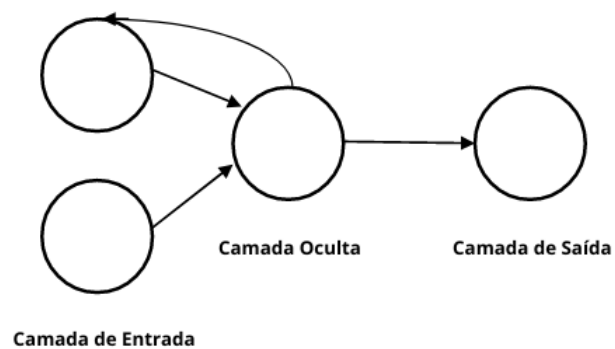
Figura 11 – Rede Neural Perceptron Multicamadas Com Duas Camadas Ocultas.



Fonte: adaptado de (HAYKIN, 2009)

tem um impacto profundo na capacidade de aprendizagem da rede e no seu desempenho. (WERBOS, 1988)

Figura 12 – Rede Neural Recorrente



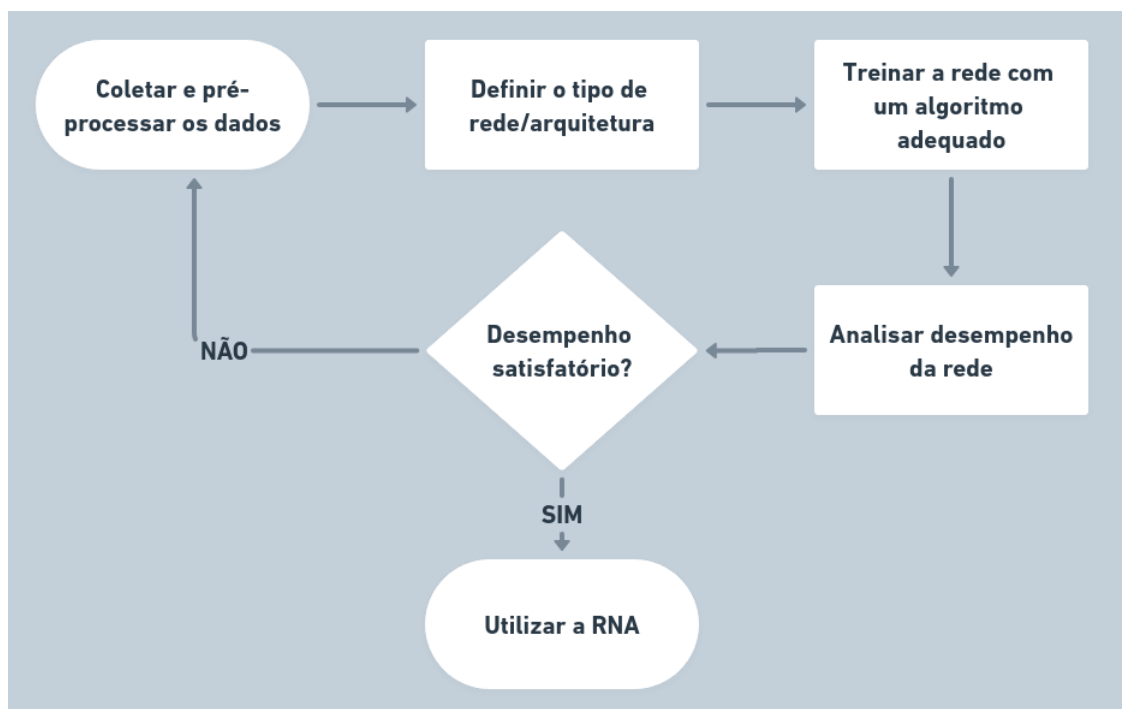
Fonte: Adaptado de (WERBOS, 1988)

4 Projeto de redes neurais para previsão

4.1 Introdução

Projetar redes neurais para previsão é um procedimento iterativo que começa com a coleta e pré-processamento de dados para tornar o treinamento mais eficiente. Em seguida, os dados de treinamento precisam ser divididos em conjuntos de treinamento, validação e teste. Depois disso, o tipo de rede e a arquitetura apropriados para a previsão são escolhidos. Uma vez que a rede e a arquitetura tenham sido decididas, o próximo passo é selecionar um algoritmo de treinamento apropriado para o problema de previsão. Após treinar a rede, é preciso analisá-la para verificar se o desempenho é satisfatório. Caso seja encontrado algum problema é necessário reiniciar o processo, conforme o fluxograma da figura 13. (HAGAN; DEMUTH; BEALE, 1996)

Figura 13 – Fluxograma do Processo de Treinamento de Rede Neural



Fonte: Elaborado pelo autor.

4.2 Coleta de dados e pré-processamento

A primeira etapa no projeto da rede neural para previsão é a coleta de dados. A escolha do conjunto de dados está intimamente relacionada à escolha do número de neurônios na rede

neural. Para tornar o problema de previsão mais eficiente, os dados são pré-processados antes de serem utilizados para treinar a rede neural. Existem vários tipos de pré-processamento de dados, como normalização, transformações não lineares, extração de recursos, codificação de entradas/destinos discretos, manipulação de dados ausentes, etc.

A normalização é o principal passo no pré-processamento de dados. Ela auxiliará a rede neural a extrair informações relevantes no processo de treinamento. Geralmente, existem dois métodos para normalização. O primeiro método é normalizar os dados para que eles obedeçam em um intervalo padrão, normalmente de -1 a 1. Isso pode ser feito por: (HAGAN; DEMUTH; BEALE, 1996)

$$p^n = \frac{2(p - p^{min})}{p^{max} - p^{min}} - 1 \quad (4.1)$$

sendo p^{min} o vetor que contém os valores mínimos e p^{max} os valores máximos de cada elemento dos vetores de entrada. O outro método de normalização é ajustar os dados para eles terem média e variância especificadas normalmente 0 e 1:

$$p^n = \frac{(p - p^{mean})}{p^{std}} \quad (4.2)$$

em que p^{mean} é a média dos vetores de entrada no conjunto de dados e p^{std} é o vetor que contém os desvios padrão de cada elemento do vetor de entrada.

Outra alternativa de pré-processamento é a remoção de valores discrepantes, valores ausentes ou quaisquer irregularidades. As redes neurais podem ser sensíveis a dados defeituosos. Não é possível garantir o desempenho da rede quando as entradas das redes estão fora do alcance do conjunto de treinamento. Dados ausentes podem ocorrer nas entradas e/ou destinos (por exemplo, dados históricos de carga não disponíveis por um mês). Se os dados ausentes estiverem em uma variável de entrada, uma alternativa é substituir os dados ausentes pelo valor médio da variável de entrada específica. Quando ocorrem dados ausentes em um elemento do destino, o índice de desempenho pode ser modificado para que os erros associados aos valores de destino ausentes não contribuam para o índice de desempenho. (HAGAN; DEMUTH; BEALE, 1996)

Uma vez que a coleta de dados e o pré-processamento tenham sido implementados, os dados serão divididos em três conjuntos: treinamento, validação e teste. O conjunto de treinamento geralmente compõe aproximadamente 70% do conjunto de dados e, a validação e teste compõem aproximadamente 15% cada. O gradiente é calculado no conjunto de treinamento e o erro quadrático médio no conjunto de treinamento é minimizado. Durante o treinamento, o erro de validação é monitorado. O treinamento é interrompido quando o erro de validação aumenta, para evitar o sobreajuste (overfitting). Após a conclusão do treinamento, o

erro no conjunto de teste é calculado. Isso fornece uma indicação de como a rede funcionará em novos dados. (HAGAN; DEMUTH; BEALE, 1996)

4.3 Tipo de rede e arquitetura

O objetivo da previsão é prever o valor futuro de algumas séries temporais, por exemplo, a previsão de carga de curto prazo prevê a carga até 24 horas à frente. Como muitas previsões requerem modelos não lineares, redes neurais dinâmicas podem ser usadas. Neste trabalho, utilizaremos o modelo feedforward. Para esta rede é necessário selecionar o número de neurônios nas camadas ocultas e o número de camadas ocultas.

4.3.1 Número de camadas ocultas

Para determinar o número de camadas ocultas, o procedimento padrão é iniciar a rede com apenas uma camada oculta. Se o desempenho da rede de duas camadas não for bom o suficiente, uma rede de três camadas poderá ser usada. Foi demonstrado que uma ou duas camadas ocultas são boas o suficiente para a previsão de redes neurais. (HIPPERT; PERDREIRA; SOUZA, 2001) É incomum utilizar mais de duas camadas ocultas haja visto que, se houver várias camadas ocultas, o treinamento se torna mais complexo. Cada camada oculta executa uma operação de esmagamento (squashing), pois a função de ativação na camada oculta deve ser diferenciável e não decrescente, ou seja, tangente logística ou hiperbólica. Isso faz com que as derivadas do desempenho em relação aos pesos nas camadas iniciais sejam bastante pequenas, o que pode causar convergência lenta para otimização de descida mais acentuada.

4.3.2 Número de neurônios nas camadas ocultas

Também é fundamental definir o número de neurônios em cada camada. Esta tarefa no entanto não é tão simples. Caso o número de neurônios seja escasso, o modelo não será flexível o suficiente para modelar adequadamente os dados. Por outro lado, se for abundante, o modelo irá sobreajustar os dados. O procedimento padrão na escolha do número de neurônios nas camadas ocultas é começar com mais neurônios do que o necessário, o que pode fazer com que a rede superajuste os dados. Para evitar o sobreajuste, pode-se utilizar a parada antecipada ou regularização bayesiana. Na regularização bayesiana, a norma dos pesos é penalizada, reduzindo o número efetivo de parâmetros. Se, após o treinamento, o número efetivo de parâmetros for muito menor que o número total de parâmetros, o número de neurônios pode ser reduzido e a rede treinada novamente. (HIPPERT; PERDREIRA; SOUZA, 2001)

4.4 Treinamento da rede neural

Existem alguns procedimentos para realizar um treinamento de uma rede neural, são eles: inicialização dos pesos, escolha do algoritmo de treinamento e escolha do critério para interromper o treinamento.

4.4.1 Pesos iniciais

A inicialização dos pesos dependerá da rede. Em geral, os pesos e os desvios são definidos como pequenos valores aleatórios (por exemplo, uniformemente distribuídos entre -0,5 e 0,5, se as entradas forem normalizadas para cair entre -1 e 1). Não é recomendado definir os pesos e desvios para zero, porque a condição inicial pode cair em um ponto de sela da superfície de desempenho. Por outro lado, se os pesos iniciais forem muito grandes, a condição inicial pode cair na parte plana da superfície de desempenho, devido à saturação das funções de transferência sigmóides. Uma ótima abordagem para definição dos pesos iniciais e as tendências para uma rede de duas camadas foi introduzida por [NGUYEN; WIDROW](#) em 1990. Este método gera pesos iniciais e valores de polarização para uma camada, de modo que as regiões ativas dos neurônios sejam distribuídas aproximadamente uniformemente pelo espaço de entrada.

4.4.2 Classificação de algoritmos de treinamento

Os algoritmos de aprendizagem de máquina podem ser classificados em três divisões de acordo com seu objetivo:

1. Aprendizado supervisionado

No aprendizado supervisionado a regra de aprendizado é fornecida com um conjunto de dados de treinamento de comportamento de rede adequado. À medida que as entradas são aplicadas à rede, as saídas são comparadas com os alvos. A regra de aprendizado é então utilizada para ajustar os pesos e desvios da rede para aproximar as saídas da rede dos alvos. ([WIDROW; LEHR, 1990](#))

2. Aprendizado não supervisionado

No aprendizado não supervisionado os pesos e desvios são modificados em resposta apenas às entradas da rede. Não há saídas de destino disponíveis. A maioria desses algoritmos executa algum tipo de operação de agrupamento. Eles aprendem a categorizar os padrões de entrada em um número finito de classes. ([HAYKIN, 2009](#))

3. Aprendizado por reforço

O aprendizado por reforço é aproveitado por diferentes aplicações para alcançar a melhor solução e escolher o melhor caminho em cada caso. Os dados de treinamento

no aprendizado supervisionado têm a saída e o algoritmo usa essas chaves de resposta no processo de treinamento. No entanto, o aprendizado por reforço não possui essas chaves de resposta. Portanto, não há escolha de algoritmo para aprendizado, exceto a experiência. (HAYKIN, 2009)

4.4.3 Algoritmo de treinamento

Algoritmo de treinamento é um conjunto de regras bem definidas para a solução de um problema de treinamento. Este procedimento é usado para modificar os pesos e desvios de uma rede neural artificial sendo objetivo do algoritmo treinar a RNA para realizar determinada tarefa.

A retropropagação, também conhecida como backpropagation, foi proposta por WERBOS, em 1974, e é um dos algoritmos mais utilizados no treinamento de redes neurais. Este algoritmo é fundamentado na regra de aprendizado delta de Widrow-Hoff. Neste método, inicialmente, todos os pesos e tendências para diferentes camadas ocultas são atribuídos aleatoriamente. Os dados ou informações são fornecidos à rede, a RNA é ativada e uma saída é registrada. Esta saída é então comparada com a saída alvo e um erro é gerado. Em seguida, o erro gerado é propagado de volta para as camadas ocultas e os pesos são ajustados de acordo. O processo se repete até que o erro mínimo seja alcançado. O fluxograma do treinamento da rede neural utilizando o algoritmo de retropropagação é apresentado na figura 14.

O processo de treinamento por retropropagação é descrito abaixo em dez etapas: (RAMEZANI; FALAGHI; HAGHIFAM, 2005)

1. Inicializar os pesos da rede
2. Adicionar a entrada ponderada e aplicar a função de ativação para calcular a saída da camada oculta:

$$h_j = f [\sum_i X_i W_{ij}] \quad (4.3)$$

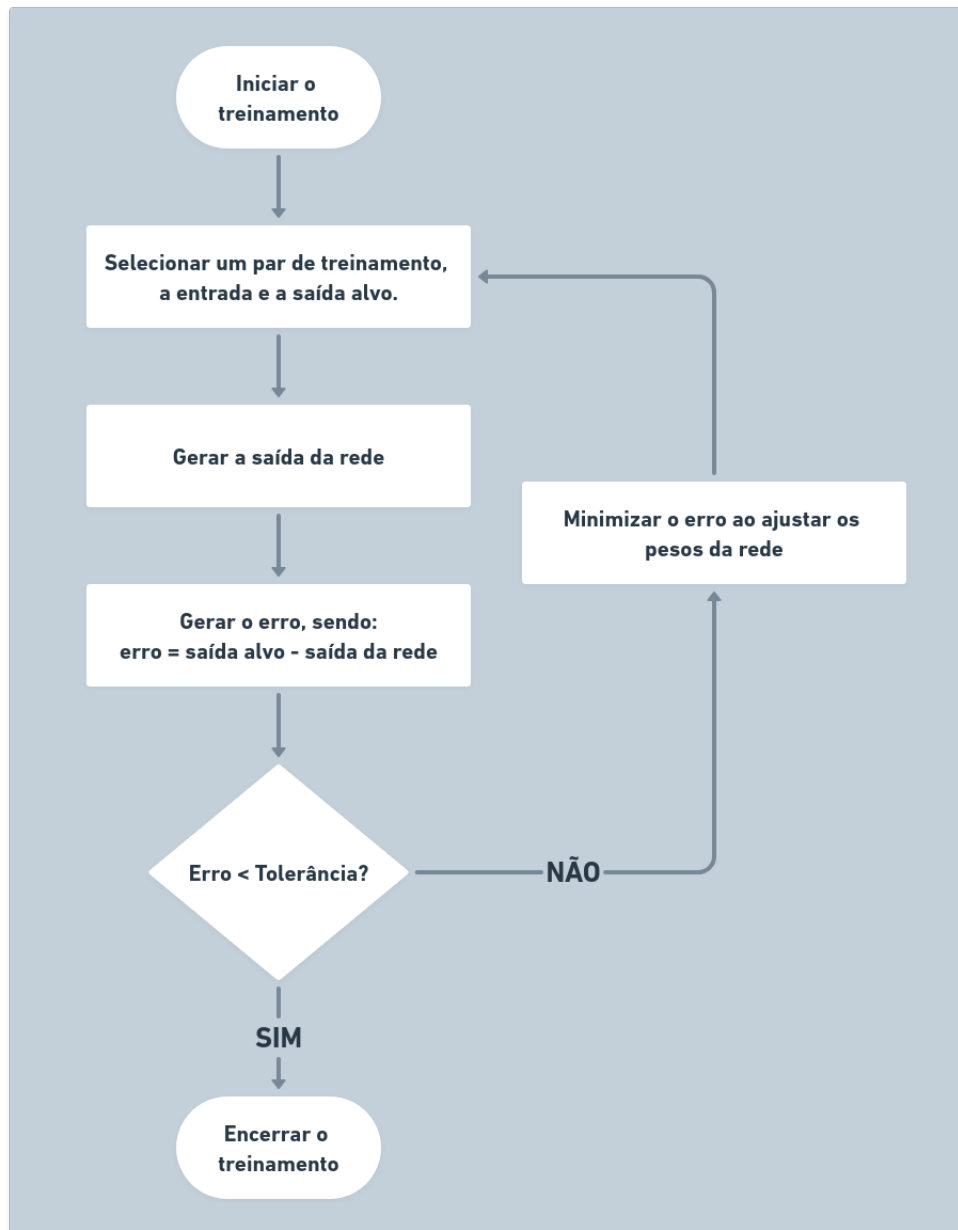
em que,

- h_j = saída do neurônio oculto j para entrada i;
- f = função de ativação;
- X_i = dados de entrada do neurônio de entrada i;
- W_{ij} = pesos sinápticos entre o neurônio de entrada i e neurônio oculto j.

3. Adicionar a saída ponderada da camada oculta e aplicar a função de ativação para calcular a saída da camada de saída:

$$O_j = f [\sum_j h_j W_{jk}] \quad (4.4)$$

Figura 14 – Fluxograma Treinamento Retropapagação



Fonte: Adaptado de (RAMEZANI; FALAGHI; HAGHIFAM, 2005)

sendo,

O_k = saída do neurônio de saída k;

f = pesos sinápticos entre o neurônio oculto j e o neurônio de saída k.

4. Calcular o erro de retropropagação:

$$\delta_k = (t_k - O_k)f'(O_k) \quad (4.5)$$

em que,

f' = derivada da função de ativação;

t_k = alvo do neurônio de saída k.

5. Calcular o termo de correção de peso:

$$\Delta W_{jk}(n) = \eta \delta_k h_j + a \Delta W_{jk}(n-1) \quad (4.6)$$

sendo,

η = proporção de treinamento;

a = coeficiente de momento.

6. Adicionar a entrada delta para cada neurônio oculto e gerar um erro:

$$\delta_j = \sum_k \delta_k W_{jk} f'(\sum_i X_i W_{ij}) \quad (4.7)$$

7. Calcular termo de correção de peso:

$$\Delta W_{ij}(n) = \eta \delta_j X_i + a \Delta W_{ij}(n-1) \quad (4.8)$$

8. Atualizar os pesos:

$$W_{jk}(n+1) = W_{jk}(n) + \Delta W_{jk}(n) \quad (4.9)$$

$$W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij}(n) \quad (4.10)$$

9. Repetir a etapa 2 para o seguinte número de erros:

$$MSE = \frac{1}{2p} \left[\sum_P \sum_K (D_K^P - o_K^P)^2 \right] \quad (4.11)$$

em que,

p = padrões no conjunto de treinamento;
 MSE = erro quadrático médio.

10. Treinamento completo.

Existem diversos algoritmos de otimização que podem ser utilizados. De acordo com [HAGAN; DEMUTH; BEALE](#) o algoritmo de Levenberg-Marquardt e algumas variações do algoritmo de gradiente conjugado produzem os melhores resultados. Levenberg-Marquardt combina o método gradiente de descida com o método de Gauss-Newton. Este algoritmo herda a vantagem de velocidade do algoritmo de Gauss-Newton e a estabilidade da descida mais íngreme.

4.4.4 Critérios para interromper o treinamento

Após a escolha do algoritmo de treinamento, é necessário definir alguns critérios para interromper o treinamento. A rede deve interromper o treinamento quando o erro atingir alguns limites especificados. O critério mais simples é interromper o treinamento após um número fixo de iterações ter sido alcançado. Mas como também é difícil saber quantas iterações serão necessárias, o número máximo de iterações é geralmente definido razoavelmente alto. Outro critério de parada é a norma do gradiente do índice de desempenho. Se essa norma atingir um limite suficientemente pequeno, o treinamento poderá ser interrompido. Como o gradiente deve ser zero no mínimo do índice de desempenho, esse critério interromperá o algoritmo quando ele se aproximar do mínimo. O limite para a norma mínima deve ser definido para um valor muito pequeno (por exemplo, 10^{-6} para índices de erro quadrático médio, com alvos normalizados) para que o treinamento não termine prematuramente. Quando a parada antecipada é utilizada, o treinamento será interrompido quando o desempenho no conjunto de validação aumentar para um número definido de iterações. Além de evitar o sobreajuste, esse procedimento de parada também proporciona uma redução significativa na computação. Para a maioria dos problemas práticos, o erro de validação aumentará antes que qualquer outro critério de parada seja alcançado. [HAGAN; DEMUTH; BEALE](#)

4.5 Análise de desempenho da rede

É necessário testar o desempenho da rede após a conclusão do treinamento. Existem dois conceitos importantes ao avaliar uma rede treinada:

1. o erro de previsão não deve ser correlacionado no tempo
2. o erro de previsão não deve ser correlacionado com a sequência de entrada

Se os erros de previsão estão correlacionados no tempo, é necessário melhorar a previsão original. Para testar a correlação dos erros de previsão no tempo, pode-se utilizar a autocorrelação amostral:

$$R_e(\tau) = \frac{1}{Q - \tau} \sum_{t=1}^{Q-\tau} e(t)e(t + \tau) \quad (4.12)$$

Caso os erros de previsão não forem correlacionados (ruído branco), então é esperado que $R_e(\tau)$ seja próximo de zero, exceto quando $\tau = 0$. Para determinar se $R_e(\tau)$ é próximo de zero, é definido um intervalo de confiança aproximado de 95% a partir do intervalo:

$$-\frac{2R_e(0)}{\sqrt{Q}} < R_e(\tau) < \frac{2R_e(0)}{\sqrt{Q}} \quad (4.13)$$

O erro $e(t)$ é branco, se $R_e(\tau)$ satisfaz 4.12 para $\tau \neq 0$. Para testar a correlação entre os erros de previsão e a sequência de entrada, utiliza-se a função de correlação cruzada amostral:

$$R_{pe}(\tau) = \frac{1}{Q - \tau} \sum_{t=1}^{Q-\tau} p(t)e(t + \tau) \quad (4.14)$$

Se não houver correlação entre os erros de previsão e a sequência de entrada, então é esperado que $R_{pe}(\tau)$ seja próximo de zero para todos os τ . Para determinar se $R_{pe}(\tau)$ é próximo de zero, é definido um intervalo de confiança aproximado de 95% a partir do intervalo:

$$-\frac{2\sqrt{R_e(0)}\sqrt{R_p(0)}}{\sqrt{Q}} < R_{pe}(\tau) < \frac{2\sqrt{R_e(0)}\sqrt{R_p(0)}}{\sqrt{Q}} \quad (4.15)$$

Em suma, uma rede neural pode ser considerada adequadamente treinada se: (ADYA; COLLOPY, 1998)

- está bem ajustada aos dados de treinamento
- seus desempenhos na amostra de treinamento e nas amostras de teste são semelhantes
- seus desempenhos em diferentes amostras de teste são coerentes.

4.6 Erro de previsão

O erro de previsão é calculado comparando a carga prevista gerada pela RNA com a carga real. Este pode ser apresentado em termos do erro quadrático médio (RMSE) com unidades de kW conforme apresentado em 4.16. (HIPPERT; PERDREIRA; SOUZA, 2001) Todavia o erro de previsão é geralmente mostrado em termos do erro percentual médio absoluto

(MAPE) com unidades de porcentagem conforme mostrado em 4.17. (SANTOS; MARTINS; PIRES, 2004)

$$RMSE = \sqrt{\frac{1}{N} \times \sum_{t=1}^n (y_t - \bar{y}_t)^2} \quad (4.16)$$

$$MAPE = \frac{1}{N} \times \sum_{t=1}^N \left(\frac{|y_t - \bar{y}_t|}{y_t} \times 100 \right) \quad (4.17)$$

sendo,

N = número de amostras;

y_t = carga alvo no tempo t;

$\bar{y}_t$ = carga prevista no tempo t.

Na literatura o erro percentual médio absoluto deve ser igual ou inferior a 5% para que a previsão realizada seja utilizada de forma confiável. (SRINIVASAN et al., 1998)

Outro erro importante a considerar durante uma previsão é o erro máximo percentual (Emax):

$$Emax (\%) = \max \left(\frac{|y_t - \bar{y}_t|}{y_t} \right) * 100 \quad (4.18)$$

5 Metodologia e resultados

5.1 Introdução

Neste capítulo, será apresentada a metodologia empregada e os resultados da rotina de previsão de carga elétrica global a curto prazo utilizando uma rede neural feedforward e algoritmo de treinamento retropapagação Levenberg-Marquardt.

5.2 Metodologia

Para realizar a rotina de previsão de carga elétrica global foi usada a ferramenta Neural Network Training (nntraintool) do software Matlab R2018b. (MATLAB, 2018) Além disso, todos os gráficos apresentados neste capítulo também foram coletados do Matlab sendo que alguns foram gerados automaticamente pela ferramenta nntraintool e outros gerados via código.

O conjunto de dados utilizados nesta rotina de previsão foram disponibilizadas pela Comissão de Eletricidade da Nova Zelândia e coletadas no Centralized Dataset (CDS, 2010). A base de dados conta com dados do período de 1 de janeiro de 2008 a 8 de janeiro de 2009 totalizando um conjunto com 17952 amostras (374 dias x 48 amostras por dia) que foram divididas em 70% para o conjunto de treinamento, 15% para o conjunto de validação e 15% para o conjunto de teste. A base de dados contém informações sobre o dia da semana, dias atípicos, horário do dia a cada 30 minutos (48 amostras/dia), carga (h-2), carga(h-1), carga(h) e carga real.

O dia escolhido para realizar a previsão foi 8 de janeiro de 2009. A fim de observar a influência do número de amostras na previsão, adotou-se três períodos distintos apresentados na tabela 1.

Tabela 1 – Períodos utilizados para treinamento da RNA

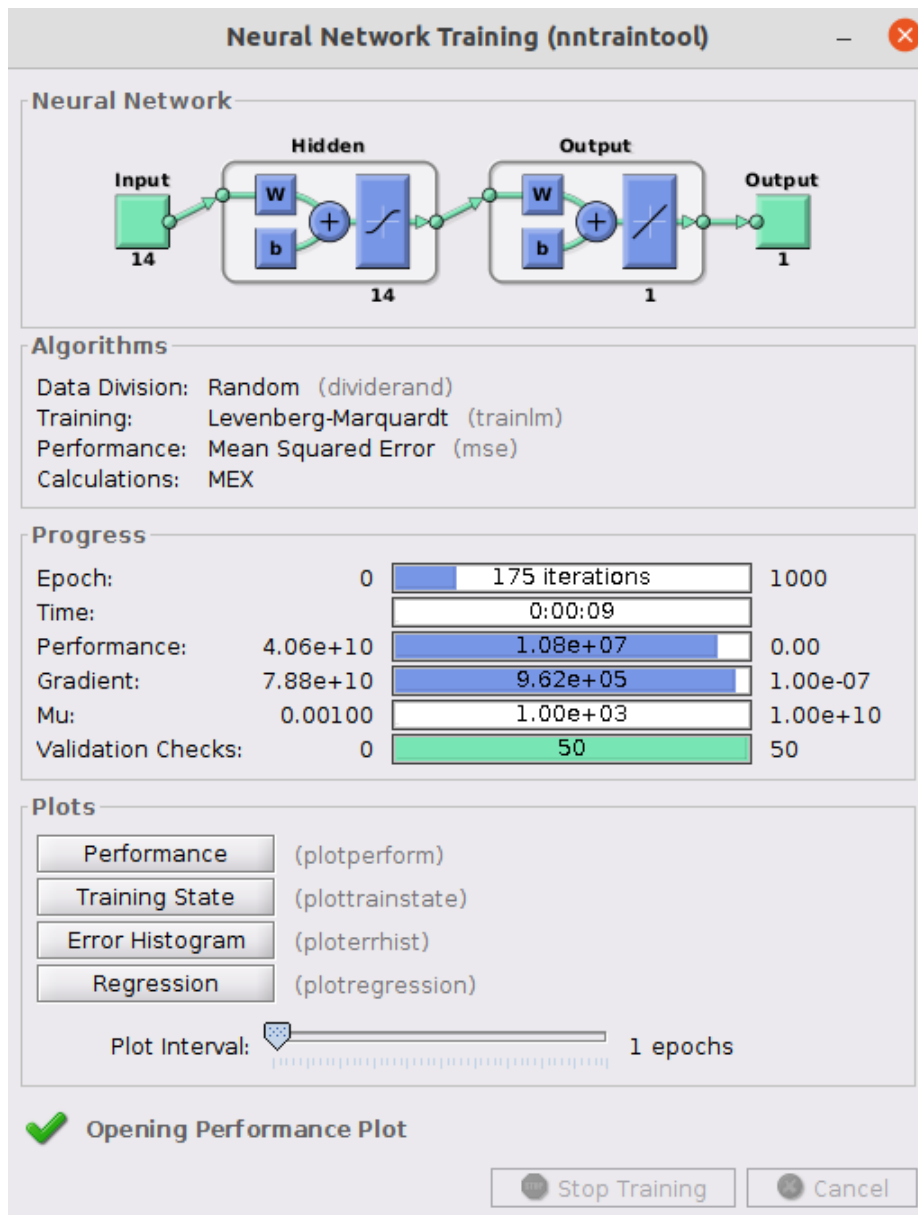
Período	Intervalo de treinamento	Número de Amostras	Dia Previsto
1	01/09/08 a 07/01/09	6240	08/01/09
2	01/05/08 a 07/01/09	12145	08/01/09
3	01/01/08 a 07/01/09	17952	08/01/09

Fonte: Elaborado pelo autor.

Para avaliar o desempenho da rede feedforward em relação à configuração de suas camadas, foram testadas as configurações apresentadas na tabela 2.

As redes neurais das configurações 1-4 seguem o modelo da figura 16 sendo n o número de neurônios na camada oculta. As redes neurais das configurações 5-8 seguem o modelo

Figura 15 – Ferramenta Neural Network Training (nntraintool)



Fonte: (MATLAB, 2018)

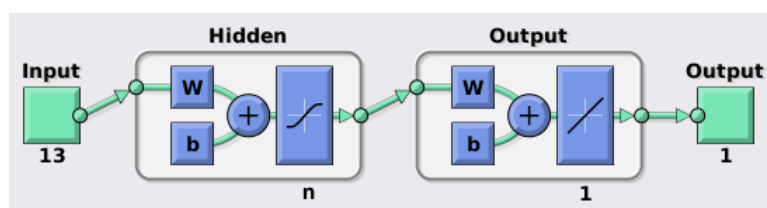
Tabela 2 – Configurações de Camada

Configuração	Número de Camadas	Número de neurônios por camada
1	2	5-1
2	2	14-1
3	2	20-1
4	2	30-1
5	3	14-5-1
6	3	14-14-1
7	3	14-20-1
8	3	20-20-1

Fonte: Elaborado pelo autor.

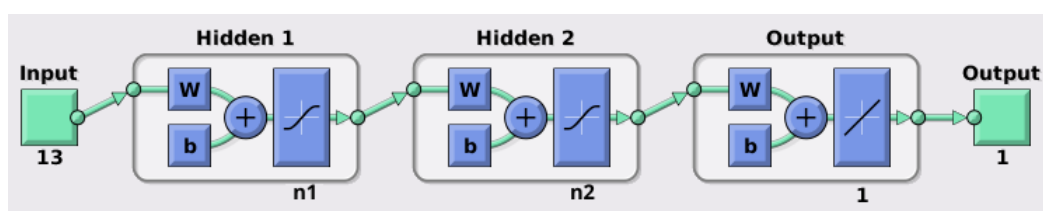
da figura 17 sendo n_1 o número de neurônios na primeira camada oculta e n_2 o número de neurônios na segunda camada oculta. A função de ativação das camadas ocultas é uma função sigmoide simétrica, enquanto na camada de saída, a função de ativação é uma função linear.

Figura 16 – Rede Neural Feedforward com 2 Camadas



Fonte: Adaptado de (MATLAB, 2018)

Figura 17 – Rede Neural Feedforward com 3 Camadas



Fonte: Adaptado de (MATLAB, 2018)

O algoritmo de treinamento utilizado (trainFcn) foi o algoritmo Levenberg-Marquardt Backpropagation. Os demais de treinamento estão apresentados na tabela 3.

Por fim, para verificar a precisão da rotina de previsão de carga elétrica foram utilizados os erros percentuais médios absolutos e os erros máximos percentuais conforme apresentado na seção 4.6.

Tabela 3 – Parâmetros de Treinamento da Rede Neural

Parâmetro	Nome da Variável (MATLAB)	Valor
Número máximo de épocas	epochs	1000
Erro Mínimo de Validação	max_fail	20
Descida de Gradiente Mínimo	min_grad	10^{-7}
Mu - Parâmetro de Controle	mu	10^{-3}
Mu - Fator de Diminuição	mu_dec	10^{-1}
Mu - Fator de Aumento	mu_inc	10
Mu - Máximo Valor	mu_max	10^{10}
Taxa de Treinamento	trainRatio	0.7
Taxa de Validação	valRatio	0.15
Taxa de Teste	testRatio	0.15

Fonte: Elaborado pelo autor.

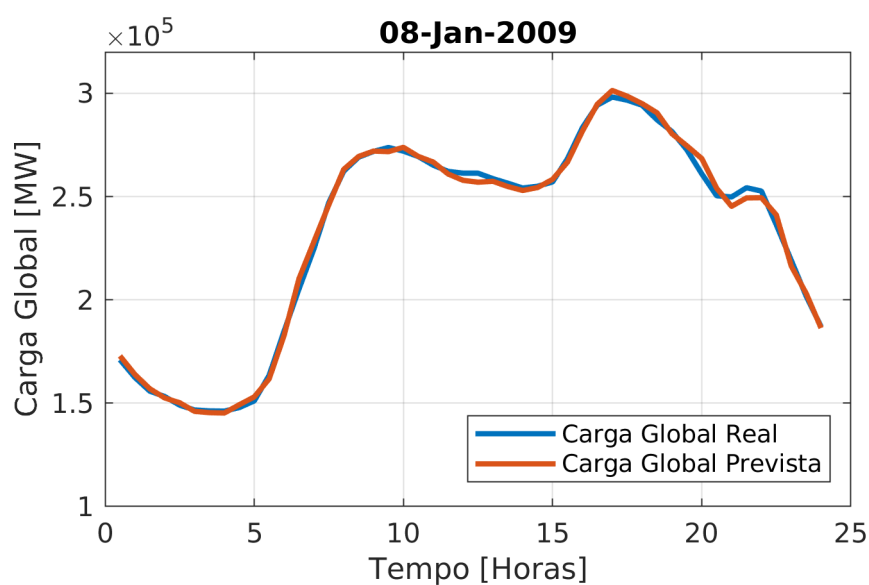
5.3 Resultados

Os resultados obtidos com a rede neural feedforward estão apresentados a seguir. Foram considerados 24 aplicações de previsão, sendo cada aplicação a combinação de períodos da tabela 1 com configurações da tabela 2. Além disso, para aumentar a precisão da previsão foram realizados 10 previsões para cada aplicação, totalizando em 240 previsões.

5.3.1 Curvas de Carga Global

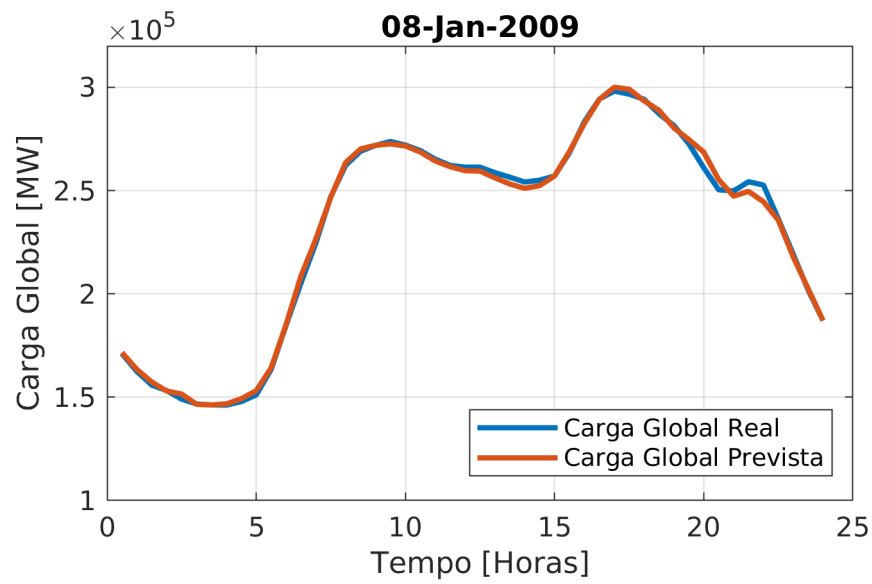
As figuras 18, 19, ..., 41 apresentam as curvas de carga global real e prevista no dia 8 de janeiro de 2009 para cada aplicação.

Figura 18 – Previsão de Carga Elétrica Global - Aplicação 1: Período 1 e Configuração 1



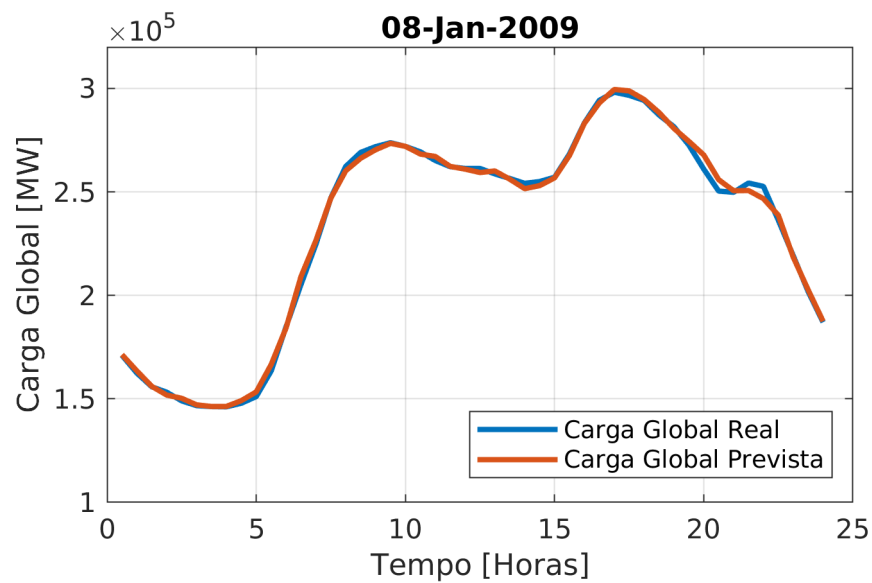
Fonte: Elaborado pelo autor.

Figura 19 – Previsão de Carga Elétrica Global - Aplicação 2: Período 1 e Configuração 2



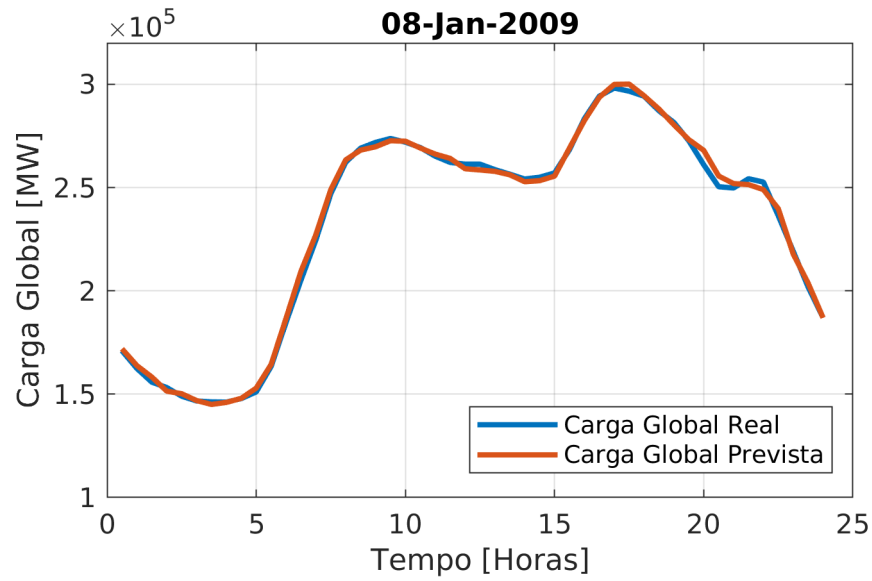
Fonte: Elaborado pelo autor.

Figura 20 – Previsão de Carga Elétrica Global - Aplicação 3: Período 1 e Configuração 3



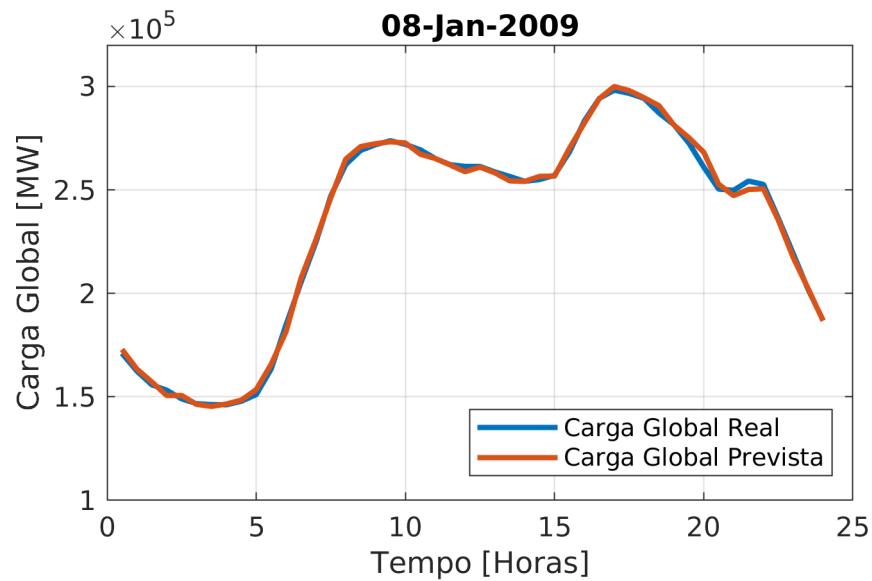
Fonte: Elaborado pelo autor.

Figura 21 – Previsão de Carga Elétrica Global - Aplicação 4: Período 1 e Configuração 4



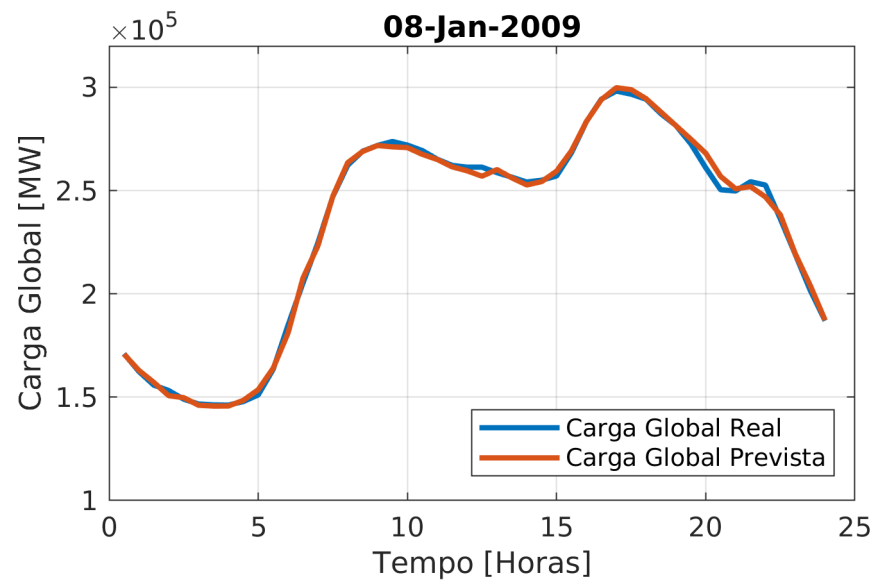
Fonte: Elaborado pelo autor.

Figura 22 – Previsão de Carga Elétrica Global - Aplicação 5: Período 1 e Configuração 5



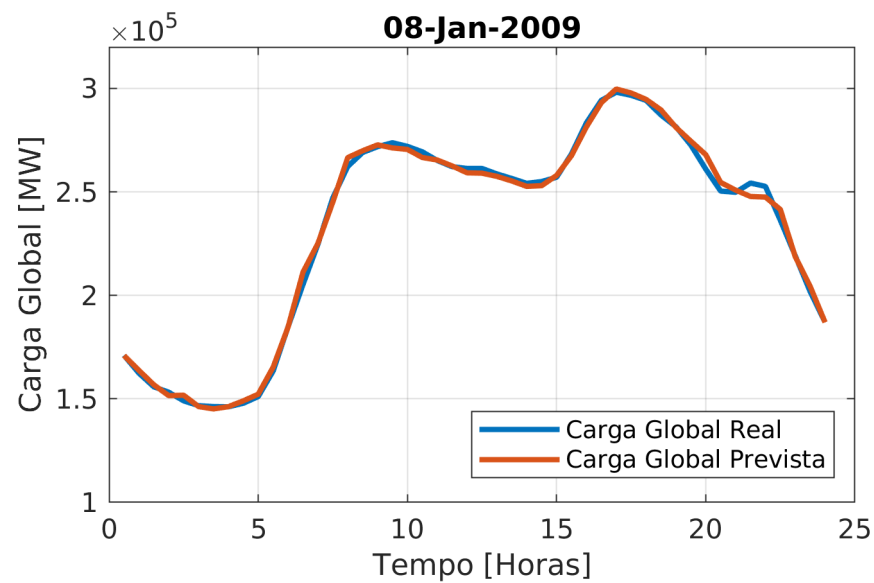
Fonte: Elaborado pelo autor.

Figura 23 – Previsão de Carga Elétrica Global - Aplicação 6: Período 1 e Configuração 6



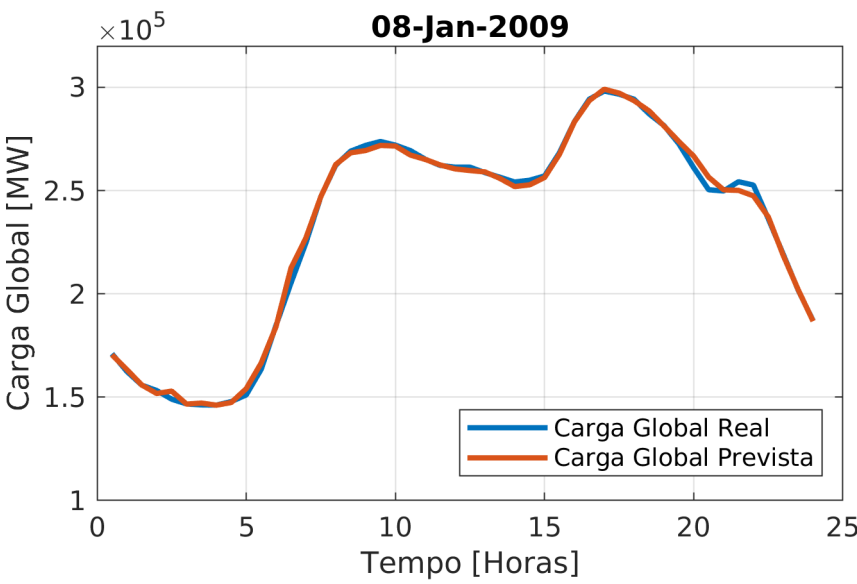
Fonte: Elaborado pelo autor.

Figura 24 – Previsão de Carga Elétrica Global - Aplicação 7: Período 1 e Configuração 7



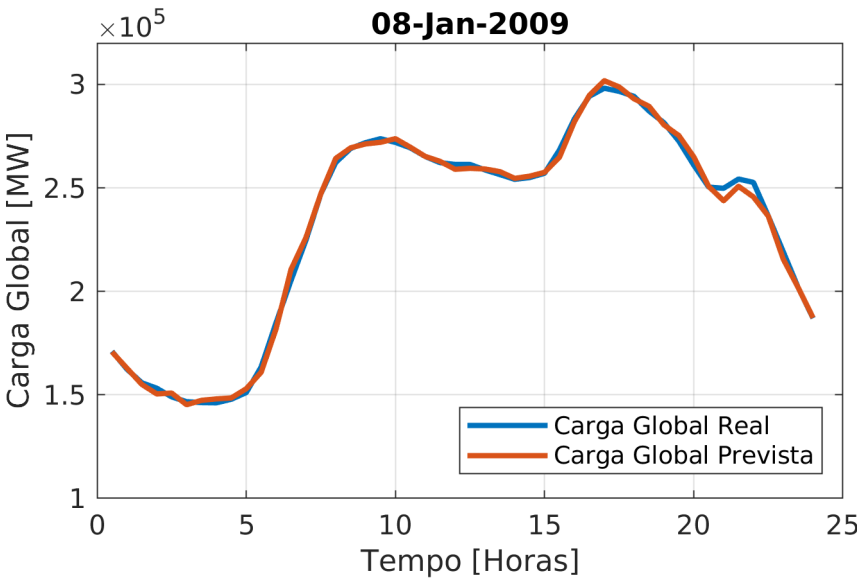
Fonte: Elaborado pelo autor.

Figura 25 – Previsão de Carga Elétrica Global - Aplicação 8: Período 1 e Configuração 8



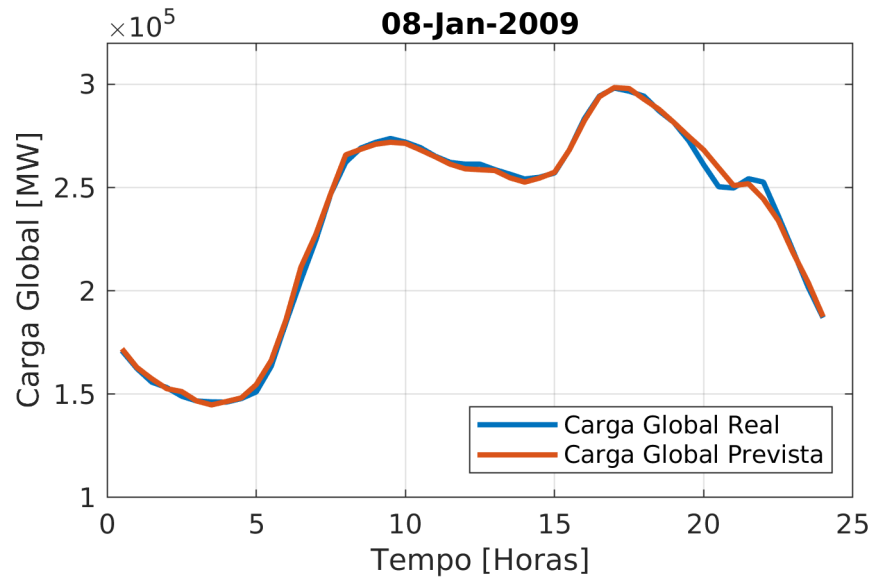
Fonte: Elaborado pelo autor.

Figura 26 – Previsão de Carga Elétrica Global - Aplicação 9: Período 2 e Configuração 1



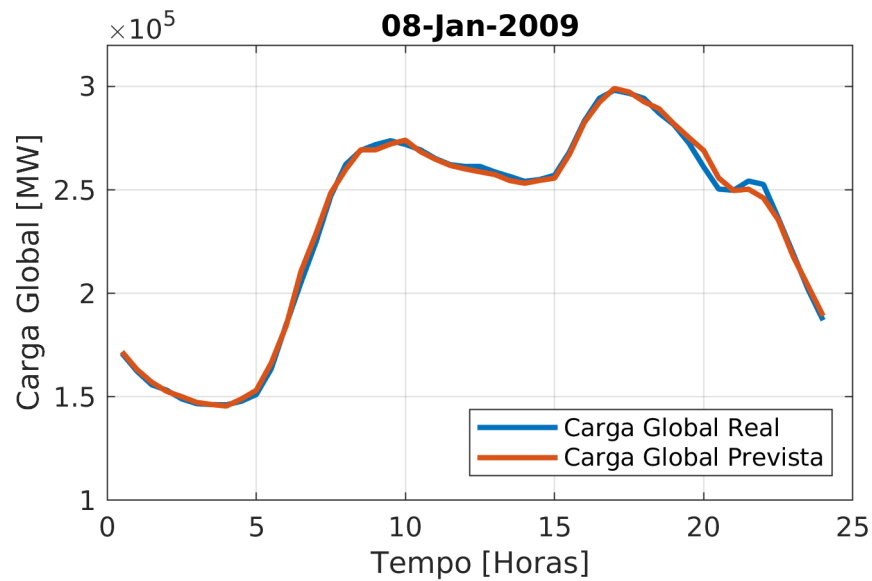
Fonte: Elaborado pelo autor.

Figura 27 – Previsão de Carga Elétrica Global - Aplicação 10: Período 2 e Configuração 2



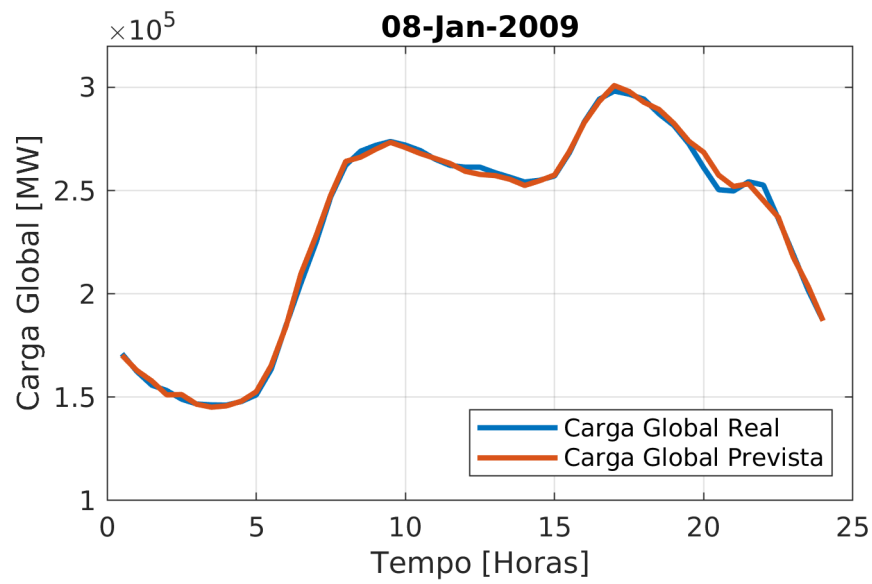
Fonte: Elaborado pelo autor.

Figura 28 – Previsão de Carga Elétrica Global - Aplicação 11: Período 2 e Configuração 3



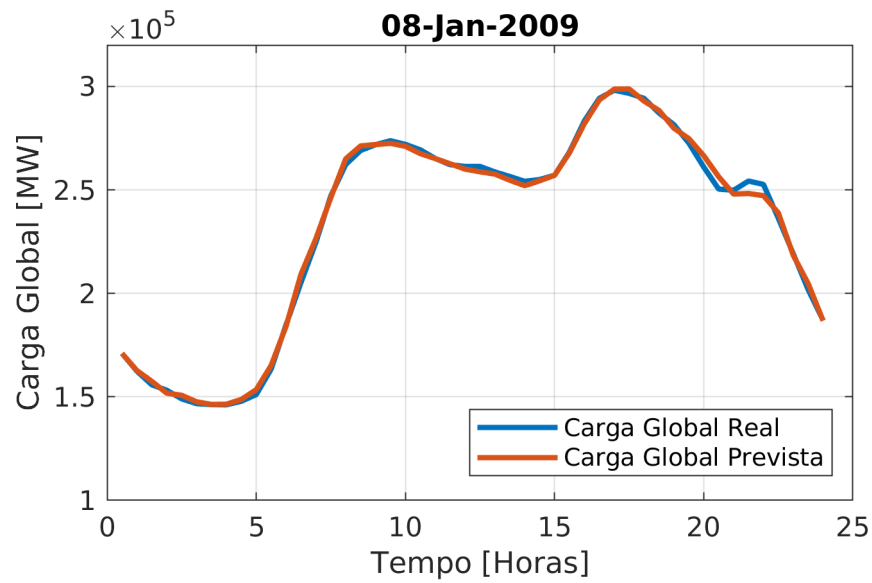
Fonte: Elaborado pelo autor.

Figura 29 – Previsão de Carga Elétrica Global - Aplicação 12: Período 2 e Configuração 4



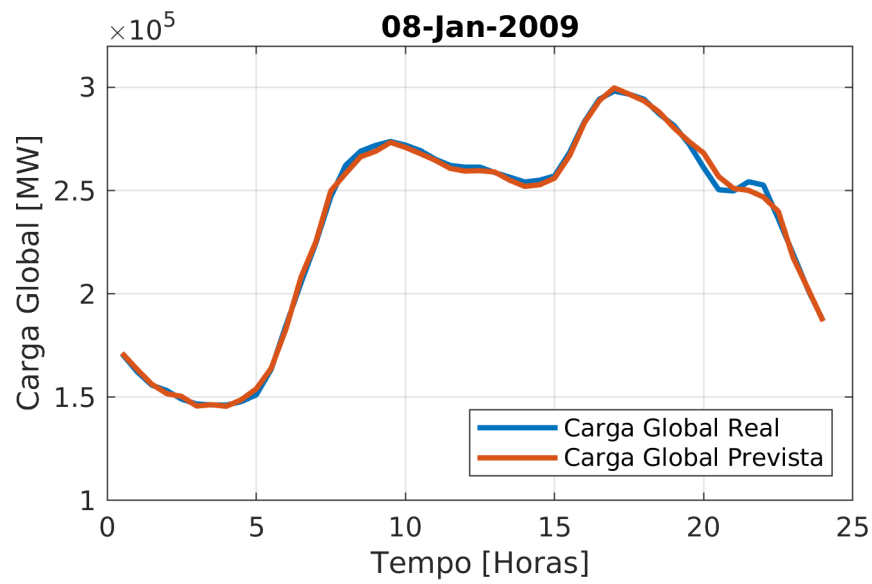
Fonte: Elaborado pelo autor.

Figura 30 – Previsão de Carga Elétrica Global - Aplicação 13: Período 2 e Configuração 5



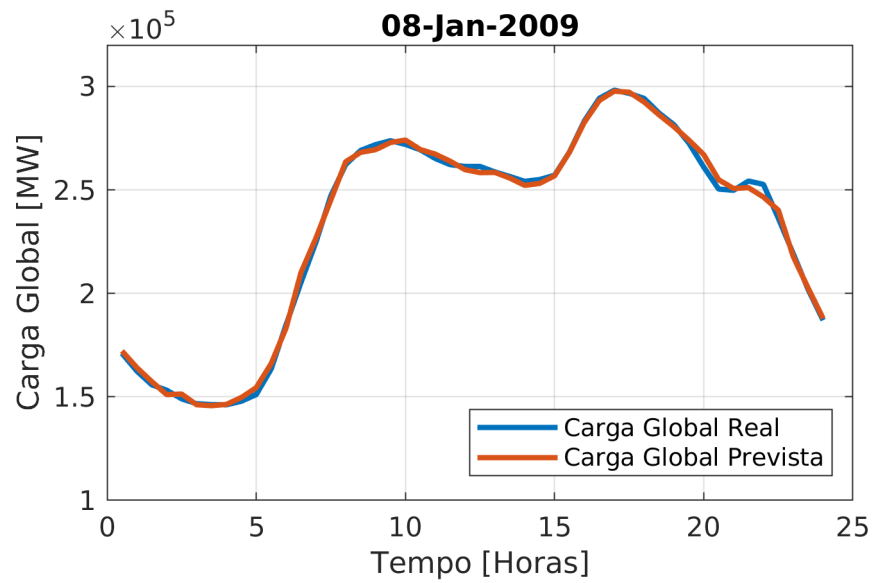
Fonte: Elaborado pelo autor.

Figura 31 – Previsão de Carga Elétrica Global - Aplicação 14: Período 2 e Configuração 6



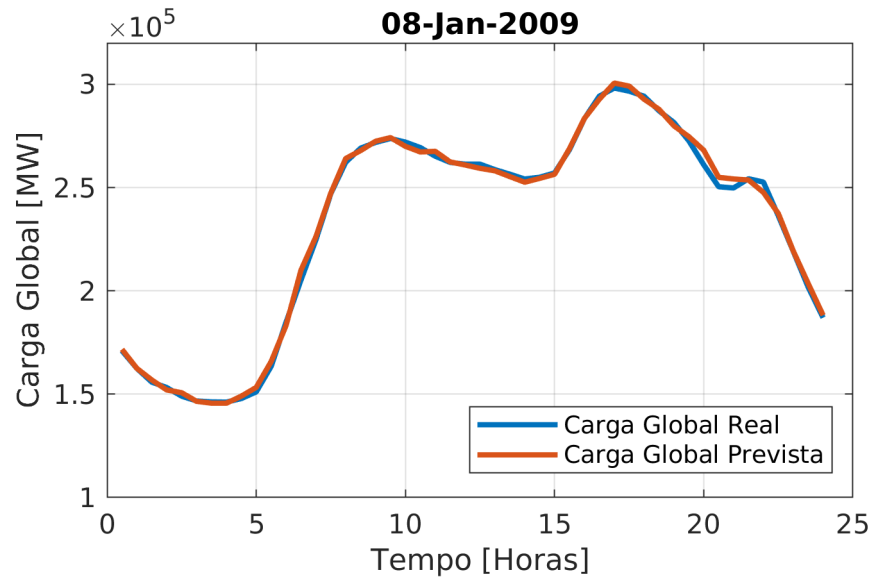
Fonte: Elaborado pelo autor.

Figura 32 – Previsão de Carga Elétrica Global - Aplicação 15: Período 2 e Configuração 7



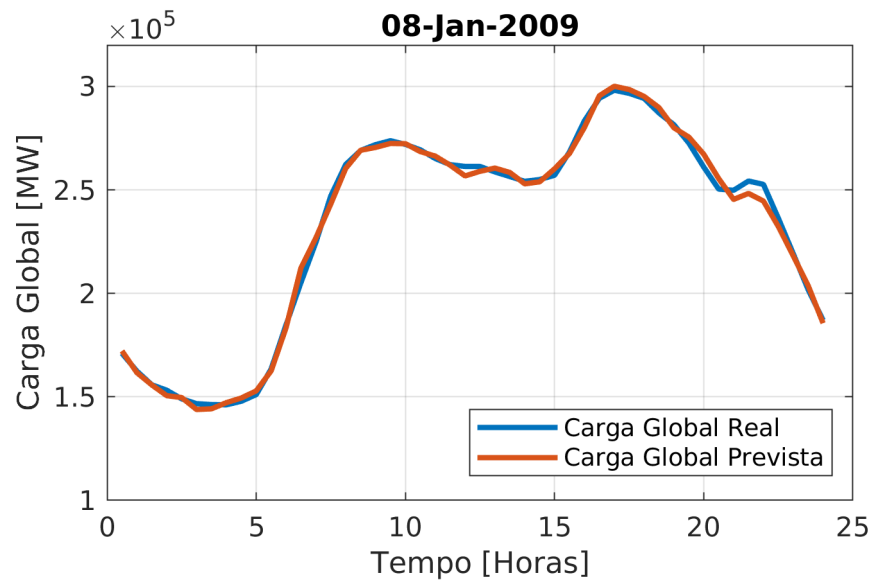
Fonte: Elaborado pelo autor.

Figura 33 – Previsão de Carga Elétrica Global - Aplicação 16: Período 2 e Configuração 8



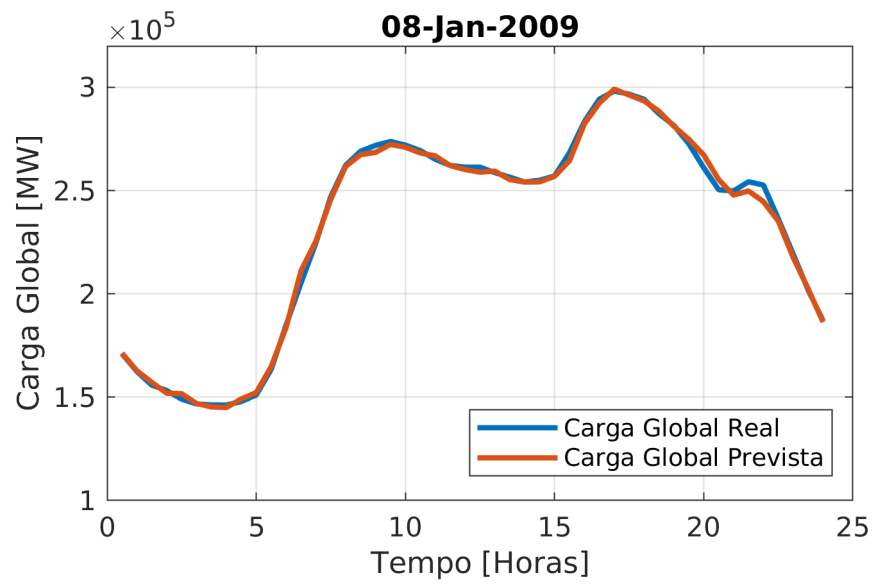
Fonte: Elaborado pelo autor.

Figura 34 – Previsão de Carga Elétrica Global - Aplicação 17: Período 3 e Configuração 1



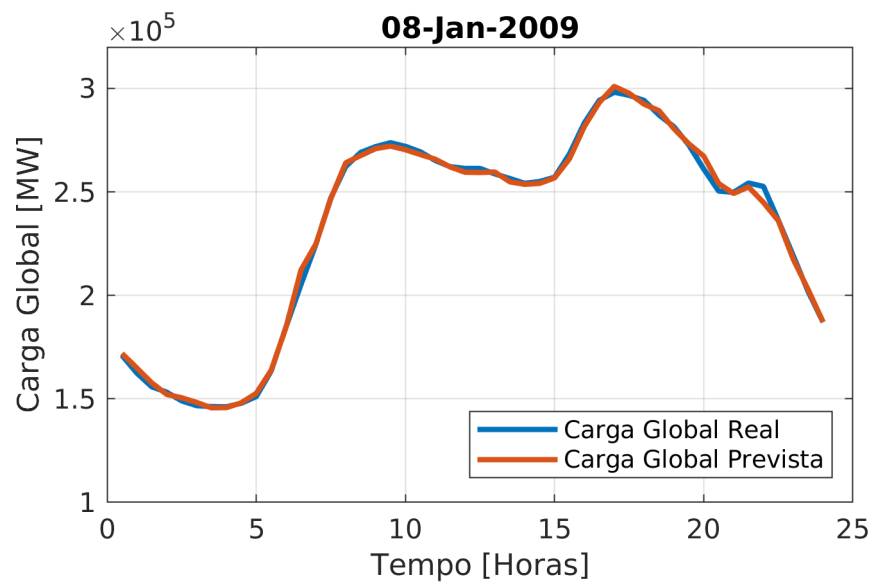
Fonte: Elaborado pelo autor.

Figura 35 – Previsão de Carga Elétrica Global - Aplicação 18: Período 3 e Configuração 2



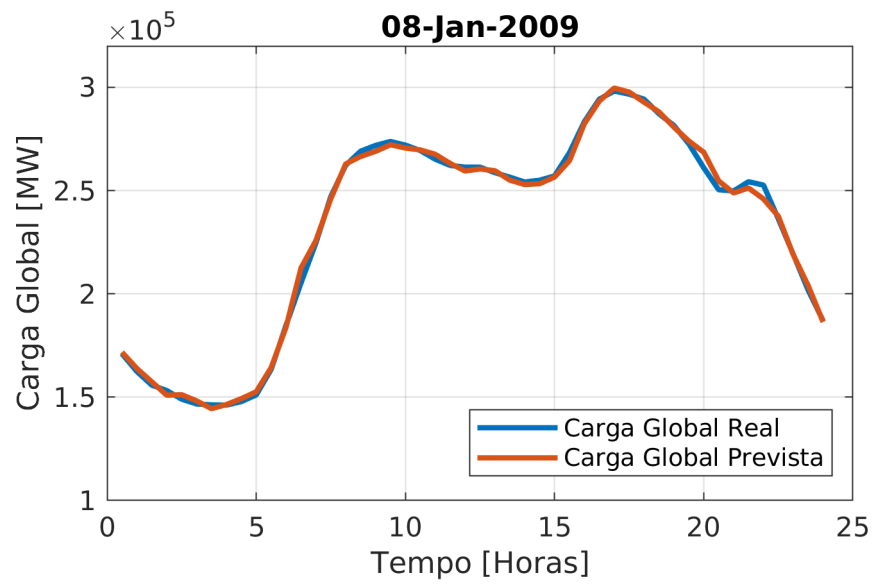
Fonte: Elaborado pelo autor.

Figura 36 – Previsão de Carga Elétrica Global - Aplicação 19: Período 3 e Configuração 3



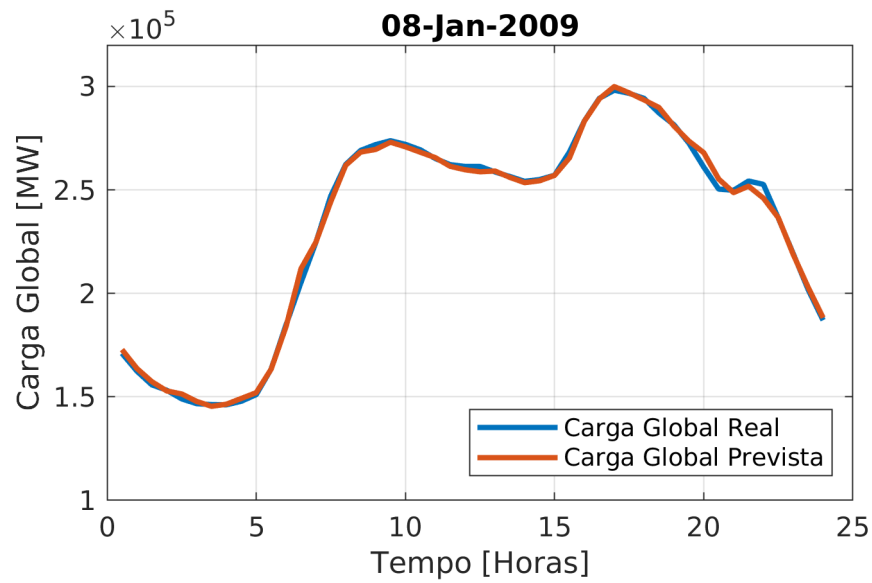
Fonte: Elaborado pelo autor.

Figura 37 – Previsão de Carga Elétrica Global - Aplicação 20: Período 3 e Configuração 4



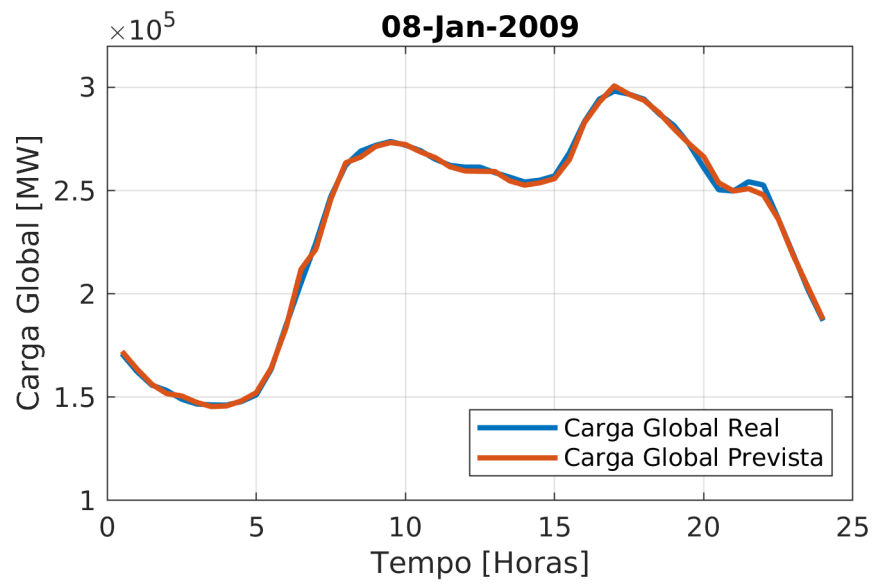
Fonte: Elaborado pelo autor.

Figura 38 – Previsão de Carga Elétrica Global - Aplicação 21: Período 3 e Configuração 5



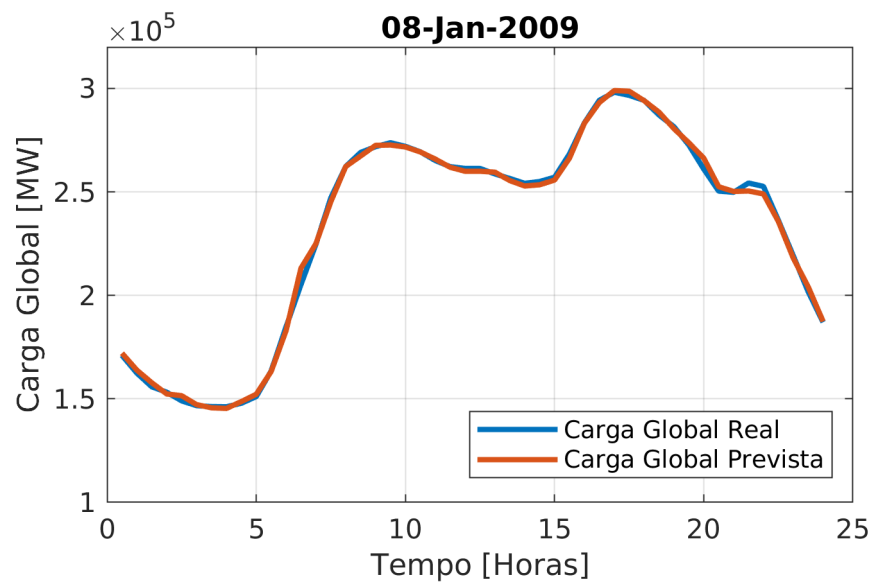
Fonte: Elaborado pelo autor.

Figura 39 – Previsão de Carga Elétrica Global - Aplicação 22: Período 3 e Configuração 6



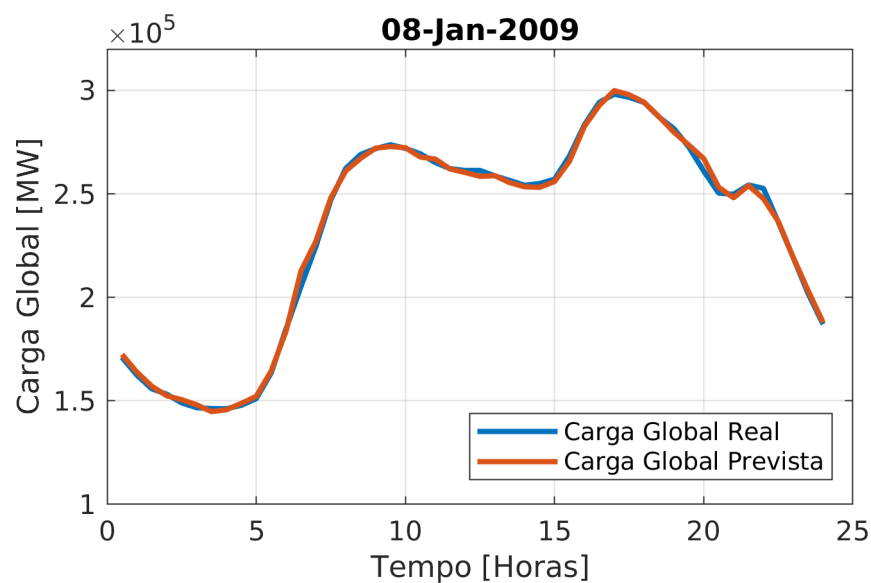
Fonte: Elaborado pelo autor.

Figura 40 – Previsão de Carga Elétrica Global - Aplicação 23: Período 3 e Configuração 7



Fonte: Elaborado pelo autor.

Figura 41 – Previsão de Carga Elétrica Global - Aplicação 24: Período 3 e Configuração 8



Fonte: Elaborado pelo autor.

5.3.2 Análise comparativa de treinamento e desempenho

O comparativo dos resultados obtidos da previsão de carga por cada aplicação estão apresentados nas tabelas 4, 5 e 6, sendo estas relativas, respectivamente, aos períodos 1, 2 e 3. Os resultados apresentados nestas tabelas são a média de 10 previsões realizadas para cada configuração e por este motivo o número de iterações não é inteiro.

Tabela 4 – Resultados Comparativos Médios Para Previsões do Período 1

Período 1				
Configuração	N.º de Iterações (épocas)	Tempo de Treinamento (s)	E _{max} (%)	MAPE (%)
1	165.60	1.18	12.46	0.87
2	129.60	2.90	11.19	0.68
3	82.90	3.06	11.30	0.66
4	81.80	5.63	10.82	0.65
5	136.00	4.25	11.12	0.70
6	69.40	4.17	10.99	0.67
7	89.40	8.08	10.96	0.71
8	50.40	7.53	10.96	0.66

Fonte: Elaborado pelo autor.

Tabela 5 – Resultados Comparativos Médios Para Previsões do Período 2

Período 2				
Configuração	N.º de Iterações (épocas)	Tempo de Treinamento (s)	E _{max} (%)	MAPE (%)
1	129.60	1.74	16.41	1.15
2	159.30	5.68	14.13	0.92
3	184.80	12.17	13.67	0.89
4	116.60	15.00	13.24	0.85
5	147.30	9.27	14.25	0.92
6	88.90	10.54	14.88	0.86
7	105.70	17.30	13.25	0.86
8	80.60	23.47	13.02	0.83

Fonte: Elaborado pelo autor.

Tabela 6 – Resultados Comparativos Médios Para Previsões do Período 3

Período 3				
Configuração	N.º de Iterações (épocas)	Tempo de Treinamento (s)	E _{max} (%)	MAPE (%)
1	168.80	3.14	17.04	1.11
2	242.50	13.49	16.08	0.92
3	140.40	13.03	15.10	0.88
4	164.40	29.45	14.14	0.84
5	137.60	12.02	24.23	2.69
6	136.80	23.08	15.23	0.84
7	145.80	30.99	14.45	0.85
8	105.60	38.11	13.23	0.82

Fonte: Elaborado pelo autor.

6 Conclusão

Neste trabalho, foi desenvolvida uma rotina de previsão de carga elétrica global a curto prazo utilizando uma rede neural feedforward via algoritmo de Levenberg-Marquardt retropropagado, foram executados e analisados diversos testes, com a base de dados históricos de uma companhia do setor elétrico da Nova Zelândia.

Para atribuir pesos corretos, é necessário treinamento. O algoritmo retropropagação é utilizado principalmente para treinar RNAs em classe supervisionada. Os algoritmos de retropropagação dão muita estabilidade durante o processo de treinamento, mas não são rápidos na realização do treinamento e isso se deve à natureza do gradiente descendente. O algoritmo de Levenberg-Marquardt é usado nesta pesquisa para treinar a rede feedforward porque este combina a velocidade do método de Gauss-Newton com a estabilidade do algoritmo de retropropagação de erros durante as etapas de treinamento. Treinar um neurônio em uma RNA ajuda a entender as tendências em uma carga com as entradas específicas fornecidas.

Com as curvas de cargas globais previstas obtidas e apresentadas na subseção 5.3.1 pode-se observar que, em geral, todas tiveram um comportamento semelhante em relação às curvas de cargas globais reais. A maior imprecisão ocorreu entre o intervalo das 19 horas a 23 horas. A medida que a complexidade da RNA aumenta (diferente configuração das camadas) e número de amostras (diferentes períodos), nota-se um aumento da precisão durante o intervalo mencionado, como pode ser observado na curva prevista da figura 41.

O erro percentual médio absoluto (MAPE) foi utilizado, para medir o desempenho da previsão de carga, ao invés do erro médio quadrático (RMSE) pois o MAPE permite a comparação entre diferentes séries o que não é possível com o RMSE. O MAPE mínimo alcançado foi 0.65% durante o período 1 e o máximo alcançado foi 2.69% durante o período 3, o que é uma faixa aceitável conforme discutido na literatura de PCCP. Quanto ao erro máximo percentual (Emax), o mínimo alcançado foi 10.82% durante o período 1 e o máximo alcançado foi 24.23% durante o período 3. No geral os resultados de MAPE e Emax foram bem próximos, sendo a aplicação 21 (Período 3 e Configuração 5) o único "ponto fora da curva". O número de amostras (períodos) não se mostrou ser um fator determinante para o desempenho da RNA em termos de MAPE e Emax. No entanto, quanto a complexidade da rede (tipo de configuração), notou-se que quanto maior a complexidade da rede, mais preciso foi o desempenho, ou seja, menor foram os valores médios de MAPE e Emax.

Em relação ao tempo de treinamento a relação é direta quanto ao número de amostras e a complexidade da rede. Quanto maior o número de neurônios na camada oculta, maior foram os tempos médios de treinamento. O mesmo comportamento pode ser observado em relação ao número de camadas ocultas. As configurações 1 a 4 são RNAs com apenas uma

camada oculta enquanto as configurações 5 a 8 são RNAs com duas camadas ocultas. Embora seja difícil comparar RNAs com camadas distintas, pode-se observar que os tempos médios de treinamento das configurações com apenas 1 camada oculta foram menores que os tempos das configurações com 2 camadas ocultas. Por fim, outro fator determinante para o tempo de treinamento é o número de amostras. Quanto maior o número de amostras maior foram os tempos médios de treinamento observados, sendo este comportamento evidenciado ao comparar os resultados de cada período.

Por fim, quanto ao número de iterações não observou-se um comportamento padrão. O período 1 teve uma relação indireta em relação a complexidade da rede, no entanto esse comportamento não se repetiu para os períodos seguintes. O que pôde-se concluir em relação ao número de iterações é que critério de parada de treinamento determinante foi o erro mínimo de validação e não o número máximo de épocas apresentados na tabela 3.

Referências

- ADYA, M.; COLLOPY, F. How effective are neural networks at forecasting and prediction? a review and evaluation. *Journal of forecasting.*, v. 17, n. 4, p. 481–495, 1998.
- AL-KANDARI, A. M.; SOLIMAN, S. A.; EL-HAWARY, M. E. Fuzzy short-term electric load forecasting. *Electrical Power & Energy Systems*, Guildford, v. 26, n. 2, p. 111–122, 2004.
- AMJADY, N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Transactions on Power Systems*, n. 16, p. 498–505, 2001.
- AMRAL, N.; OZVEREN, C.; KING, D. Short term load forecasting using multiple linear regression. *Universities Power Engineering Conference*, Brighton, p. 1192–1198, 2007.
- ASHOUR, Z. H.; ABU, M. A.; EL-FATTAH, M. A. Modelling and forecasting of electric load using weather information. *The First ICEMP*, Cairo, Egypt, p. 304–316, 1991.
- BENGIO, Y.; DUCHARME, R.; VINCENT, P. Neural probabilistic language model. *Journal of Machine Learning Research.*, v. 3, p. 1137–1155, 2003.
- BOX, G. E.; JENKINS, G. M. *Times Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1976.
- C., G.; A., K. Learning task-dependent distributed structure- representations by backpropagation through structure. *Proceedings of the IEEE International Conference on Neural Networks.*, Washington, p. 347–352, 1996.
- CDS. *Centralized Dataset, New Zealand Electricity Commission for the load dataset*. 2010.
- CHEN, H.; CANIZARES, C.; SINGH, A. Ann-based short-term load forecasting in electricity markets. *Proceedings of the IEEE Power Engineering Society Winter Meeting.*, v. 2, p. 411–415, 2001.
- CHRISTIANSE, W. Short term load forecasting using general exponential smoothing. *IEEE Transactions on PAS*, p. 900–910, 1971.
- DESOUKY, A. A.; ELKATEB, M. M. Hybrid adaptive techniques for electric-load forecast using ann and arima. *IEE Proceedings of Generation, Transmission and Distribution*, p. 213–217, 2000.
- EUGENE, A. F.; DORA, G. Load forecasting. *University of New York*, New York, 2004.
- FILHO, M. E. F.; CARVALHO, A. C. P. L. Evolutionary design of mlp neural network architectures. IV SBRN - Simpósio Brasileiro de Redes Neurais, Campos do Jordão - São Paulo, p. 58–65, 1997.
- FRASCONI, P.; GORI, M.; SPERDUTI, A. A general framework for adaptive processing of data structures. *IEEE Transactions on Neural Networks.*, v. 9, p. 768–786, 1998.
- GROSS, G.; GALIANA, F. D. Short term load forecasting. *Proceedings of the Institute of Radio Engineers*, Institute of Electrical and Electronics Engineers Inc, v. 75, n. 12, p. 1558–1573, 1987.

- HAGAN, M.; DEMUTH, H.; BEALE, M. *Neural networks design*. 3. ed. [S.l.]: PWS Publishing Co, 1996.
- HAHN, S. M.; PICKL, S. Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research*, v. 199, p. 902–907, 2009.
- HAYKIN, S. *Neural networks and learning machines*. 3. ed. New York, USA: Pearson Education, 2009.
- HIPPERT, H.; PERDREIRA, C.; SOUZA, R. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Trans. Power Syst.*, v. 16, p. 44–55, 2001.
- HONG, T.; GUI, M.; BARAN, M.; WILLIS, H. Modeling and forecasting hourly electric load by multiple linear regression with interactions. *Power and Energy Soc. General Meeting*, Minneapolis, p. 1–8, 2010.
- HWAN, K. J.; KIM, G. A short-term load forecasting expert system. *Proceedings of The Fifth Russian-Korean International Symposium on Science and Technology*, p. 112–116, 2001.
- KARPATHY, A. *CS231n Convolutional Neural Networks for Visual Recognition*. n.d. Disponível em: <<https://cs231n.github.io/neural-networks-1/>>. Acesso em: 13 jul. 2022.
- KHATOON, S.; IBRAHEEM; SINGH, A. K.; PRITI. Effects of various factors on electric load forecasting: An overview. In: *2014 6th IEEE Power India International Conference (PIICON)*. [S.l.: s.n.], 2014. p. 1–5.
- KIM, K. H.; YOUN, H. A.; KANG, Y. C. Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method. *IEEE Transactions on Power Systems*, n. 15, p. 559–565, 2000.
- KYRIAKIDES, E.; POLYCARPOU, M. Short term electric load forecasting: A tutorial, vol. 35 of studies in computational intelligence. *Springer*, Berlin, Heidelberg, 2007.
- LI, Y. T.; FANG, E. Y. Short-term electrical load forecasting using least squares support vector machines. *International Conference on Power System Technology*, p. 230–233, 2002.
- LOPES, M. L. M. *Desenvolvimento de redes neurais para previsão de cargas elétricas de sistemas de energia elétrica*. 169 f. Tese (Doutorado) — Universidade Estadual Paulista "Julio de Mesquita Filho", Ilha Solteira, 2005.
- MANOHAR, T.; VEERA, V. Load forecasting by a novel technique using ann. *ARPN J. of Eng. And Appl. Sci.*, v. 3, p. 19–25, 2008.
- MATLAB. “*MathWorks Products*”, *MathWorks Inc*. 2018.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, n. 9, p. 127–147, 1943.
- MIRANDA, V.; MONTEIRO, C. Fuzzy inference in spatial load forecasting. *Power Engineering Winter Meeting, IEEE*, v. 2, p. 1063–1068, 2000.
- MOGHARAM, I. R. Analysis and evaluation of five short-term load forecasting techniques. *IEEE Transactions on Power Systems*, New York, v. 4, n. 4, p. 1484–1491, 1989.

- MURTO, P. *Neural network models for short-term load forecasting*. Dissertação (Mestrado) — Department of Engineering Physics and Mathematics, University of Technology, Helsinki, 1998.
- NGUYEN, D.; WIDROW, B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Proceeding of IJCNN.*, v. 3, p. 21–26, 1990.
- O'DONOVAN, T. M. *Short term forecasting: An introduction to the box-jenkins approach*. John Wiley & Sons, New York, 1983.
- PAPALEXOPOULOS, A. D.; HESTERBERG, T. A regression based approach to short term load forecasting. *IEEE Transactions on Power Systems*, n. 5, p. 40–45, 1990.
- PARK, D.; EL-SHARKAWI, M.; MARKS, R.; ATLAS, L.; DAMBORG, M. Electric load forecasting using an artificial neural network. *IEEE Transactions on Power Systems*, v. 6, n. 2, p. 442–449, 1991.
- RAMEZANI, M.; FALAGHI, H.; HAGHIFAM, M. Short-term electric load forecasting using neural networks. *Int. Conf. on Computer as a Tool.*, Belgrade, p. 1525–1528, 2005.
- RUDER, S. *An overview of gradient descent optimization algorithms*. 2016. Disponível em: <<https://ruder.io/optimizing-gradient-descent/>>. Acesso em: 14 jul. 2022.
- SANTOS, P.; MARTINS, A.; PIRES, A. Short-term load forecasting based on ann applied to electrical distribution substations. *Universities Power Engineering Conf.*, Bristol, UK, v. 1, p. 427–432, 2004.
- SCHMIDT C. A. J.; LIMA, M. A. M. A demanda por energia elétrica no brasil. *Revista Brasil Econômico*, Campinas, v. 58, n. 1, p. 68–98, 2004.
- SKARMAN, S.; GEORGIOPOULOUS, M. Short-term electrical load forecasting using a fuzzy artmap neural network. *Proceedings of the SPIE*, p. 181–191, 1998.
- SRINIVASAN, D.; TAN, S. S.; CHANG, C. S.; CHAN, E. K. Practical implementation of a hybrid fuzzy neural network for one-day-ahead load forecasting. *IEEE Proceedings Generation, Transmission And Distribution.*, New York, v. 145, n. 6, p. 687–692, 1998.
- VEHVILÄINEN I; KEPPO, J. Managing electricity market price risk. *European Journal of Operational Research*, Amsterdam, v. 145, n. 1, p. 136–147, 2003.
- VILLALBA, S. A.; BEL, C. A. Hybrid demand model for load estimation and short-term load forecasting in distribution electrical systems. *IEEE Transactions on Power Delivery*, n. 1, p. 764–769, 2000.
- WERBOS, P. J. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Tese (Doutorado) — Harvard University, 1974.
- WERBOS, P. J. Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, v. 1, n. 4, p. 339–356, 1988. ISSN 0893-6080.
- WIDROW, B.; LEHR, M. A. 30 years of adaptive neural networks: Perceptron, madaline, and backpropagation. *Proceedings of the IEEE.*, v. 78, n. 9, p. 1415–1442, 1990.
- YANG, J.; CHENG, H. Application of svm to power system short-term load forecast. *Power System Automation Equipment China*, n. 24, p. 30–32, 2004.