UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

HAROLDO LUIZ MORETTI DO AMARAL

**Desenvolvimento de uma nova metodologia
para previsão do consumo de energia elétrica
de curto prazo utilizando redes neurais
artificiais e decomposição de séries temporais**

SÃO PAULO
2020

HAROLDO LUIZ MORETTI DO AMARAL

**Desenvolvimento de uma nova metodologia
para previsão do consumo de energia elétrica
de curto prazo utilizando redes neurais
artificiais e decomposição de séries temporais**

Versão Corrigida

(Versão original encontra-se na unidade que aloja o Programa de Pós-graduação)

Tese apresentada à Escola Politécnica da Universidade de São Paulo para a obtenção do título de Doutor em Ciências.

Área de Concentração: Sistemas de Potência

Orientador: Prof. Dr. André Nunes de Souza

SÃO PAULO

2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Amaral, Haroldo Luiz Moretti do

Desenvolvimento de uma nova metodologia para previsão do consumo de energia elétrica de curto prazo utilizando redes neurais artificiais e decomposição de séries temporais / H. L. M. Amaral -- versão corr. -- São Paulo, 2020.

113 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Energia e Automação Elétricas.

1.Previsão do consumo de energia elétrica 2.Redes neurais artificiais 3.Decomposição de série temporal 4.Redes inteligentes 5.Eficiência energética I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Energia e Automação Elétricas II.t.

Dedico este trabalho com todo carinho aos meus pais, Haroldo P. Amaral e Cleide Moretti, que sempre estiveram presentes e me apoiaram na concretização desta conquista. Dedico também à querida Prof.^a Dra. Maria Goretti Zago (*in memoriam*) que me apoiou e incentivou em todas as etapas da minha caminhada acadêmica.

Agradecimentos

Primeiramente agradeço a Deus pela saúde, força e perseverança durante caminhada e pela família maravilhosa que me acompanha e apoia em todos os meus passos, incentivando, corrigindo e confortando quando necessário.

Agradeço ao Prof. Dr. André Nunes de Souza pela orientação, conselhos, amizade e profissionalismo durante todos estes anos em que tenho contato com o grupo de pesquisa do Laboratório de Sistemas de Potência e Técnicas Inteligentes - LSISPOTI. Agradeço também ao Prof. Dr. Luiz Lebensztajn que me acolheu durante o início dessa etapa, tornando possível a realização do doutorado. Também agradeço ao Prof. Dr. Pedro da Costa Jr. pelos auxílios no desenvolvimento deste trabalho.

Agradeço a todos os integrantes do LSISPOTI que nos auxiliam no dia a dia, em especial ao amigo Prof. Dr. Danilo Gastaldello. As contribuições realizadas por todos eles em todos estes anos com certeza trouxeram excelentes frutos de conhecimento e amizade. Agradeço também aos meus amigos do Portal Embarcados que a cada dia proporcionam novos conhecimentos na área de sistemas embarcados, além das diversas conversas descontraídas.

Agradeço a Escola Politécnica da USP pela oportunidade de poder integrar este renomado programa de pós graduação, que muito agregou ao meu desenvolvimento acadêmico e pessoal. Agradeço também a CAPES pelo incentivo e apoio financeiro dado para a execução deste trabalho de doutorado, através da concessão de bolsa pelo programa de pós-graduação da USP.

Por fim, agradeço imensamente e especialmente aos meus pais, Haroldo P. Amaral e Cleide Moretti que sempre estão ao meu lado me incentivando. A luta de vocês para que nunca desistisse dos estudos me trouxe até este caminho, sou grato a vocês.

“O desejo profundo da humanidade
pelo conhecimento é justificativa
suficiente para nossa
busca contínua.”
(STEPHEN HAWKING)

Resumo

Os consumidores residenciais estão cada vez mais ganhando acesso a novas tecnologias de medição e operação remota, com a disseminação de redes inteligentes. Essas oportunidades trouxeram melhorias em termos de qualidade de serviço e o estabelecimento de novos níveis de economia financeira. Estudos recentes demonstraram que um aumento significativo na eficiência energética está relacionado às mudanças nos hábitos de consumo, induzidas por novas tecnologias que dão acesso às informações sobre a qualidade e a quantidade da energia consumida. Nesse contexto, a previsão de consumo de energia de curto prazo é uma ferramenta de conscientização baseada no fornecimento de informações úteis prontas para ajudar na busca dos mais eficientes padrões de consumo e economia de energia. Esta tese apresenta o desenvolvimento de uma nova metodologia para a previsão do consumo de energia residencial, utilizando redes neurais artificiais e decomposição de séries temporais. Vários testes foram realizados para obter a melhor configuração para o problema de previsão do consumo de eletricidade. Os testes também foram realizados comparativamente com os métodos combinados disponíveis, permitindo uma avaliação dos resultados obtidos. Os resultados demonstrados por essa metodologia atingiram um baixo erro percentual médio e um erro percentual absoluto médio baixo no horizonte de previsão, que são as métricas mais utilizadas nesse tipo de estudo.

Palavras-chave: Previsão do consumo de energia elétrica. Redes neurais artificiais. Decomposição de série temporal. Redes inteligentes. Eficiência energética.

Abstract

Residential consumers are increasingly gaining access to new measurement technologies and remote operation with the spread of smart grids. These opportunities have brought improvements in the quality of service and the establishment of new levels of financial savings. Recent studies have shown that a significant increase in energy efficiency is related to changes in consumption habits induced by new technologies that provide access to information on the quality and quantity of energy consumed. In this context, short-term energy consumption forecasting is an awareness tool based on providing useful information ready to assist in the pursuit of the most efficient energy consumption and saving patterns. This thesis presents the development of a new methodology for the prediction of residential energy consumption, using artificial networks and time series decomposition. Several tests were performed to obtain the best configuration for the electricity consumption forecasting problem. The tests were also performed compared to the available combined methods, allowing an evaluation of the obtained results. The results demonstrated by this methodology reached a low average percentage error and a low average absolute percentage error in the forecast horizon, which are the most used metrics in this type of study.

Keywords: Forecast of electricity consumption. Artificial neural networks. Time series decomposition. Smart grids. Energy efficiency.

Lista de ilustrações

Figura 1 – Publicações nos últimos 10 anos.....	27
Figura 2 – Evolução das redes elétricas.....	31
Figura 3 – Comparativo entre tarifa convencional e tarifa branca.....	31
Figura 4 – Desagregação da variação do consumo 2013-2050.....	34
Figura 5 – Impacto da eficiência energética no setor residencial.....	36
Figura 6 – Apresentação das informações em uma casa inteligente	39
Figura 7 – Estrutura de um neurônio simples	41
Figura 8 – Perceptron Multicamadas.....	42
Figura 9 – <i>Smart meter</i> utilizado para coleta inicial de dados.....	48
Figura 10 – Trecho de consumo residencial	48
Figura 11 – Trecho em destaque	49
Figura 12 – Teste 1 – 100 amostras de treinamento	50
Figura 13 – Teste 2 – 400 amostras de treinamento	50
Figura 14 – Previsão a partir da série compactada	51
Figura 15 – Fluxograma das etapas de desenvolvimento	54
Figura 16 – Diagrama do protótipo	56
Figura 17 – <i>Flag</i> indicativo de utilização do <i>buffer</i>	56
Figura 18 – Protótipo para coleta de dados	57
Figura 19 – Curva de carga – Original vs <i>Downsampling</i>	59
Figura 20 – Curva de energia acumulada – Original vs <i>Downsampling</i>	60
Figura 21 – Entrada de dados para a rede neural	61

Figura 22 – Teste 01 – Previsão com a própria série de treinamento	63
Figura 23 – Teste 01 – Previsão com série diferente	63
Figura 24 – Teste 02 – Previsão com a própria série de treinamento	64
Figura 25 – Teste 02 – Previsão utilizando série diferente.....	64
Figura 26 – Teste 03 – Previsão com a própria série de treinamento	65
Figura 27 – Teste 03 – Previsão utilizando série diferente.....	65
Figura 28 – Teste 04 – Previsão com a própria série de treinamento	66
Figura 29 – Teste 04 – Previsão utilizando série diferente.....	66
Figura 30 – Teste utilizando a função sigmoidal na camada de saída.....	67
Figura 31 – Teste 05 – Previsão com a própria série de treinamento	68
Figura 32 – Teste 05 – Previsão utilizando série diferente.....	68
Figura 33 – Teste 06 – Previsão com a própria série de treinamento	69
Figura 34 – Teste 06 – Previsão utilizando série diferente.....	69
Figura 35 – Teste 07 – Previsão com a própria série de treinamento	70
Figura 36 – Teste 07 – Previsão utilizando série diferente.....	70
Figura 37 – Teste 08 – Previsão com a própria série de treinamento	71
Figura 38 – Teste 08 – Previsão utilizando série diferente.....	71
Figura 39 – Teste 09 – Previsão com a própria série de treinamento	73
Figura 40 – Teste 09 – Previsão utilizando série diferente.....	73
Figura 41 – Teste 10 – Previsão com a própria série de treinamento	74
Figura 42 – Teste 10 – Previsão utilizando série diferente.....	74
Figura 43 – Teste 11 – Previsão com a própria série de treinamento	75
Figura 44 – Teste 11 – Previsão utilizando série diferente.....	76
Figura 45 – Teste 12 – Previsão com a própria série de treinamento	76
Figura 46 – Teste 12 – Previsão utilizando série diferente.....	77
Figura 47 – Teste 13 – Previsão com a própria série de treinamento	77

Figura 48 – Teste 13 – Previsão utilizando série diferente.....	78
Figura 49 – Teste 14 – Previsão com a própria série de treinamento	78
Figura 50 – Teste 14 – Previsão utilizando série diferente.....	79
Figura 51 – Representação do modelo gerada pelo Keras	80
Figura 52 – Gráfico com as métricas de perda (<i>loss</i>).....	81
Figura 53 – Teste de previsão com a própria série de treinamento	82
Figura 54 – Teste de previsão utilizando série diferente.....	83
Figura 55 – Entrada de dados para a decomposição temporal	84
Figura 56 – Teste 01 – Previsão com decomposição temporal – Ciclos sequenciais	86
Figura 57 – Teste 02 – Previsão com decomposição temporal – Ciclos sequenciais	86
Figura 58 – Teste 03 – Previsão com decomposição temporal – Ciclos sequenciais	87
Figura 59 – Teste 04 – Previsão com decomposição temporal – Ciclos sequenciais	87
Figura 60 – Entrada de dados para os métodos de previsão.....	91
Figura 61 – Dados utilizados na previsão e resultados das previsões – Teste 01 – Dias sequenciais.....	92
Figura 62 – Destaque para o período da previsão – Teste 01 – Dias sequenciais	92
Figura 63 – Curva real vs. curva da previsão combinada – Teste 01 – Dias sequenciais	93
Figura 64 – Dados utilizados na previsão e resultados das previsões – Teste 02 – Dias sequenciais.....	94
Figura 65 – Destaque para o período da previsão – Teste 02 – Dias sequenciais	94
Figura 66 – Curva real vs. curva da previsão combinada – Teste 02 – Dias sequenciais	95
Figura 67 – Dados utilizados na previsão e resultados das previsões – Teste 01 – Dias iguais.....	96
Figura 68 – Destaque para o período da previsão – Teste 01 – Dias iguais.....	96
Figura 69 – Curva real vs. curva da previsão combinada – Teste 01 – Dias iguais...	97

Figura 70 – Dados utilizados na previsão e resultados das previsões – Teste 02 – Dias iguais.....	98
Figura 71 - Destaque para o período da previsão – Teste 02 – Dias iguais.....	98
Figura 72 – Curva real vs. curva da previsão combinada – Teste 02 – Dias iguais...	99

Lista de tabelas

Tabela I - Consumo por classe (GWh) – Brasil – Ano base 2014	33
Tabela II - Consumo por classe (GWh) - Brasil – Ano base 2017	34
Tabela III – Características da rede NAR utilizada	49
Tabela IV – <i>Datasets</i> encontrados após as pesquisas	53
Tabela V – Configuração da rede e resultados – Teste 01	62
Tabela VI – Configuração da rede e resultados – Teste 02	63
Tabela VII – Configuração da rede e resultados – Teste 03	64
Tabela VIII – Configuração da rede e resultados – Teste 04	65
Tabela IX – Configuração da rede e resultados – Teste 05	68
Tabela X – Configuração da rede e resultados – Teste 06	69
Tabela XI – Configuração da rede e resultados – Teste 07	70
Tabela XII – Configuração da rede e resultados – Teste 08	71
Tabela XIII – Configuração da rede e resultados – Teste 09	72
Tabela XIV – Configuração da rede e resultados – Teste 10	73
Tabela XV – Configuração da rede e resultados – Teste 11	75
Tabela XVI – Configuração da rede e resultados – Teste 12	76
Tabela XVII – Configuração da rede e resultados – Teste 13	77
Tabela XVIII – Configuração da rede e resultados – Teste 14	78
Tabela XIX – Configuração final da rede neural artificial utilizada	80
Tabela XX – Configuração da rede selecionada e resultados	82
Tabela XXI – Configuração da rede selecionada e resultados – Parte 1	85

Tabela XXII – Configuração da rede selecionada e resultados – Parte 2.....	85
--	----

Lista de siglas

AFE - Analog Front End

ANN - Artificial Neural Networks

ARIMA - Autoregressive Integrated Moving Average

ARMA - Autoregressive Moving Average

CMA - Centered Moving Average

CNN - Convolutional Neural Network

DIY - Do It Yourself

DSP - Digital Signal Processing

GP - Genetic Programming

GPU - Graphics Processing Unit

HVAC - Heating, Ventilating and Air Conditioning

IoT - Internet of Things

LSISPOTI - Laboratório de Sistemas de Potência e Técnicas Inteligentes

LSTM - Long Short-Term Memory

MA - Moving Average

MAPE - Mean Absolute Percentage Error

MLP - Multiplayer Perceptron

MPE - Mean Percentage Error

MR - Multiple Regression

MSE - Mean Squared Error

NAR - Nonlinear Autoregressive

RNN - Recurrent Neural Network

SARIMA - Seasonal Autoregressive Integrated Moving Average

SF - Seasonal Factor

SVM - Support Vector Machines

TC - Transformador de Corrente

THD - Total Harmonic Distortion

TP - Transformador de Potencial

Sumário

1	INTRODUÇÃO	20
1.1	Motivação	20
1.2	Objetivo	22
1.3	Ineditismo	22
1.4	Publicações Relacionadas à Tese	23
1.4.1	Trabalhos Publicados	23
1.4.2	Outros Trabalhos e Contribuições	25
1.5	Estrutura da Tese	25
2	REVISÃO BIBLIOGRÁFICA	27
2.1	Estudo de Anterioridade e Estado da Arte	27
2.2	<i>Smart Grids</i> e Eficiência Energética	30
2.3	Previsão do Consumo de Energia Elétrica	37
2.4	Redes Neurais Artificiais	40
2.5	Decomposição Temporal	43
3	DESENVOLVIMENTO DA METODOLOGIA PROPOSTA	47
3.1	<i>Smart Meter</i> Utilizado para Coleta de Dados Iniciais	47
3.2	Testes Utilizando o Matlab	49
3.3	Escolha de uma Nova Plataforma Computacional	52
3.4	Desenvolvimento do Processo de Previsão	54
3.4.1	Desenvolvimento de um Protótipo para Coleta de Dados	54
3.4.2	Tratamento dos Dados	58
3.4.3	Escolha da Topologia da Rede Neural Artificial	60
3.4.3.1	Previsão um Passo à Frente (Wh)	62
3.4.3.2	Previsão um Passo à Frente (W)	67
3.4.3.3	Previsão N Passos à Frente (Wh)	72
3.4.3.4	Previsão N Passos à Frente (W)	75
3.4.3.5	Configuração Final - Previsão N Passos à Frente - W	79

3.4.4	<i>Previsão Utilizando Decomposição Temporal</i>	83
3.4.5	<i>Nova Abordagem Proposta para Previsão do Consumo</i>	88
4	RESULTADOS EXPERIMENTAIS	90
4.1	Teste Utilizando Dias Sequenciais	91
4.2	Teste Utilizando Dias Iguais	95
4.3	Análise da Função ReLU e do Dropout nos Resultados	99
5	CONCLUSÕES	101
5.1	Aspectos Gerais	101
5.2	Desdobramentos da Pesquisa	103
	REFERÊNCIAS BIBLIOGRÁFICAS	105
	GLOSSÁRIO	111

1 INTRODUÇÃO

1.1 Motivação

Os dados do Plano Decenal de Expansão 2027 [1] apontam um crescimento de 2,1% do consumo de energia total até 2027. Este crescimento é o somatório de diversos aspectos como o aumento na renda média das famílias, o aumento no número de domicílios, além das políticas de eficiência energética. Em 2017 a energia elétrica era responsável por 46% do consumo final de energia no setor residencial. Para 2027 projeta-se que a energia elétrica corresponda a 55% do consumo final de energia, um crescimento de aproximadamente 20%. Este crescimento no consumo de eletricidade está relacionado diretamente com a posse média de eletrodomésticos, decorrente da elevação da renda familiar.

O aumento no consumo projetado só não é maior devido às políticas que determinam índices mínimos de eficiência energética, bem como, o programa de etiquetagem que acaba conscientizando e informando o consumidor sobre a eficiência elétrica dos dispositivos. De todas as fontes de energia, a energia elétrica é a que possui maior potencial de conservação no horizonte 2017-2027.

Fica evidente que a ação dos consumidores possui um impacto muito grande sobre o consumo de energia elétrica, e grande parte dessas ações podem ser melhoradas através do binômio informação e educação.

Com o crescimento das *smart grids*, ou redes inteligentes, os consumidores passaram a ter acesso a novas tecnologias que possibilitaram melhorias na qualidade dos serviços recebidos de uma maneira geral. A introdução de novas tecnologias, diretamente ou indiretamente, facilitam alcançar novos patamares de eficiência energética de um modo muito mais natural e intuitivo.

Grande parte da população no Brasil ainda não tem as *smart grids* como uma realidade em seu dia a dia. O processo de medição e tarifação da energia elétrica é massivamente dependente dos medidores eletromecânicos, que por sua vez dependem da ação humana para coleta do registro realizado [2].

Paralelamente pode-se citar a disseminação e crescimento gradual na utilização da geração de energia fotovoltaica. Em geral essas tecnologias são adotadas inicialmente por consumidores entusiastas que aceitam pagar mais para ter acesso antecipado às novidades tecnológicas, também citados como “*early adopters*”. Esses consumidores já passam a ter contato com parte das tecnologias e conceitos relacionados às redes inteligentes, além de possuir um conjunto de informações sobre seu consumo de energia que a maioria dos consumidores residenciais não possuem. Sendo assim, pode-se esperar que esses consumidores tendem a consumir energia de uma maneira mais eficiente.

A dissertação desenvolvida em [3] apresenta uma pesquisa aberta aplicada aos consumidores residenciais. Como resultado revelou informações consideradas como interessantes pelos consumidores residenciais. Dentre elas destaca-se: a) consumo atual, visto que esse dado não é intuitivo de ser obtido em um medidor eletromecânico por um leigo; b) preço atual da energia elétrica, principalmente em mercados que possuem preços dinâmicos ou tarifas variáveis; c) histórico de consumo; e d) previsão de consumo.

Essas informações possibilitam aos consumidores um melhor entendimento sobre seus padrões e hábitos de consumo, ao mesmo tempo em que influenciam na busca por padrões mais eficientes de consumo, mantendo um balanço entre bem estar e custo-benefício.

Como apresentado nos parágrafos anteriores, a informação é um dos alicerces da eficiência energética, ao mesmo tempo que atualmente pode ser considerada como restrita ou escassa para os consumidores residenciais. Ou seja, a falta de informações é um problema a ser solucionado com o objetivo de melhorar os índices de eficiência energética.

Embora diversos artigos tenham sido publicados com o foco na previsão do consumo de energia elétrica, em geral focam na aplicação do método em grandes instalações ou mesmo no consumo da concessionária de energia. Poucos artigos

focam nos consumidores residenciais. Além disso, a abordagem utilizada em diversas publicações não é clara com relação a estrutura utilizada, bem como, se os dados apresentados no horizonte de previsão refletem a previsões uma etapa ou múltiplas etapas a frente, tornando difícil estabelecer comparações [4].

1.2 Objetivo

Essa tese traz à tona a relação entre eficiência energética e a quantidade e qualidade das informações fornecidas aos consumidores residenciais para destacar o uso da previsão do consumo de curto prazo como uma ferramenta auxiliar no processo de gerenciamento do consumo de uma residência.

O objetivo desta tese foi desenvolver uma nova metodologia para previsão do consumo de energia elétrica de curto prazo utilizando redes neurais artificiais e decomposição de séries temporais para obtenção de informações que serão utilizadas pelo consumidor, na busca por melhores padrões de consumo e eficiência energética.

1.3 Ineditismo

O ineditismo da proposta surge de uma nova metodologia alternativa para previsão do consumo de energia elétrica de curto prazo, ou seja, para o dia seguinte. Essa abordagem combina a utilização de redes neurais artificiais (*Artificial Neural Networks* - ANN) com a decomposição de séries temporais visando minimizar o erro percentual médio (*Mean Percentage Error* - MPE) e o erro percentual absoluto médio (*Mean Absolute Percentage Error* - MAPE) ao longo do horizonte de previsão.

Para a avaliação das técnicas utilizadas foi necessária a aquisição de dados reais relativos ao consumo de energia elétrica, sendo assim, desenvolveu-se um protótipo para coleta de dados em longo prazo que suprisse as características necessárias à aplicação. O protótipo desenvolvido possuiu classe de exatidão e precisão comparáveis aos medidores encontrados no mercado, além de incorporar táticas que evitam a perda de dados em situações indesejadas, como uma queda de

conexão com o servidor na nuvem, uma característica presente na maioria dos bancos de dados disponíveis na internet.

Para o desenvolvimento da nova metodologia focou-se na utilização de uma ferramenta de código aberto (*open source*) que vem se popularizando muito nos últimos anos, evitando que a solução fique dependente de uma ferramenta paga como o Matlab, o que poderia no futuro inviabilizar a aplicabilidade dessa abordagem.

Além disso, torna possível implementações futuras fora do ambiente computacional tradicional, permitindo a integração da metodologia proposta em um sistema de gerenciamento de uma residência inteligente (*smart home*).

Os dados resultantes dessa previsão mostram-se de grande interesse para o consumidor residencial, ao passo que fornecem informações e métricas que podem auxiliar no controle e gerenciamento dos custos.

Em geral, os trabalhos relacionados ao tema abordam a previsão de consumo de energia elétrica de maneira genérica, sem um foco específico, ou então, focando no lado do fornecedor de energia. Diferentemente, esta tese foca na previsão de consumo em uma residência, ou seja, no consumidor final.

O desenvolvimento da metodologia utilizando uma abordagem alternativa que combina as previsões geradas por redes neurais artificiais e pelo método de decomposição temporal, utilizando ferramentas de código aberto, além do enfoque no consumidor residencial evidencia o ineditismo desta tese.

1.4 Publicações Relacionadas à Tese

Seguem artigos provenientes de resultados parciais da pesquisa em diversos fóruns.

1.4.1 Trabalhos Publicados

- AMARAL, H. L. M.; MAGINADOR, J. A. G.; AYRES, R. M. J.; SOUZA, A. N.; GASTALDELLO, D. S. Integration of consumption forecasting in Smart Meters and Smart Home Management Systems. In: Simpósio

Brasileiro de Sistemas Elétricos - SBSE 2018, 2018, Niterói. Anais SBSE 2018, 2018.

- AMARAL, H. L. M. do; SOUZA, A. N.; GASTALDELLO, D. S.; PALMA, T. X. da S.; MARANHO, A. da S.; PAPA, J. P. Use of virtual load curves for the training of neural networks for residential electricity consumption forecasting applications. In: IEEE/IAS INDUSCON 2018 - XIII International Conference on Industry Applications, 2018, São Paulo/SP. IEEE/IAS INDUSCON 2018, 2018.
- AYRES, R. M. J.; SOUZA, A. N.; GASTALDELLO, D. S.; AMARAL, H. L. M. do; IKESHOJI, M. A.; SANTANA, G. V. A New Intelligent System Architecture for Energy Saving in Smart Homes. In: IEEE/IAS INDUSCON 2018 - XIII International Conference on Industry Applications, 2018, São Paulo/SP. IEEE/IAS INDUSCON 2018, 2018.
- IKESHOJI, M. A.; SOUZA, A. N.; GASTALDELLO, D. S.; SANTANA, G. V.; AMARAL, H. L. M.; MAGINADOR, J. A. G. Políticas Estratégicas em Eficiência Energética e Contribuições da Qualidade da Energia Elétrica. In: Conferência Brasileira sobre Qualidade da Energia Elétrica - CBQEE, 2017, Curitiba. CBQEE 2017, 2017.
- GASTALDELLO, D. S.; SOUZA, A. N.; AMARAL, H. L. M. do; GOTTI, M. A.; PAPA, J. P.; MAGINADOR, J. A. G. Estudo e análise dos padrões de curvas de carga de consumidores residenciais. In: Congresso Brasileiro de Automática - CBA2016, 2016, Vitória/ES. CBA2016, 2016.
- MANCUZO, E.; SOUZA, A. N.; GASTALDELLO, D. S.; AMARAL, H. L. M. Redes Inteligentes com Foco em Eficiência Energética e Conservação de Energia. In: 12th IEEE/IAS International Conference on Industry Applications, 2016, Curitiba/PR. XII INDUSCON - 2016, 2016.
- FERNANDES, F.; CARREIRO, A.; MORAIS, H.; VALE, Z.; GASTALDELLO, D. S.; do AMARAL, H. L. M.; SOUZA, A. N. Management of Heating, Ventilation and Air Conditioning System for

SHIM Platform. In: ISGT LA 2015 - Innovative Smart Grid Technologies Conference Latinamerica, 2015, Montevideo. ISGT LA 2015, 2015.

- GASTALDELLO, D. S.; do AMARAL, H. L. M.; SOUZA, A. N.; RAMOS, C. C. O. Income Classification of Residential Consumers Through Intelligent Techniques and Load Curves. In: ISGT LA 2015 - Innovative Smart Grid Technologies Conference Latinamerica, 2015, Montevideo. ISGT LA 2015, 2015.
- GASTALDELLO, D. S.; SOUZA, A. N.; AMARAL, H. L. M.; ZAGO, M. G. Técnicas Inteligentes para Classificação de Curvas de Carga Residenciais em Relação à Classe de Consumo e Faixa Salarial. In: CBQEE - Conferência Brasileira sobre Qualidade da Energia Elétrica, 2015, Campina Grande/Paraíba. Anais do Evento, 2015.
- GASTALDELLO, D. S.; SOUZA, A. N.; AMARAL, H. L. M. do; ZAGO, M. G. Classification of Residential Load Curves in Consumption Classes Using Intelligent Techniques. In: CLAGTEE - Congresso Latino-Americano de Geração e Transmissão de Energia Elétrica, 2015, São José dos Campos/SP. Anais do Evento, 2015.

1.4.2 Outros Trabalhos e Contribuições

- Biblioteca para processamento de sinais (DSP) em sistemas embarcados, inicialmente desenvolvida para ser aplicada no protótipo para coleta de dados. Disponível publicamente em:
 - https://github.com/agaelema/DSP_and_Math
- Contribuições para a biblioteca do AFE ATM90E26 utilizado no protótipo para coleta de dados. Dentre elas, melhorias na biblioteca e melhorias na planilha para configuração dos registradores. Disponível publicamente em:
 - https://github.com/agaelema/ATM90E26_Arduino

1.5 Estrutura da Tese

Esta tese está estruturada em cinco capítulos:

- Capítulo 2 – REVISÃO BIBLIOGRÁFICA. Apresenta um estudo de anterioridade contendo a cronologia dos estudos relacionados com a previsão do consumo de energia elétrica e séries temporais até o presente momento. Também são apresentados conceitos importantes relacionados ao tema, como às redes inteligentes, os medidores inteligentes e a eficiência energética;
- Capítulo 3 – DESENVOLVIMENTO DA METODOLOGIA PROPOSTA. Neste capítulo será discutida a metodologia utilizada no desenvolvimento da nova abordagem proposta para previsão do consumo de energia elétrica de curto prazo. Inicialmente é realizada uma introdução sobre o desenvolvimento contendo resultados de alguns testes iniciais utilizando a ferramenta *Matlab* que levaram ao desenvolvimento dessa tese. Além disso é discutida a escolha da nova plataforma que acabou substituindo o uso do *Matlab*. Também é abordado o desenvolvimento do protótipo para coleta dos dados de longo prazo utilizados nesta tese. São apresentados testes com as duas técnicas utilizadas que ajudaram a balizar na seleção das suas características e configurações. Ao final é apresentada a nova abordagem e suas características;
- Capítulo 4 – RESULTADOS EXPERIMENTAIS. O capítulo apresenta os resultados dos testes realizados com a nova abordagem proposta. Todos os testes foram realizados através de dados reais permitindo verificar e validar o comportamento do sistema em uma situação real; e
- Capítulo 5 – CONCLUSÕES. Neste capítulo são apresentadas as conclusões finais e um levantamento de possibilidades de desdobramentos futuros da pesquisa.

2 REVISÃO BIBLIOGRÁFICA

2.1 Estudo de Anterioridade e Estado da Arte

A previsão de séries temporais é aplicada nas mais diversas áreas, desde a previsão climática [5], passando pelo setor financeiro [6], chegando até o setor de vendas e controle de estoque [7].

Não diferente, a previsão de séries temporais também é empregada no setor elétrico onde o objeto de estudo geralmente está relacionado a previsão de consumo em curto, médio ou longo prazo, dependendo do objetivo. Por exemplo, com a chegada das redes inteligentes e a disseminação da geração fotovoltaica de energia elétrica, a previsão de séries temporais pode ser utilizada para prever a irradiação solar [8].

Realizou-se uma busca através do portal *Web of Science* para verificar a quantidade de publicações relacionadas com este tema nos últimos 10 anos. O termo buscado foi “previsão de carga” (*load forecasting*), e depois realizou-se uma busca um pouco mais restritiva utilizando a combinação de termos “previsão de carga” e “redes neurais” (*neural network*). O resultado é apresentado na Figura 1.

Figura 1 – Publicações nos últimos 10 anos
Fonte: *Web of Science* – Pesquisa por termos

Observa-se que existe um crescimento quase que constante do tema no período, sendo que aproximadamente metade das pesquisas estão de alguma forma, relacionadas com a utilização de redes neurais artificiais, apesar deste ser apenas um dos métodos possíveis para solução do problema.

O trabalho desenvolvido por [9] apresenta uma abordagem para previsão do consumo mensal de energia aplicando o método auto-regressivo Integrado de Média Móvel (*Autoregressive Integrated Moving Average* - ARIMA) para técnica clássica de decomposição de séries temporais.

Nesse artigo, o método da decomposição multiplicativa é utilizado para decompor a série temporal em uma componente relacionada à “tendência-ciclo” e outra relacionada à sazonalidade. A componente de “tendência-ciclo” é utilizada para, através do método ARIMA agora denominado SARIMA, encontrar o melhor modelo que descreve os dados. Os métodos são novamente combinados para realizar a previsão da série temporal.

É interessante notar que o comportamento de uma série temporal relacionada ao consumo mensal de uma distribuidora de energia tende a ter um comportamento mais bem definido, tornando o método uma boa alternativa.

Na publicação [10] é apresentada uma revisão com diversos métodos utilizados para previsão do consumo de energia em edifícios, uma área um pouco mais ampla, envolvendo não somente o consumo da energia elétrica em si, mas também o consumo relacionado ao processo de aquecimento e refrigeração da edificação.

São apresentadas técnicas que envolvem modelagem física da construção, levando em consideração diversos aspectos construtivos, geográficos e climáticos para prever o consumo associado a um sistema de aquecimento e refrigeração (*Heating, Ventilating and Air Conditioning* – HVAC). Também são apresentados métodos estatísticos como o ARMA e ARIMA. As ANN's possuem um grande destaque dentro do trabalho que relaciona diversos outros artigos que utilizam ANN's para previsão de consumo, temperatura e mesmo otimização dos controles relacionados ao sistema de HVAC. Ao final é apresentado um comparativo entre os métodos relacionando complexidade, facilidade de uso, tempo de execução, tipo de dado de entrada e exatidão do método. Cada um dos métodos possui seus prós e

contras, tornando difícil elencar um único método para solução de todos os problemas. Mesmo assim, apesar da complexidade associada as ANN's, o tempo de execução durante o processo de inferência costuma ser bastante baixo, com uma exatidão bastante elevada, contribuindo para sua aplicação na maioria dos casos.

O artigo desenvolvido em [11] aborda a utilização de ANN's para previsão do consumo de energia em uma construção residencial. Os autores afirmam que ainda não existe uma solução ótima e robusta para o problema, mesmo com os significativos avanços nas áreas de computação e simulação. É proposta uma nova abordagem para o processo de treinamento da rede neural, que então é comparada ao tradicional algoritmo *Levenberg-Marquardt*. O resultado de ambas as abordagens é bastante semelhante segundo os autores, não sendo possível elencar um vencedor em função da quantidade de dados disponível para treinamento e teste.

O tamanho da série temporal disponível para o processo de treinamento é um fator bastante importante nas abordagens relacionadas às redes neurais artificiais. Em geral, o volume de dados necessário é grande, o que pode dificultar os estudos devido às longas séries temporais necessárias. A série deve ser longa o suficiente para fornecer os dados utilizados no processo de treino e testes utilizando a técnica de validação cruzada, na qual, os dados de testes não foram aplicados no treinamento.

O trabalho desenvolvido em [12] apresenta um sistema inteligente para previsão do consumo de energia utilizando redes neurais do tipo Perceptron Multicamadas (*Multiplayer Perceptron* – MLP) e *Long Short Term Memory* (LSTM), bastante utilizadas em problemas envolvendo séries temporais já que a estrutura interna deste tipo de rede consegue lidar melhor com a relação temporal dos dados, o que tende a melhorar o resultado da previsão dependendo da estrutura utilizada. Este tipo de rede neural tem se tornado bastante comum em aplicações relacionadas como *deep-learning*, ou, de aprendizado profundo. O estudo compara o resultado das previsões realizadas com ANN's MLP e LSTM com métodos utilizando regressão linear e Máquinas de Vetores de Suporte (*Support Vector Machines* – SVM), ao final, concluindo que é possível a previsão através dos métodos apresentados com erro de 1,411%, entretanto, não é especificada qual a métrica para o cálculo do erro apresentado.

Novamente é abordada a questão da quantidade de dados, havendo a citação que quanto maior for a série temporal de treinamento, melhores serão os resultados dos testes.

O artigo desenvolvido em [13] utilizou uma abordagem de *Deep Neural Networks*, redes neurais profundas em uma tradução direta, para previsão horária do consumo para um período de 24 horas. A série temporal é aplicada à múltiplas redes convolucionais (*Convolutional Neural Network* – CNN) em paralelo para extração de características para então passar por uma rede neural recorrente (*Recurrent Neural Network* – RNN). Em paralelo são recebidos dados externos como temperatura, indicação de feriado, hora do dia e dia da semana, que são concatenados e aplicados a uma rede MLP. Os resultados da rede RNN e dos dados externos são concatenados para novamente passarem por mais estágios de redes MLP, gerando ao final a previsão da curva de carga.

No artigo desenvolvido em [14] são comparados 5 métodos para previsão de consumo em um prédio administrativo localizado em Londres. Os métodos comparados são: Regressão Múltipla (*Multiple Regression* – MR), programação genética (*Genetic Programming* – GP), ANN, redes neurais profundas e SVM. Os resultados indicam que as ANN's realizam melhor a previsão de que as outras técnicas relacionadas.

Conforme mencionado no início desta seção, grande parte das abordagens utilizadas para previsão do consumo de energia estão relacionadas com redes neurais artificiais, com uma tendência recente de utilização de métodos baseados em *deep-learning* para solução deste tipo de problema. Os métodos utilizando redes profundas têm demonstrado grande potencial, mas ao mesmo tempo possuem um custo computacional elevado quando comparado a uma rede MLP tradicional com poucos elementos. Em geral, as abordagens utilizando *deep-learning* dependem de computadores com maior capacidade de processamento, em especial, da presença de poderosas placas de vídeo (*Graphics Processing Unit* – GPU) para que o processo de treinamento seja acelerado.

2.2 Smart Grids e Eficiência Energética

As redes elétricas convencionais foram concebidas com o objetivo principal de levar a energia, geralmente gerada em grandes unidades geradoras, até os consumidores finais através das redes de transmissão, distribuição e todo o aparato tecnológico associado [15].

As *smart grids* por sua vez trazem à tona um novo conceito para a rede elétrica, incorporando características como: vias bidirecionais tanto para os canais de comunicação como para a distribuição de energia, conceitos relacionados às áreas de tecnologia da informação e segurança cibernética [2], [15], [16].

A Figura 2 apresenta de forma resumida o novo conceito das redes elétricas inteligentes.

Figura 2 – Evolução das redes elétricas

Fonte: Adaptada de [https://www.energypro-site.com/geracao-distribuida] e [3]

As tarifas variáveis como a tarifa branca possuem um grande apelo dentro de uma rede inteligente, visto sua capacidade de auxiliar no deslocamento de cargas para fora dos horários de pico, minimizando as perdas no sistema. Este deslocamento das cargas ocorre mediante tarifas diferenciadas de acordo com o horário, sendo classificadas como fora de ponta, intermediária e ponta [17], conforme a Figura 3.

Figura 3 – Comparativo entre tarifa convencional e tarifa branca
Fonte: [17]

As redes inteligentes vão ao encontro com outros conceitos relativamente recentes, dentre eles os conceitos de casas inteligentes e prédios inteligentes. Uma casa inteligente busca uma completa integração entre os dispositivos eletrônicos, permitindo um melhor gerenciamento e aproveitamento dos recursos, ao mesmo tempo em que provém maior qualidade de vida aos seus habitantes, conceito semelhante ao das redes inteligentes se for levado em consideração a questão de integração, gerenciamento e controle de recursos.

Dentro de uma casa inteligente pode-se relacionar o conceito de Internet das Coisas (*Internet of Things* – IoT), que compreende uma grande área de estudos e inovações nos últimos anos.

É uma tendência que o IoT seja parte integrante de uma casa inteligente, possibilitando a comunicação entre os mais diversos dispositivos, bem como, com a nuvem, onde estarão concentrados grande parte dos dados e recursos computacionais, envolvendo conceitos de *big data* e aprendizado de máquina (*machine-learning*) [18], [19]. Dentre estes dispositivos pode-se citar o *smart meter* e sistemas auxiliares de gerenciamento de energia.

Antes mesmo das redes inteligentes se tornarem uma realidade, a *internet* das coisas já está se tornando parte do dia-a-dia da população, em especial para os entusiastas em tecnologia, além de um grupo conhecido como *makers*, que está relacionado à pessoas que adotam a filosofia do “faça você mesmo” (*do it yourself* – DIY) para resolver problemas do dia-a-dia. O desenvolvimento de projetos DIY envolvendo conceitos de IoT é um dos temas que vem crescendo bastante nos últimos anos dentro deste grupo de consumidores.

Além disso, também já se encontram comercializadas soluções que envolvem os conceitos de IoT, dentre elas pode-se citar alguns “botões inteligentes” que são instalados na casa dos consumidores e que, quando pressionados, disparam alguma ação, como a entrega de gás ou outro produto.

Outro dispositivo que tem se popularizado são os assistentes virtuais inteligentes como a *Alexa* da *Amazon*, ou o *Google Home* da *Google*. Ambos possuem capacidade interagir naturalmente com o usuário, geralmente através da fala, além de possuírem a capacidade de controlar dispositivos conectados,

permitindo controlar remotamente lâmpadas, dispositivos eletrônicos, entre outros. Um conceito que se relaciona diretamente com o de uma casa inteligente.

A eficiência energética é um tema que se destacou na década de 70 durante os choques no preço do petróleo, quando este passou a ter um custo crescente. Ficava evidente a necessidade de se desempenhar um mesmo serviço com menor gasto energético, gerando menos impactos ao meio ambiente. Nos últimos anos a busca pela eficiência energética ganhou novo destaque em função de fatores ecológicos, como as mudanças climáticas relacionadas ao aquecimento global [20] ou mesmo a dificuldade e os custos associados à construção de grandes unidades geradoras, em especial em um país como o Brasil com a sua elevada dependência da matriz hidroelétrica.

O setor residencial possui grande impacto sobre o consumo total de energia elétrica no Brasil. Em 2014 o setor foi responsável por aproximadamente 28% do consumo total de energia elétrica, perdendo apenas para o setor industrial [21]. Esse montante justifica a aplicação de técnicas e programas que incentivem a busca por melhores padrões de eficiência energética. A Tabela I apresenta o valor absoluto de consumo de 2010 a 2014 e o percentual do ano de 2014.

Tabela I - Consumo por classe (GWh) – Brasil – Ano base 2014

	2010	2011	2012	2013	2014	Variação % (2014/2013)	Participação % (2014)
Brasil	415.667	433.015	448.177	463.134	475.432	2,7	100,0
Residencial	107.215	111.971	117.646	124.908	132.399	6,0	27,8
Industrial	179.478	183.576	183.475	184.685	179.618	-2,7	37,8
Comercial	69.170	73.482	79.226	83.704	89.840	7,3	18,9
Rural	18.906	21.027	22.952	23.455	25.671	9,4	5,4
Poder público	12.817	13.222	14.077	14.653	15.354	4,8	3,2
Iluminação pública	12.051	12.478	12.916	13.512	14.043	3,9	3,0
Serviço público	13.589	13.983	14.525	14.847	15.242	2,7	3,2
Próprio	2.441	3.277	3.360	3.371	3.265	-3,4	0,7

Fonte: Anuário Estatístico de Energia Elétrica 2015 ano base 2014 [21]

Para efeito de comparação, a Tabela II apresenta o percentual de consumo dos principais setores entre os anos 2013 e 2017.

Tabela II - Consumo por classe (GWh) - Brasil – Ano base 2017

	2013	2014	2015	2016	2017	Variação % (2017/2016)	Participação % (2017)
Brasil	463.142	474.823	465.708	461.780	467.161	1,2	100
Residencial	124.908	132.302	131.190	132.872	134.368	1,1	28,8
Industrial	184.685	179.106	169.289	165.314	167.398	1,3	35,8
Comercial	83.704	89.840	90.768	87.873	88.292	0,5	18,9
Rural	23.455	25.671	25.899	27.266	28.136	3,2	6,0
Poder público	14.653	15.355	15.196	15.096	15.052	-0,3	3,2
Iluminação pública	13.512	14.043	15.333	15.035	15.443	2,7	3,3
Serviço público	14.847	15.242	14.730	14.969	15.196	1,5	3,3
Próprio	3.379	3.265	3.304	3.355	3.277	-2,3	0,7

Fonte: Anuário Estatístico de Energia Elétrica 2018 ano base 2017 [22]

O setor industrial é responsável pelo maior percentual de consumo de energia elétrica, entretanto, o nível de eficiência energética associado a este setor tende a ser elevado, visto que desperdícios durante qualquer uma das etapas do processo de produção irá influenciar no valor final de um produto, e consequentemente, na sua competitividade perante outros produtores. Assim, abordagens com foco na eficiência energética através do consumidor residencial podem representar maior economia, visto que o padrão de consumo das residências é elástico, diferente do consumo de uma indústria que tende a ser constante.

Projeta-se que em 2050 o setor residencial brasileiro consuma aproximadamente 212 TWh a mais do que 2013, passando de 124 TWh para 352 TWh, um aumento de cerca de 183,87% [20]. O gráfico da Figura 4 apresenta os principais motivos para esse crescimento.

Figura 4 – Desagregação da variação do consumo 2013-2050

Fonte: Estudos da Demanda de Energia - Demanda da Energia 2050 [20]

O incremento no consumo está relacionado principalmente com o aumento no número de domicílios e crescimento populacional esperado para o período, seguido pela posse e pelo uso de dispositivos eletroeletrônicos, respectivamente.

Contrapondo essa tendência de crescimento, observa-se uma diminuição na potência média dos dispositivos, efeito que pode ser diretamente relacionado a evolução tecnológica e ao aumento da eficiência energética intrínseca. O aumento na eficiência é uma evolução natural com o desenvolvimento dos dispositivos eletrônicos, que tendem a consumir cada dia menos. Um exemplo bastante ilustrativo são as lâmpadas incandescentes que a alguns anos passaram a ser gradativamente substituídas por lâmpadas compactas fluorescentes, e mais recentemente, pelas Lâmpadas LED. Esse ponto é ilustrado no gráfico da Figura 4, ao apresentar a diminuição na potência consumida em uma residência no horizonte 2013-2050.

Geralmente relaciona-se a eficiência energética apenas com a eficiência intrínseca de um dispositivo eletrônico, deixando-se de lado o fator comportamental relacionado ao usuário. Neste ponto, é importante salientar que mesmo um dispositivo eletrônico com elevado nível de eficiência energética pode elevar consideravelmente o consumo de uma residência. Assim, o consumo desnecessário pode ser considerado como um comportamento ineficiente, mesmo que tecnicamente o dispositivo possua eficiência elevada.

Em 2014 o setor residencial brasileiro registrou crescimento de 5,6%, enquanto o setor industrial registrou queda de 2,0% em relação ao ano anterior [23].

Através do aumento nos níveis de eficiência energética evita-se o consumo de uma certa parcela de energia, ao mesmo tempo em que se impacta sobre o planejamento a longo prazo relacionado à expansão do sistema elétrico. Ou seja, a eficiência energética e a redução no consumo associada podem minimizar a necessidade de expansão em curto prazo, dando mais flexibilidade e margem de segurança na operação do sistema [3].

Segundo o Plano Decenal [1], no ano 2027 a eficiência energética pode acarretar em uma economia de cerca de 4% da energia consumida no setor residencial, conforme apresentado na Figura 5 a seguir.

Figura 5 – Impacto da eficiência energética no setor residencial.
Fonte: Ministério de Minas e Energia [1].

Um estudo conduzido por [24] entre 1991 e 2014 analisou a eficiência energética relacionada ao consumo de energia elétrica do setor residencial. Como resultado constatou-se elevada ineficiência no consumo de energia elétrica, enfatizando a importância de políticas públicas e privadas que estimulem o comportamento mais eficiente.

No artigo citado anteriormente, são relacionados os programas de substituição tecnológica. Esses programas foram amplamente empregados pelas concessionárias ao longo dos últimos anos, assim, grande parte das melhorias relacionadas a esse tipo de programa já foram alcançados.

Neste ponto, alguns trabalhos afirmam que as *smart grids* podem elevar consideravelmente os níveis de eficiência energética [25]–[27]. Um dos pontos chave tem relação à quantidade e qualidade dos dados obtidos e disponibilizados aos consumidores. Consumidores bem informados tendem a agir de maneira mais proativa e eficiente, gerando benefícios para o sistema além do próprio consumidor.

Diversos projetos relacionados às redes inteligentes apresentaram resultados expressivos com relação à eficiência energética. Um deles, realizado na Austrália entre 2009 e 2013 alcançou uma redução de consumo entre 4-10%, outro projeto realizado em 2007 na Califórnia (EUA) levou a redução de consumo entre 11-13% [28]. Esses resultados podem ser explicados devido ao elevado nível de informações disponibilizadas aos consumidores, afetando positivamente em sua capacidade de gestão e tomada de decisão diante de suas necessidades e prioridades energéticas.

O *smart meter* desenvolvido em [3] foi instalado em três residências para coleta de dados. Com relação as características das residências, a primeira possuía um morador adulto, a segunda 3 moradores adultos e a terceira com dois moradores adultos e um bebê.

Para estes testes, a instalação do dispositivo seguida pela explicação aos residentes sobre suas possibilidades e disponibilização de dados básicos como consumo acumulado mensal, despertou a atenção dos participantes, que reportaram ao final do período de coleta terem se policiado em virtude das informações adicionais disponibilizados pelo medidor e por saberem que estavam sendo monitorados, mesmo que indiretamente. Complementarmente, o *feedback* mais informal obtido com diversos consumidores sobre este tipo de tecnologia é bastante promissor, em especial quando relacionado ao acesso dos dados registrados.

Há de se destacar que diversos fatores podem impactar quanto ao aproveitamento das informações pelo usuário, dentre elas, sua escolaridade, grau de familiaridade e aceitação de novas tecnologias, além da sua idade.

Além disso, a chegada e popularização das *smart grids* e suas tecnologias abre portas para novos negócios e serviços oferecidos ao consumidor. Novos serviços tendem a surgir em função das necessidades dos consumidores, bem como do mercado, semelhante ao que aconteceu com o surgimento e desenvolvimento dos *smartphones*.

Os dados apresentados corroboram para a importância do setor residencial, evidenciando as perspectivas no aumento de consumo em longo prazo. Em virtude deste aumento considerável, táticas que proporcionem maiores níveis de eficiência energética são de suma importância, viabilizando que a demanda seja atendida pelas concessionárias ao mesmo tempo em que diminuiria os riscos de sobrecarga no sistema ou eventos como o racionamento de energia que ocorreu durante os anos de 2001 e 2002.

2.3 Previsão do Consumo de Energia Elétrica

A previsão de séries temporais é amplamente aplicada nas áreas relacionadas ao setor financeiro, estudo e previsão do clima, bem como nos diversos ramos da engenharia. Dentro do setor elétrico, a previsão de consumo é utilizada principalmente no processo de gestão a longo prazo do sistema elétrico, auxiliando na escolha das melhores estratégias de investimento e expansão do sistema para cada cenário possível [29].

Usualmente a previsão de demanda é aplicada em larga escala, alcançando um erro médio absoluto percentual (MAPE) inferior a 2%, neste caso considerando-se como grande escala o consumo de um país, estado ou cidade. Quando aplicado em média e pequena escala tende a originar erros percentuais mais elevados [30] e esse comportamento pode ser explicado pela variabilidade de consumo de uma ou poucas unidades consumidoras isoladas. Quando se trata de grandes escalas com muitos consumidores, o comportamento médio tende a ser mais controlado.

A possibilidade do consumidor poder verificar antecipadamente qual deverá ser o consumo de energia até o final de um determinado período – um dia, uma semana, um mês – possibilita que este realize o controle das cargas menos prioritárias com o objetivo de enquadrar seu consumo dentro de um limite desejado.

Pode-se relacionar esse tipo de ação como uma forma de educação ativa do consumidor, levando-o a buscar melhorias em seus hábitos de consumo (mesmo que inconscientemente). A partir deste comportamento pode-se esperar um aumento na eficiência.

A integração da previsão de consumo em um medidor inteligente ou mesmo em um sistema auxiliar torna-se um grande atrativo por unificar o processo em um único dispositivo ou sistema integrado. Como sistema auxiliar pode-se considerar, por exemplo, o controlador central de uma casa inteligente que receberá de um medidor inteligente os dados referentes ao consumo de energia, ou que realize as medições por conta própria, para então poder realizar a previsão do consumo que pode ser apresentada ao consumidor através de uma interface gráfica dedicada ou mesmo através de um APP para o *smartphone*. Exemplos de apresentação das informações em uma casa conectada com um sistema de gerenciamento são apresentados em [31], conforme é apresentado na Figura 6.

Figura 6 – Apresentação das informações em uma casa inteligente
Fonte: Monitoramento do consumo de energia elétrica em residências conectadas [31].

O resultado da previsão de consumo pode ser expresso através do consumo acumulado para o próximo período, através da curva de carga ou através da quantia monetária respectiva ao próximo período. Entretanto, o resultado financeiro tende a possuir um maior apelo ao consumidor, influenciando mais no seu comportamento. Assim, a combinação do custo estimado para o próximo período juntamente com a curva de carga prevista, apresenta informações de simples entendimento sobre o comportamento esperado para o período.

Além disso, as possibilidades dessa integração são bastante amplas. Os dados de previsão podem fazer parte de um sistema de gerenciamento das cargas dentro de uma residência, aprimorando sua capacidade de tomada de decisão, como apresentado em [27], [32]. Mesmo em um sistema com menor integração dos recursos ou capacidade de controle direto sobre as cargas, é possível a utilização de um sistema de recomendação que forneça sugestões com relação à utilização das cargas, buscando mais eficiência.

Outro fator interessante deste tipo de sistema tem relação com o volume de dados transmitidos através da rede. Com o processamento dos dados ocorrendo na própria residência, a quantidade de dados trafegando externamente será drasticamente reduzida, algo que se assemelha muito com o conceito de dispositivo de borda (*edge device*) em uma rede de IoT, onde grande parte do processamento é realizado no próprio dispositivo de borda, que por sua vez, envia apenas os dados úteis para a nuvem.

Diversas técnicas podem ser utilizadas para previsão de séries temporais, dentre as mais comuns [15], [30].

- Métodos estatísticos: tradicionalmente relacionados com modelos lineares de previsão, modelos auto-regressivos (AR), modelos auto-regressivos de média móvel (ARMA), modelos auto-regressivos integrados de médias móveis (ARIMA) e decomposição de série temporal.
- Redes neurais artificiais: baseados nos modelos de redes neurais artificiais (ANN), apresentam grande capacidade de generalização, além de resultados confiáveis em sistemas com comportamentos não lineares.
- Lógica Fuzzy: Uma extensão da lógica booleana tradicional que permite resultados parciais - ou verdades parciais - assemelhando-se muito com a tomada de decisão por um ser humano.
- Neuro Fuzzy: Combina a capacidade de aprendizado e generalização de uma rede neural artificial com as inferências Fuzzy realizadas sobre uma série de dados.

2.4 Redes Neurais Artificiais

Redes neurais artificiais são modelos matemáticos que buscam reproduzir o comportamento de um cérebro biológico, possuindo capacidade de aquisição, manutenção e generalização do conhecimento. A estrutura mais básica de uma ANN é o neurônio artificial.

Assim como na estrutura biológica, um neurônio artificial possui n entradas referentes aos estímulos externos. Estes sinais são ponderados através de pesos sinápticos para então serem linearmente combinados. O resultado desta combinação sofre a ação de uma função de ativação que tem como principal característica ser uma função diferenciável [33], conforme a Figura 7.

Figura 7 – Estrutura de um neurônio simples
 Fonte: Apostila de Tópicos Especiais: Aprendizado de Máquina [34]

A função de ativação controla em que nível o neurônio é ativado, bem como, a intensidade do sinal na saída do neurônio. Em geral são utilizadas funções de ativação não lineares, o que se traduz em uma rica capacidade para aproximar funções [35]. Duas das funções de ativação mais comuns são a função sigmoide, ou logística (1) e a função tangente hiperbólica ($\tanh$) (2).

$$g(z) = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (1)$$

$$g(z) = \tanh(z) = \frac{(e^z - e^{-z})}{(e^z + e^{-z})} \quad (2)$$

Onde

$$y = w^T * x \quad (3)$$

O *Perceptron* foi proposto por *McCulloch* e *Pitts* com o objetivo de modelar o neurônio biológico, entretanto, sua capacidade de aprendizado é limitada visto que pode lidar com apenas um hiperplano [34].

Para sobrepor essa limitação, os neurônios são então interconectados formando as redes, tradicionalmente organizadas em camadas, com cada camada possuindo uma determinada quantidade de neurônios.

A camada de entrada (*input layer*) geralmente é representada como um neurônio, apesar de ser apenas um elemento de passagem, transmitindo o dado externo para os elementos da próxima camada.

Internamente a rede existem as chamadas camadas escondidas (*hidden layers*), que são camadas existentes entre a camada de entrada e a camada de saída. A camada de saída (*output layer*) é a camada que externaliza o resultado da rede neural.

A Figura 8 representa uma rede Perceptron multicamadas.

Figura 8 – Perceptron Multicamadas

Fonte: Elaborada pelo autor

Para a rede anterior:

$$a_1^{(2)} = g(w_{01}^{(1)} * x_0 + w_{11}^{(1)} * x_1 + \dots + w_{n1}^{(1)} * x_n) \quad (4)$$

$$a_2^{(2)} = g(w_{02}^{(1)} * x_0 + w_{12}^{(1)} * x_1 + \dots + w_{n2}^{(1)} * x_n) \quad (5)$$

$$h_w(x) = a_1^{(3)} = g(w_{02}^{(2)} * a_0^{(2)} + w_{11}^{(1)} * a_1^{(2)} + w_{21}^{(2)} * a_2^{(2)}) \quad (6)$$

Onde:

$w^{(t)}$ é o valor de w na iteração t

A quantidade de camadas, neurônios e função de ativação utilizadas depende do problema em questão, geralmente sendo obtido através de testes empíricos. A literatura [36] apresenta diversos métodos para se obter o número ideal de neurônios na camada escondida, entretanto, os métodos possuem suas limitações podendo não alcançar o resultado esperado de acordo com o problema em questão.

Para que a rede neural aprenda o comportamento de uma função é necessário que passe pelo processo de treinamento. Existem diversos algoritmos de treinamento, sendo o Gradiente Descendente um dos mais utilizados. Durante o treinamento os dados de entrada e saída são apresentados a rede neural. A saída

da rede é comparada com a saída real, gerando um erro. Este erro é propagado da saída para a entrada, uma camada por vez, e a cada camada os pesos são atualizados de acordo com sua contribuição para o erro na saída. Este processo é conhecido como *Backpropagation*.

O processo se repete para todos os dados de treinamento, concretizando o que é chamado de época (*epoch*). Neste caso, onde dados de entrada e saída são apresentados a rede neural, o processo é conhecido como aprendizado supervisionado.

O poder das redes neurais artificiais vem da sua capacidade de aprender as características dos dados apresentados em sua entrada e como relacioná-las com as variáveis de saída que se deseja prever. Devido à elevada capacidade de generalização e de adaptação, comumente são utilizadas para classificação de padrões, *clustering*, aproximação de funções, otimização, além da previsão ou estimação de valores.

2.5 Decomposição Temporal

Séries temporais são definidas como um conjunto de dados gerados sequencialmente no tempo. Pode-se associar a esta definição o conjunto de dados sequenciais relacionados à potência consumida ou energia consumida.

Modelos temporais assumem que um evento futuro tem relação com eventos passados e por isso podem ser representados através de modelos criados a partir de dados passados [37].

Um dos modelos tradicionalmente utilizados é o Modelo de Decomposição Multiplicativa. A forma geral da decomposição pode ser expressa como em (7).

$$Y_t = f(T_t, S_t, e_t) \quad (7)$$

Onde:

Y_t é o resultado da série no tempo t

T_t é a componente de tendência no tempo t

S_t é a componente de sazonalidade no tempo t

e_t é a componente residual (erro) no tempo t

O modelo multiplicativo assume que a série temporal pode ser descrita como em (8) [9], [37], [38].

$$Y_t = T_t \times S_t \times e_t \quad (8)$$

A componente de tendência indica as mudanças nos dados no longo prazo, enquanto a componente de sazonalidade indica mudanças que dependem de alguma sazonalidade, geralmente baseadas no clima ou época do ano [9], neste caso representado intervalos característicos dentro do período de análise de curto prazo.

O cálculo das componentes relativas à decomposição temporal pode ser dividido em uma sequência de etapas que devem ser realizadas.

Na primeira etapa são calculados os índices sazonais através da aplicação de uma média móvel (*moving average* - MA). A média móvel aplica uma janela aos valores de entrada, sendo esta janela deslocada no tempo. O objetivo da média móvel é suavizar os dados pertencentes a série temporal, removendo qualquer aleatoriedade associada ao mesmo, resultando em uma componente “tendência-ciclo” com a componente sazonal. A fórmula para implementação da média móvel é descrita em (9).

$$MA(k) = \frac{1}{k} \sum_{t=1}^k Y_t \quad (9)$$

Como a quantidade de pontos relacionados a cada iteração da média móvel é um número par de pontos, torna-se necessário uma nova média móvel, entretanto, contendo apenas dois pontos, sendo o ponto atual e o seu seguinte. Este processo é conhecido como média móvel centrada (*centered moving average* – CMA). No caso da utilização de k pontos, sendo k um número ímpar, essa segunda média móvel não é necessária pois o resultado da primeira média móvel já representa um resultado centrado em relação ao total de pontos.

Na próxima etapa calcula-se o fator de sazonalidade (*seasonal factor* – SF), obtido através da relação entre o valor real da série temporal no tempo t e o valor dessazonalizado resultante de média móvel centralizada, conforme em (10).

$$SF_t = \frac{Y_t}{CMA_t} \quad (10)$$

Cada fator de sazonalidade calculado está associado a um ponto da série temporal, que por sua vez está associado a um horário. Através da média dos fatores de sazonalidade referentes ao mesmo horário remove-se qualquer componente residual, originando o índice sazonal (*seasonal index* – S_t) que irá possuir m resultados, referentes a cada horário da janela temporal utilizada na média móvel.

Na próxima etapa são calculados os dados dessazonalizados através da relação entre o valor real da série temporal no tempo t e o índice sazonal, conforme em (11).

$$d = \frac{Y_t}{S_t} \quad (11)$$

A próxima etapa utiliza o método dos mínimos quadrados (*least squares*) para obtenção da regressão linear, conforme em (12), (13) e (14).

$$m = \frac{n \sum(xy) - \sum x \sum y}{n \sum(x^2) - (\sum x)^2} \quad (12)$$

$$b = \frac{\sum y - m \sum x}{n} \quad (13)$$

$$T_t = m \times t + b \quad (14)$$

Onde:

m é o coeficiente angular da reta

b é o coeficiente linear da reta

n é o número de amostras

x é a contagem da amostra (eixo do tempo)

y é o valor da amostra no tempo x

T_t previsão através da regressão no tempo t

Os dados dessazonalizados são utilizados como variável dependente para a regressão linear, com a variável t variando de 1 até a quantidade número de pontos a serem previstos.

A previsão pelo método da decomposição temporal é finalmente obtida através da multiplicação do resultado obtido com a regressão linear pelo índice de sazonalidade, conforme em (15).

$$\text{Forecast} = T_t \times S_t \quad (15)$$

O método de decomposição temporal traz consigo os benefícios de apresentar resultados interessantes com baixo custo computacional e facilidade de implementação do algoritmo. Apesar das diversas etapas envolvidas no processo, os cálculos envolvidos são de baixa complexidade e custo além de necessitarem de pouca memória para implementação.

No próximo capítulo será abordado o desenvolvimento da metodologia proposta.

3 DESENVOLVIMENTO DA METODOLOGIA PROPOSTA

3.1 *Smart Meter* Utilizado para Coleta de Dados Iniciais

O desenvolvimento inicial se baseou no *software* Matlab, utilizando como base os dados coletados com o medidor desenvolvido na dissertação de mestrado de [3]. Esse medidor possuía características técnicas que suprem todas as necessidades com relação à medição, dentre elas, exatidão superior a 1%, tanto para medição da tensão como da corrente, amplo *range* para medição da corrente, processamento de grandezas relacionadas à qualidade da energia como o fator potência (fp) e a taxa de distorção harmônica relativa às formas de onda de tensão e corrente, se assemelhando muito aos dados fornecidos por um analisador de qualidade da energia convencional.

O dispositivo é apresentado na Figura 9 e possui a capacidade de coletar os seguintes dados:

- Data e hora da coleta;
- Consumo acumulado total em Wh e VAh;
- Medição RMS de tensão e corrente;
- Cálculo das potências aparente, ativa, reativa e reativa de distorção harmônica;
- Taxa de distorção harmônica (*Total Harmonic Distortion* - THD) de tensão e corrente; e
- Formas de onda de tensão e corrente.

Figura 9 – *Smart meter* utilizado para coleta inicial de dados
Fonte: Arquivo pessoal

Uma abordagem diferente das tradicionais foi utilizada, adotando-se a curva de energia consumida já que o objetivo final era prever o consumo de energia ao final de um determinado período, por exemplo, ao final de um dia. Além disso, a característica da curva demonstra ser mais linear quando comparada com uma curva de carga, o que em teoria, poderia melhorar os resultados das previsões.

Através da observação de um longo período da curva de energia consumida, seu comportamento aparenta ser “quase” linear, entretanto em alguns pontos pode-se notar não linearidades decorrentes do amplo range dinâmico relacionado às cargas de uma residência.

A Figura 10 apresenta um gráfico de consumo e a Figura 11 apresenta o trecho em destaque com mais detalhes.

Figura 10 – Trecho de consumo residencial
Fonte: Elaborada pelo autor

Figura 11 – Trecho em destaque
Fonte: Elaborada pelo autor

3.2 Testes Utilizando o Matlab

Para o teste utilizou-se o *Neural Time Series toolbox* do *Matlab* com a opção *Nonlinear Autoregressive (NAR) model*, conforme apresentado na referência [39]. Optou-se por esta estrutura por depender apenas dos dados passados de uma série histórica, utilizando-se neste caso os valores referentes ao consumo em Wh e VAh. Empiricamente, encontrou-se a seguinte topologia.

Tabela III – Características da rede NAR utilizada

Configuração	
Entradas	1
Número de delays	5
Camada escondida	5 neurônios (Sigmoidal/tanh)
Saída	1 neurônio (Linear)
Épocas	1000

O primeiro parâmetro avaliado foi o erro quadrático médio, ou *mean squared error (MSE)*. Selecionou-se um pequeno grupo com 100 amostras e outro com 1500 amostras. Para o primeiro caso obteve-se um MSE de $1,0 \times 10^{-2}$ e para o segundo caso $3,49 \times 10^{-3}$, após o treinamento. Apesar de os valores variarem a cada treinamento, demonstraram a esperada diminuição no erro em função do aumento no número de amostras para o treinamento da rede.

Novos grupos com 100, 400 e 1500 amostras, foram separados e utilizados para treinar as redes temporais. Estes dados eram referentes a uma coleta diferente dos dados posteriormente utilizados para a previsão.

Para cada grupo realizou-se o treinamento e posterior teste de previsão. Um ponto importante é que o *offset* inicial das amostras de treinamento e teste eram diferentes, permitindo avaliar o comportamento da rede neste caso. Observou-se que o *offset* acabava interferindo no processo levando a um erro bastante elevado. Para resolver este problema subtraiu-se o *offset* dos dados de teste, simulando como se ambos iniciassem com valores semelhantes no eixo y.

A Figura 12 apresenta o resultado do treinamento para 100 amostras.

Figura 12 – Teste 1 – 100 amostras de treinamento

Fonte: Elaborada pelo autor

A Figura 13 apresenta o resultado para 400 amostras de teste.

Figura 13 – Teste 2 – 400 amostras de treinamento

Fonte: Elaborada pelo autor

Os resultados indicavam uma elevada capacidade de previsão para o modelo de rede utilizada. Aparentemente a rede tinha a capacidade de conseguir prever séries com até duas vezes a duração dos dados de entrada. Melhores resultados

foram obtidos para o grupo de 400 amostras que permitiu a previsão de aproximadamente 1500 dados com baixo erro, conforme observado nos gráficos.

O grupo com 1500 amostras foi reservado para um teste diferente. Os dados sofreram uma compactação em seu volume. Uma a cada seis amostras era separada, o que seria equivalente a salvar os dados a cada 2 minutos ao invés de a cada 20 segundos. Esta compactação influencia diretamente na quantidade de dados e respectiva memória necessária para seu armazenamento, além de impactar no tempo de treinamento.

Estes dados foram utilizados para treinar uma nova rede e os resultados obtidos chamam atenção pela exatidão em função da redução de recursos. Foi possível prever com grande exatidão os 1500 pontos da série utilizando apenas 250 pontos da série compactada. Apesar da compactação o método previu os dados com intervalos inferiores aos das amostras, apresentando maiores erros apenas durante as primeiras previsões, conforme a Figura 14.

Figura 14 – Previsão a partir da série compactada
Fonte: Elaborada pelo autor

Os resultados dos testes preliminares se mostraram bastante promissores. Entretanto, surgiram problemas relacionados à implementação das redes fora do *toolbox*. As tentativas de implementação apresentavam erros semelhantes ao apresentado no *toolbox* apenas para previsão de um ponto à frente, diferente dos múltiplos pontos apresentados no teste interno ao *toolbox*. Mesmo após exaustivas buscas por soluções na *internet*, nenhuma resolução efetiva foi alcançada.

3.3 Escolha de uma Nova Plataforma Computacional

As dificuldades encontradas para implementação da rede neural fora do *toolbox* levou a uma busca por uma plataforma alternativa para continuar o desenvolvimento do sistema.

Através das pesquisas realizadas tomou-se o conhecimento de algumas bibliotecas e *frameworks* focados em aplicações relacionadas ao aprendizado de máquina, dentre elas: *TensorFlow*, uma biblioteca de código aberto para aprendizado de máquina mantida pelo *Google*; *Theano*, uma biblioteca para aprendizado de máquina; *Caffe*, um *framework* com foco em *deep-learning*;

As bibliotecas *TensorFlow* e *Theano* possuem um ponto em comum, ambas são compatíveis com a linguagem *Python* o que se mostrou um ponto positivo, visto que essa linguagem tem se disseminado cada vez mais entre os cientistas de dados e *machine-learning* [40]. Os principais pontos que tem contribuído para sua popularização tem relação com a qualidade das bibliotecas disponíveis, a rápida evolução dessas bibliotecas, além de ser uma plataforma de código aberto. Atualmente *Python* é uma das linguagens mais populares nas áreas relacionadas a ciências de dados e *machine-learning*.

Outro ponto favorável tem relação com a utilização da biblioteca *Keras* para *Python*, uma biblioteca para *deep-learning*, de código aberto, que pode utilizar como biblioteca base (*backend*), tanto a famosa biblioteca *TensorFlow* como a biblioteca *Theano*. O principal benefício da sua utilização está relacionado ao elevado nível de abstração das funções intrínsecas de cada *backend*, agregando versatilidade caso seja necessário realizar algum teste comparativo entre os *backends*.

Os testes desenvolvidos serão detalhados na seção 3.4. Nos testes preliminares utilizando a biblioteca *Keras* com *backend TensorFlow* ficou evidente a necessidade de um volume de dados muito maior que o esperado.

O *smart meter* anteriormente desenvolvido serviu como base de testes e coleta de dados para alguns trabalhos do grupo de pesquisa pertencente ao Laboratório de Sistemas de Potência e Técnicas Inteligentes – LSISPOTI, entretanto, apesar das séries armazenadas possuírem um elevado volume de dados

(muitas amostras individuais), o intervalo temporal entre elas era pequeno. Após o tratamento dos dados para curva de carga padrão com 15 minutos, a quantidade de dados úteis diminuía drasticamente, inviabilizando qualquer treinamento ou teste conclusivo.

Para permitir a execução de testes mais conclusivos iniciou-se uma pesquisa em busca de séries temporais relacionadas ao consumo de energia elétrica. As pesquisas retornaram alguns poucos bancos de dados (*datasets*) relacionados ao consumo de energia elétrica. Dentre os principais pode-se listar:

Tabela IV – *Datasets* encontrados após as pesquisas

Nome do <i>dataset</i>	Características	Link para acesso
UCI - Individual household electric power consumption Data Set	Aproximadamente 4 anos de medição em uma residência. Intervalo de 1 minuto entre as amostragens.	Referência [41]
Smart* Data Set for Sustainability	Dados coletados de 3 casas reais, incluindo dados ambientais. Dados de mais de 400 residências anônimas minuto a minuto.	Referência [42]
REFIT: Electrical Load Measurements	Dados coletados em 20 residências Intervalos de 8 segundos	Referência [43]

O volume de dados fornecido por esses *datasets* é enorme. Apesar do volume de dados, todos eles possuíam algumas características em comum, dentre elas: perdas de dados, ausência de dados relacionados à energia consumida (Wh), além de serem dados com origem diferentes da realidade de uma residência brasileira. Dos pontos listados, a perda de dados e ausência da energia consumida se mostraram limitações a serem consideradas.

A perda de dados em séries temporais longas pode ser considerada um problema comum e recorrente, apesar de indesejado. Em geral os dados são transmitidos da residência para um servidor via *internet*. Na ausência de conexão ou mesmo falta de energia os dados podem ser perdidos dependendo da abordagem utilizada.

Dependendo da aplicação, a falta de dados pode ser contornada através da interpolação dos dados de consumo em Wh, corrigindo pontos ausentes no *dataset*.

Devido à ausência dos dados energia consumida (Wh) e a presença de perda de dados, as séries encontradas não se mostraram úteis.

Apesar dos problemas, as séries encontradas se mostraram extremamente interessantes para projetos e testes envolvendo simulação de consumo e estratégias de controle em uma *smart home* que dependam de dados relacionada à potência instantânea dentro de uma residência.

Devido à necessidade de mais dados, foi desenvolvido um protótipo para coleta de dados, apresentado na seção a seguir.

3.4 Desenvolvimento do Processo de Previsão

A Figura 15 apresenta de forma resumida todo o processo de previsão do consumo de energia e suas principais etapas.

Figura 15 – Fluxograma das etapas de desenvolvimento
Fonte: Elaborada pelo autor

Nos itens a seguir são apresentadas as etapas de desenvolvimento do processo de previsão.

3.4.1 Desenvolvimento de um Protótipo para Coleta de Dados

O medidor desenvolvido em [3] possui características técnicas que suprem todas as necessidades em relação à medição, dentre elas, exatidão superior a 1% tanto para medição da tensão como da corrente, amplo *range* para medição da corrente, processamento de grandezas relacionadas à qualidade da energia como o

fator potência (fp) e a taxa de distorção harmônica relativas às formas de onda de tensão e corrente, se assemelhando muito aos dados fornecidos por um analisador de qualidade da energia. Entretanto, sua capacidade de armazenamento de dados é limitada em função do meio físico utilizado, um *chip* de memória *Flash* conectada à placa do medidor.

A limitação na quantidade de dados úteis disponíveis para os testes levou ao desenvolvimento de um protótipo para coleta de dados por longos períodos, que suprisse as necessidades relacionadas aos testes futuros e contornasse as limitações existentes no outro medidor, em especial com relação ao armazenamento dos dados e capacidade de comunicação.

O dispositivo deveria ter como característica principal a capacidade de armazenar os dados de medição diretamente na nuvem, não ficando dependente de um meio físico de armazenamento. Essa característica se mostrou muito útil a fim de evitar perda de dados no caso de alguma falha inesperada.

Com relação ao algoritmo de medição, eram possíveis duas abordagens: implementação dos algoritmos de medição em um microcontrolador, ou, utilizar um *Analog Front End* (AFE) específico para este tipo de aplicação.

Apesar de uma biblioteca para processamento de sinais (*Digital Signal Processing* - DSP) e rotinas utilizadas em medição ter sido desenvolvida durante o começo deste projeto, a segunda abordagem se mostrou mais efetiva para evitar qualquer problema a longo prazo que pudesse originar perda de dados ou mesmo necessidade de longas paradas para manutenção e atualização.

O novo protótipo de medidor inteligente utilizou como base o *front end* ATM90E26 da *Microchip/Atmel*. Este AFE é compatível com as normas IEC62052-11, IEC62053-21 e IEC62053-23, possui classe de exatidão de 0,1% para medição de energia ativa e 0,2% para medição de energia reativa, conversor analógico digital com resolução de 16-bit e saída no formato pulso, além da comunicação serial SPI.

O AFE está conectado a um microcontrolador ESP8266 da *Espressif*, responsável pelas rotinas de configuração e leitura do AFE, além de realizar a comunicação com a nuvem para armazenamento contínuo dos dados coletados.

Para tornar o sistema mais portátil, facilitando a instalação em qualquer residência, adotou-se a utilização de um transformador de corrente (TC) não

invasivo para medição da corrente e um mini transformador de potencial (TP) para medição da tensão. Um diagrama do protótipo é apresentado na Figura 16.

Figura 16 – Diagrama do protótipo
Fonte: Elaborada pelo autor

O protótipo possui alimentação auxiliar através de um *powerbank*, garantindo o seu funcionamento mesmo durante quedas de energia. Além disso, apresenta um algoritmo de segurança contra perda de dados caso a conexão com a *internet* seja interrompida.

No caso de ausência de conexão os dados são armazenados em um *buffer* com capacidade para aproximadamente 4 horas de dados com armazenamento em intervalos de 1 minuto. Assim que a conexão é reestabelecida os dados armazenados no *buffer* são enviados para a nuvem com um *flag* indicativo da ocorrência, conforme destacado na Figura 17, auxiliando no tratamento dos dados nas próximas etapas.

	A	B	C	D	E	F	G	H
1	Carimbo de data/hora	Voltage (V)	Current (I)	Active Power (W)	P.F.	Energy (kWh)	Freq. (Hz)	Lost Data?
4883	26/04/2019 00:14:23	121,25	2,805	267,001	0,787	236,641456	60,04	
4884	26/04/2019 00:15:22	122,34	2,8	268,001	0,783	236,645884	60,01	
4885	26/04/2019 00:16:45	121,74	2,798	267	0,786	236,651443	60,02	0
4886	26/04/2019 00:17:21	121,43	2,8	267,001	0,785	236,654762	60,05	
4887	26/04/2019 00:18:22	121,13	2,796	267,001	0,789	236,659199	60,02	

Figura 17 – *Flag* indicativo de utilização do *buffer*
Fonte: Elaborada pelo autor

O protótipo desenvolvido para coleta dos dados pode ser visto na Figura 18.

Figura 18 – Protótipo para coleta de dados
Fonte: Elaborada pelo autor

O protótipo coleta as seguintes grandezas: (i) Tensão e Corrente; (ii) Potência ativa (W) – instantânea no momento da amostragem; (iii) Fator potência (fp); (iv) Energia consumida (kWh) - Acumulado geral; e (v) Frequência da rede (Hz).

Os dados coletados são enviados para a nuvem, onde são armazenados para posterior processamento e análises. Como objetivo são longas coletas de dados, a escolha do serviço que os armazena é de suma importância. A abordagem implementada utiliza o serviço de planilhas do *Google*, o *Google Sheets*, como um banco de dados na nuvem. Essa abordagem foi selecionada devido a confiabilidades dos serviços relacionados ao *Google*, bem como ao custo zero de manutenção de um servidor. Desde a instalação do sistema até o presente momento, o sistema se mostrou bastante confiável, apresentando poucos erros ou perdas de dados.

Além disso, o *Google* fornece API's que permitem o acesso, leitura, modificação e controle das planilhas através da linguagem *Python*. Um ponto bastante útil em futuras análises dos dados armazenados.

O medidor foi instalado em uma residência real para coleta dos dados e um ponto primordial tem relação a facilidade de acesso a instalação para verificações do dispositivo, atualização ou manutenções no protótipo. Além disso, a residência escolhida possui cargas comuns a grande maioria dos lares brasileiros com um consumo médio mensal de aproximadamente 200 kWh.

O protótipo teve sua instalação em meados de julho de 2018 e continua registrando dados desde então, contando com algumas paradas pontuais para correções de problemas.

3.4.2 Tratamento dos Dados

O tratamento dos dados é parte essencial do processo, garantindo que não existem inconsistências temporais ou erros nos dados que possam interferir negativamente no sistema proposto.

Os dados relativos ao acumulado da energia consumida (Wh) são processados e transformados em curva de carga (W) para os testes posteriores. Este processo é importante pois os dados referentes à potência que são registrados em cada intervalo de tempo estão relacionados à potência instantânea, não à potência média, não refletindo a real curva de carga.

A aquisição e armazenamento dos dados de longas séries temporais enfrenta alguns desafios relacionados ao espaço de armazenamento, fornecimento contínuo de energia ou mesmo conexão com a nuvem no caso de um dispositivo conectado.

Dispositivos com pouca memória necessitam ter os dados constantemente salvos para garantir que dados não serão perdidos. Uma operação inviável dependendo do local onde está instalado o dispositivo.

No caso do protótipo desenvolvido, os dados são disponibilizados pelo AFE de medição para então serem enviados para a nuvem. Para minimizar o volume de dados transmitidos, o campo “data e hora” é adicionado pelo sistema que recebe os dados do medidor. Os principais problemas associados com a transmissão dos dados tem relação à variação temporal entre as transmissões de dados e às quedas de conexão ou de energia.

O processo de tratamento da série temporal é composto por duas etapas principais. Na primeira etapa o objetivo é corrigir os desvios temporais no momento do armazenamento. As informações referentes à curva de energia acumulada (Wh) e Unix *timestamp* são aplicados a uma interpolação linear. O intervalo dos dados resultantes será igual ao intervalo ideal teórico de armazenamento, por exemplo: como os dados da série são armazenados a cada um minuto, a interpolação é realizada para intervalos de 1 minuto. Como resultado, obtém-se uma série com intervalos das amostras uniformemente distribuídos em toda a sua extensão, facilitando a segunda etapa de tratamento.

Na segunda etapa o objetivo é criar uma curva de carga com 24 pontos relativos à demanda média de cada intervalo. Como a interpolação já foi realizada na primeira etapa, um *downsampling* simples aplicado aos dados é o suficiente para gerar a nova curva de consumo acumulado com pontos igualmente espaçados em 1 hora. O *downsampling* é realizado através da seleção de amostras a cada n períodos. A curva de energia acumulada (Wh) gerada no último processo é então transformada em curva de carga, resultando na curva de carga média com 24 pontos no período.

Através de um processo envolvendo a interpolação e *downsampling* dos dados referentes ao consumo acumulado (Wh) e das informações de *Unix timestamp*, foi possível gerar curvas de carga relacionadas à potência média do período, além de corrigir inconsistências associadas à coleta dos dados. Caso o processo de *downsampling* fosse realizado com os dados referente à potência instantânea, a curva resultante não seria coerente com a curva de carga real, inclusive, ao agregar a curva o resultado da energia consumida pode apresentar elevado erro em relação à curva de energia acumulada.

A curva de carga resultante do processo de interpolação e *downsampling* possui 24 pontos referentes à intervalos de 1 h. A Figura 19 apresenta o comportamento de uma parte da curva de carga original, representada na cor verde, armazenada em intervalos de 1 min, junto com a curva de carga processada para intervalos de 1 h, representada na cor vermelha. Fica evidente a redução na variabilidade associada ao comportamento dinâmico das cargas, ao mesmo tempo que o padrão médio de consumo é mantido.

Figura 19 – Curva de carga – Original vs *Downsampling*
Fonte: Elaborada pelo autor

A Figura 20 apresenta o comportamento da curva original de energia consumida e da resultante após o processo de interpolação e *downsampling*, demonstrando que o padrão da curva não foi afetado, incluindo o consumo acumulado.

Figura 20 – Curva de energia acumulada – Original vs *Downsampling*
Fonte: Elaborada pelo autor

3.4.3 Escolha da Topologia da Rede Neural Artificial

A previsão de séries temporais é um problema de elevada complexidade pois, diferente dos problemas envolvendo redes neurais para classificação e regressão, envolve uma dependência temporal associadas aos dados. Essa dependência temporal deve ser levada em consideração durante a seleção do modelo de rede neural, assim como, no tratamento e separação dos dados utilizados durante o treinamento da rede neural.

Abordagens utilizando redes Perceptron multicamadas dependem de um tratamento externo aplicado aos dados de entrada e saída, enquanto redes do tipo NAR e Recorrente (*Recurrent Neural Network* – RNN) possuem elementos internos em sua estrutura que lidam com as relações temporais dos dados de entrada e saída [44].

A configuração de saída da rede neural também influencia em como os resultados da previsão são apresentados. Em geral, uma rede com um único neurônio de saída possibilita a previsão um passo à frente (*one step ahead*), já uma rede neural com múltiplas saídas permite a previsão de múltiplos passos à frente (*multi step ahead*). Em algumas abordagens uma rede com um único neurônio na

saída pode prever n pontos à frente através de um processo iterativo no qual os resultados na saída realimentam a entrada da rede.

Em geral a escolha da estrutura ou topologia da rede neural é realizado pelo método empírico. Define-se um modelo de base e a partir deste modelo são realizadas alterações nos hiperparâmetros da rede neural, modificando quantidade de neurônios por camada, quantidade de camadas, funções de ativação. Ao final são avaliados os resultados em relação à alguma das métricas associadas, como a perda (*loss*), erro médio quadrático (*mean squared error* – MSE), entre outros.

Vetores de entrada e saída aplicados a rede neural tem como origem a série temporal registrada. Adota-se duas janelas temporais, uma contendo m dados passados e outra contendo n dados futuros, assim, são criados pares de vetores contendo m dados de entrada e n dados de saída utilizados durante o processo de treinamento e teste.

A Figura 21 apresenta um exemplo de janelas contendo 12 dados passados para entrada e 24 dados futuros para a saída.

Segmentos da série temporal											
-12	-11	...	-2	-1	0	1	2	...	...	22	23
Dados de entrada para a ANN					Dados de saída para a ANN – Período de previsão						

Figura 21 – Entrada de dados para a rede neural

Fonte: Elaborada pelo autor

Complementarmente foram adotadas duas abordagens diferentes para o deslocamento da janela temporal. A primeira abordagem, intitulada “*sequential*”, cria vetores que estão unitariamente deslocados em relação ao tempo, ou seja, deslocados por uma amostra em relação ao anterior. Já a segunda abordagem, intitulada de “*reshape*”, cria vetores que estão temporalmente deslocados por n amostras em relação ao anterior, onde n é o número de entradas aplicadas simultaneamente a rede neural.

A primeira abordagem origina mais vetores de entrada, o que significa mais dados de treinamento em relação à segunda abordagem. Em contrapartida, o deslocamento unitário eleva a variabilidade dos padrões de entrada e saída. Através dos testes foram verificados qual o melhor tipo de deslocamento temporal da janela para utilização na solução final.

Entre a abordagem “*sequential*” e a abordagem “*reshape*”, juntamente com as variações de parâmetros realizadas, foram testadas mais de 100 estruturas diferentes de redes neurais *Perceptron* multicamadas.

A seguir serão apresentados os principais resultados dos testes realizados, que ao final, levaram a escolha da estrutura utilizada para a rede neural artificial.

3.4.3.1 Previsão um Passo à Frente (Wh)

Os primeiros testes envolveram a previsão de um passo à frente, semelhante ao obtido com o modelo NAR inicialmente testado no Matlab. Os testes foram realizados tanto com dados referentes ao consumo acumulado de energia como com os dados referentes a curva de carga média.

Um modelo de referência para a rede neural foi definido e a partir deste modelo foram realizadas modificações nos hiperparâmetros da rede neural. A seguir serão apresentados os modelos que alcançaram melhor desempenho, juntamente com suas características. Também são apresentados os resultados da previsão através de dois testes, um com o próprio dado de treinamento e outro utilizando uma série diferente.

Teste 01 – Reshape (Wh)

O primeiro teste utilizou a configuração apresentada na Tabela V. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 22 e Figura 23.

Tabela V – Configuração da rede e resultados – Teste 01

Configuração		Resultados Treinamento	
Entradas	4	Mean absolute error (MAE)	240,335586
Camada escondida	8 neurônios (ReLU)	Mean squared error (MSE)	146137,280005
Saída	1 neurônio (Linear)	Root mean squared error (RMSE)	382,279060
Épocas	2000	R squared (R^2)	0,998876

Figura 22 – Teste 01 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 23 – Teste 01 – Previsão com série diferente
Fonte: Elaborada pelo autor

Teste 02 – Sequential (Wh)

O segundo teste utilizou a configuração apresentada na Tabela VI. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 24 e Figura 25.

Tabela VI – Configuração da rede e resultados – Teste 02

Configuração		Resultados Treinamento	
Entradas	4	<i>Mean absolute error (MAE)</i>	127,246547
Camada escondida	8 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	53233,376478
Saída	1 neurônio (Linear)	<i>Root mean squared error (RMSE)</i>	230,723593
Épocas	2000	<i>R squared (R^2)</i>	0,999587

Figura 24 – Teste 02 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 25 – Teste 02 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 03 – Reshape (Wh)

O terceiro teste utilizou a configuração apresentada na Tabela VII. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 26 e Figura 27.

Tabela VII – Configuração da rede e resultados – Teste 03

Configuração		Resultados Treinamento	
Entradas	4	<i>Mean absolute error (MAE)</i>	0,006177
Camada escondida	100 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	0,000052
Saída	1 neurônio (Linear)	<i>Root mean squared error (RMSE)</i>	0,007233
Épocas	2000	<i>R squared (R^2)</i>	0,999486

Figura 26 – Teste 03 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 27 – Teste 03 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 04 – Sequential (Wh)

O quarto teste utilizou a configuração apresentada na Tabela VIII. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 28 e Figura 29.

Tabela VIII – Configuração da rede e resultados – Teste 04

Configuração		Resultados Treinamento	
Entradas	4	<i>Mean absolute error (MAE)</i>	0,002559
Camada escondida	100 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	0,000024
Saída	1 neurônio (Linear)	<i>Root mean squared error (RMSE)</i>	0,004849
Épocas	2000	<i>R squared (R^2)</i>	0,999767

Figura 28 – Teste 04 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 29 – Teste 04 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Os testes apresentaram os resultados para duas configurações bastante diferentes entre si, especialmente com relação à quantidade de neurônios presentes na camada escondida. Nos testes 01 e 02 foram utilizados 8 neurônios na camada escondida, sendo que as amostras utilizadas no processamento não estavam normalizadas, já nos testes 03 e 04 foram utilizados 100 neurônios na camada escondida, agora utilizando amostras normalizadas.

Nos testes 01 e 02 é possível notar alguns pontos onde o resultado da previsão para o próximo ponto atinge valor inferior ao do ponto anterior. Entretanto, este comportamento não se repete nos testes 3 e 4, onde foram utilizados mais neurônios na camada escondida. Para o problema em questão, o aumento no

número de neurônios possibilitou melhor generalização dos dados para a previsão do próximo ponto.

Outro ponto a ser destacado tem relação com a capacidade das redes testadas prever dados além dos valores utilizados no treinamento. Nos testes iniciais realizados no Matlab ocorreu um comportamento onde a previsão começava a saturar a partir de um certo ponto, ou seja, a previsão não poderia ultrapassar o limite superior (Wh máximo) dos dados de treinamento. Nos modelos propostos esse comportamento não ocorre devido à utilização da função de ativação linear na camada de saída, o que acaba não limitando o range da previsão, ou seja, pode-se prever o consumo além do limite superior dos dados de treinamento.

Para comprovar essa característica, realizou-se um teste com a mesma estrutura, apenas substituindo-se a função de ativação da saída para sigmoidal, e o resultado é apresentado na Figura 30.

Figura 30 – Teste utilizando a função sigmoidal na camada de saída
Fonte: Elaborada pelo autor

Observando a figura anterior, percebe-se que a saturação ocorre próximo ao valor 0,3 e próximo ao valor 0,9. Este comportamento é coerente com os dados utilizados no treinamento que possuíam valores dentro desse *range*. O comportamento obtido é semelhante ao apresentado na Figura 12 (seção 3.2), demonstrando que o fator limitador encontrado anteriormente tinha relação com a função de ativação sigmoidal na camada de saída.

3.4.3.2 Previsão um Passo à Frente (W)

Teste 05 – Reshape (W)

O quinto teste utilizou a configuração apresentada na Tabela IX. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 31 e Figura 32.

Tabela IX – Configuração da rede e resultados – Teste 05

Configuração		Resultados Treinamento	
Entradas	4	<i>Mean absolute error (MAE)</i>	79,587733
Camada escondida	8 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	22888,837414
Saída	1 neurônio (Linear)	<i>Root mean squared error (RMSE)</i>	151,290573
Épocas	2000	<i>R squared (R^2)</i>	0,612520

Figura 31 – Teste 05 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 32 – Teste 05 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 06 – Sequential (W)

O sexto teste utilizou a configuração apresentada na Tabela X. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 33 e Figura 34.

Tabela X – Configuração da rede e resultados – Teste 06

Configuração		Resultados Treinamento	
Entradas	4	Mean absolute error (MAE)	73,801395
Camada escondida	8 neurônios (ReLU)	Mean squared error (MSE)	19024,725462
Saída	1 neurônio (Linear)	Root mean squared error (RMSE)	137,930147
Épocas	2000	R squared (R²)	0,561716

Figura 33 – Teste 06 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 34 – Teste 06 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 07 - Reshape (W)

O sétimo teste utilizou a configuração apresentada na Tabela XI. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 35 e Figura 36.

Tabela XI – Configuração da rede e resultados – Teste 07

Configuração		Resultados Treinamento	
Entradas	4	<i>Mean absolute error (MAE)</i>	0,017377
Camada escondida	100 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	0,000503
Saída	1 neurônio (Linear)	<i>Root mean squared error (RMSE)</i>	0,022424
Épocas	2000	<i>R squared (R^2)</i>	0,985771

Figura 35 – Teste 07 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 36 – Teste 07 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 08 - Sequential (W)

O oitavo teste utilizou a configuração apresentada na Tabela XII. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 37 e Figura 38.

Tabela XII – Configuração da rede e resultados – Teste 08

Configuração		Resultados Treinamento	
Entradas	4	<i>Mean absolute error (MAE)</i>	0,041810
Camada escondida	100 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	0,004721
Saída	1 neurônio (Linear)	<i>Root mean squared error (RMSE)</i>	0,068713
Épocas	2000	<i>R squared (R^2)</i>	0,818176

Figura 37 – Teste 08 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 38 – Teste 08 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Nos testes utilizando dados referentes à potência ocorrem erros mais evidentes se comparados com os testes utilizando dados referentes à curva de energia consumida.

Em todos os testes com dados pertencentes a uma série temporal diferente, houve pontos onde o resultado da previsão foi um valor negativo. Um resultado indesejado visto que a série utilizada no teste não possui valores negativos, tão pouco a série utilizada no treinamento. Esse comportamento ocorre com maior intensidade no teste 8, onde além dos picos negativos são registrados picos positivos muito acima do valor real.

Uma explicação para esse comportamento pode ser relacionada com a quantidade de dados utilizadas no treinamento ter sido mais limitada em função do banco de dados disponível na ocasião. Em alguns testes adicionais foi possível minimizar esse comportamento através da utilização da função de ativação ReLU na camada de saída, ao invés da função linear. Devido à característica da função ReLU, valores negativos são automaticamente aproximados para zero.

Apesar dos problemas relacionados aos valores negativos associados a previsão, o comportamento geral da curva consegue ser previsto pela rede neural.

3.4.3.3 Previsão N Passos à Frente (Wh)

Teste 09 – Reshape (Wh)

O nono teste utilizou a configuração apresentada na Tabela XIII. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 39 e Figura 40.

Tabela XIII – Configuração da rede e resultados – Teste 09

Configuração		Resultados Treinamento	
Entradas	12	Mean absolute error (MAE)	1488,503037
Camada escondida	100 neurônios (ReLU)	Mean squared error (MSE)	3237782,622597
Saída	12 neurônios (Linear)	Root mean squared error (RMSE)	1799,383956
Épocas	2000	R squared (R^2)	0,971982

Figura 39 – Teste 09 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 40 – Teste 09 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 10 – Sequential (Wh)

O décimo teste utilizou a configuração apresentada na Tabela XIV. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 41 e Figura 42.

Tabela XIV – Configuração da rede e resultados – Teste 10

Configuração		Resultados Treinamento	
Entradas	12	Mean absolute error (MAE)	1557,118225
Camada escondida	100 neurônios (ReLU)	Mean squared error (MSE)	3535657,131126
Saída	12 neurônios (Linear)	Root mean squared error (RMSE)	1880,334314
Épocas	2000	R squared (R^2)	0,967675

Figura 41 – Teste 10 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 42 – Teste 10 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Os testes 09 e 10 envolvem a previsão de múltiplos pontos à frente, diferente dos testes realizados anteriormente. As redes utilizadas possuíam 12 entradas referentes aos 12 pontos anteriores, uma camada escondida com 100 neurônios utilizando a função ReLU e 12 saídas utilizando a função de ativação linear. Os resultados das métricas juntamente com a análise visual das previsões demonstraram que a rede não conseguiu generalizar o comportamento referente à curva de energia consumida, gerando previsões muito diferentes da curva real. Além disso, a Figura 42 apresenta pontos onde a variação da energia consumida é negativa em relação ao ponto passado. Assim como nos testes anteriores, uma característica indesejada visto que não ocorre no conjunto de dados utilizados no treinamento nem nos dados reais relativos à previsão.

Outros testes com hiperparâmetros diferentes foram realizados utilizando como entrada a curva de energia consumida, e em todos os casos os resultados foram semelhantes, não conseguindo prever a curva.

Como o objetivo da metodologia proposta é prever múltiplos pontos no futuro foram realizados mais testes, desta vez com os dados referentes a curva de carga, os resultados são apresentados a seguir.

3.4.3.4 Previsão N Passos à Frente (W)

Teste 11 – Reshape (W)

O décimo primeiro teste utilizou a configuração apresentada na Tabela XV. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 43 e Figura 44.

Tabela XV – Configuração da rede e resultados – Teste 11

Configuração		Resultados Treinamento	
Entradas	12	<i>Mean absolute error (MAE)</i>	178,909508
Camada escondida	100 neurônios (ReLU)	<i>Mean squared error (MSE)</i>	99688,269958
Saída	12 neurônios (Linear)	<i>Root mean squared error (RMSE)</i>	315,734493
Épocas	2000	<i>R squared (R^2)</i>	-1,278556

Figura 43 – Teste 11 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 44 – Teste 11 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 12 – Sequential (W)

O décimo segundo teste utilizou a configuração apresentada na Tabela XVI. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 45 e Figura 46.

Tabela XVI – Configuração da rede e resultados – Teste 12

Configuração		Resultados Treinamento	
Entradas	12	Mean absolute error (MAE)	180,933187
Camada escondida	100 neurônios (ReLU)	Mean squared error (MSE)	83501,646231
Saída	12 neurônios (Linear)	Root mean squared error (RMSE)	288,966514
Épocas	2000	R squared (R^2)	-0,786052

Figura 45 – Teste 12 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 46 – Teste 12 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 13 – Reshape (W)

O décimo terceiro teste utilizou a configuração apresentada na Tabela XVII. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 47 e Figura 48.

Tabela XVII – Configuração da rede e resultados – Teste 13

Configuração		Resultados Treinamento	
Entradas	12	Mean absolute error (MAE)	0,124797
Camada escondida	100 neurônios (ReLU)	Mean squared error (MSE)	0,051249
Dropout	0,3 (30%)	Root mean squared error (RMSE)	0,226383
Saída	12 neurônios (Linear)	R squared (R^2)	-0,821832
Épocas	2000

Figura 47 – Teste 13 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 48 – Teste 13 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

Teste 14 – Sequential (W)

O décimo quarto teste utilizou a configuração apresentada na Tabela XVIII. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 49 e Figura 50.

Tabela XVIII – Configuração da rede e resultados – Teste 14

Configuração		Resultados Treinamento	
Entradas	12	Mean absolute error (MAE)	0,144557
Camada escondida	100 neurônios (ReLU)	Mean squared error (MSE)	0,061560
Dropout	0,3 (30%)	Root mean squared error (RMSE)	0,248113
Saída	12 neurônios (Linear)	R squared (R ²)	-1,158179
Épocas	2000

Figura 49 – Teste 14 – Previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 50 – Teste 14 – Previsão utilizando série diferente
Fonte: Elaborada pelo autor

O teste 11 demonstrou um caso de *overfitting*. Ao testar a previsão com os próprios dados de treinamento, o resultado foi igual ao valor real esperado, tanto que se tornou impossível distinguir ambas as curvas. Ao realizar o teste com dados de origem diferente, percebeu-se que apesar da envoltória ser semelhante, existe um elevado *offset* associado à previsão, o que se traduz em energia acumulada muito mais elevada em relação à curva real.

Para os testes 12 e 14, ao utilizar os dados de treinamento a curva prevista assemelhou-se bastante a curva real. Para os dados pertencentes a outro período a rede neural conseguiu prever a envoltória de parte da curva.

3.4.3.5 Configuração Final - Previsão N Passos à Frente - W

A definição da configuração final da rede neural utilizada baseou-se nos resultados empíricos obtidos durante os testes apresentados anteriormente. Além disso, a definição da configuração final foi realizada somente após a coleta de séries temporais mais longas que as anteriormente disponíveis, reproduzindo um resultado mais confiável com relação à previsão da série temporal.

Diversas variações em torno do modelo de base foram testadas e seus resultados foram levados em consideração. O modelo final deveria possuir a capacidade de generalização dos dados de entrada e saída, fornecendo a previsão

dos próximos m pontos, idealmente acompanhando a tendência geral da série, a envoltória, além do consumo de energia elétrica acumulado ao final do período.

Os testes anteriormente apresentados mostraram os resultados de redes com hiperparâmetros bastante diferentes em alguns casos, comparando o resultado de um modelo com 4 neurônios na camada escondida com um outro possuindo 100 neurônios na mesma camada.

As redes com mais neurônios na camada de saída apresentaram melhor resultado de acordo com os testes realizados, entretanto, diversos parâmetros conduziram a rede ao processo de *overfitting*, dentre eles, a configuração dos hiperparâmetros, seja pelo excesso de neurônios na camada, seja pelo excesso de camadas escondidas ou mesmo a quantidade de épocas utilizadas no treinamento.

Após novos testes obteve-se a seguinte configuração, representada na Tabela XIX e na Figura 51:

Tabela XIX – Configuração final da rede neural artificial utilizada

Configuração	
Entradas	12
Camada escondida	32 neurônios (ReLU)
Dropout	0.3 (30%)
Saída	24 neurônios (ReLU)

Figura 51 – Representação do modelo gerada pelo Keras

Fonte: Elaborada pelo autor

Outro critério observado foi a quantidade de épocas para o treinamento já que um número excessivo pode levar a rede ao *overfitting*. A perda da capacidade de generalização da rede pode ser observada através do gráfico das métricas de treinamento e validação.

Cabe aqui destacar que ao longo do treinamento espera-se uma diminuição das perdas a cada iteração, entretanto, deve-se monitorar conjuntamente a perda para os dados de treinamento e validação. A diminuição ou perda da capacidade de generalização é identificada graficamente através de uma inversão na tendência

referente as métricas de perda utilizadas na validação [45]. Mais objetivamente, as métricas de perda obtidos na validação devem diminuir com o passar das épocas. Caso haja uma estagnação na métrica, ou mesmo, uma inversão levando ao aumento das perdas, existe um forte indício de que a capacidade de generalização foi afetada e o processo de treinamento deve ser cessado.

A Figura 52 apresenta as métricas “*loss*” e “*mae*” utilizadas como *feedback* quanto a integridade do processo de treinamento.

Figura 52 – Gráfico com as métricas de perda (*loss*)

Fonte: Elaborada pelo autor

É possível notar que para ambas as métricas ocorrem uma leve inversão nas métricas a partir de aproximadamente 750 épocas. Assim, definiu-se 750 como a quantidade de épocas para o processo de treinamento.

Também com base nos testes foi possível verificar que abordagem “*reshape*” (com deslocamento da janela temporal igual a quantidade de pontos de entrada da rede neural) alcançou melhores resultados. Pode-se associar esse resultado com a diminuição na variabilidade dos dados de entrada, já que ao invés de deslocar a janela temporal unitariamente, desloca-se de acordo com a quantidade de neurônios na entrada. Esse processo acaba dividindo o dia em subperíodos que tendem a possuir características mais delimitadas, facilitando a generalização da rede.

Além disso, adotou-se 24 neurônios na camada de saída, permitindo que sejam previstos 24 pontos além dos dados de entrada. Nos arranjos para múltiplos pontos utilizados durante os testes 11 a 14, a quantidade de saídas das redes foi 12 neurônios, uma quantidade arbitrária escolhida para testar a capacidade de prever *m* pontos à frente. A utilização de 12 neurônios na camada de saída impossibilitou

prever a curva de carga para o período de um dia, visto que o intervalo entre os pontos é de 1 h.

A seguir serão apresentados os resultados do modelo final. Nestes testes foram utilizados dados referentes à série temporal obtida com o protótipo apresentado na seção 3.4.1. Os primeiros testes anteriormente apresentados tomaram como base as séries mais curtas e antigas obtidas com o *smart meter* anteriormente desenvolvido. Apesar das séries curtas, serviu como base de teste para pré-seleção dos modelos.

Teste Modelo Selecionado - Reshape (W)

O teste com o modelo selecionado utilizou a configuração apresentada na Tabela XX. Os resultados dos testes de previsão utilizando os dados de treinamento e os dados exógenos são apresentados na Figura 53 e Figura 54.

Tabela XX – Configuração da rede selecionada e resultados

Configuração		Resultados Treinamento	
Entradas	12	Mean absolute error (MAE)	0,040438
Camada escondida	32 neurônios (ReLU)	Mean squared error (MSE)	0,004756
Dropout	0,3 (30%)	Root mean squared error (RMSE)	0,063626
Saída	24 neurônios (Linear)	R squared (R^2)	0,243020
Épocas	750

Figura 53 – Teste de previsão com a própria série de treinamento
Fonte: Elaborada pelo autor

Figura 54 – Teste de previsão utilizando série diferente
Fonte: Elaborada pelo autor

Para a configuração final, no teste de previsão utilizando a própria série de treinamento, a curva gerada pela rede neural acompanha toda a envoltória da curva original, com diferença em relação ao valor dos picos previsto que são menores que os reais.

Para o teste com dados de uma série diferente, o mesmo comportamento se repetiu. A envoltória geral da curva foi prevista pela rede neural.

Os melhores resultados destes testes em relação aos anteriores podem ser relacionados com a otimização dos hiperparâmetros através da busca empírica pelas melhores configurações. Também pode ser relacionada com a maior quantidade de dados de treinamento utilizados no teste. Diferente dos testes iniciais, a série temporal utilizada no teste da configuração final possuía muito mais dados, melhorando os resultados do treinamento.

3.4.4 Previsão Utilizando Decomposição Temporal

Nesta seção são apresentados os resultados dos testes utilizando a técnica de decomposição temporal abordada na seção 2.5.

O processo de decomposição temporal possui extremo baixo custo computacional além de boa assertividade em dados que possuem componentes de tendência e sazonalidade melhor definidos.

Apesar da variabilidade de comportamento, os consumidores residenciais tendem a possuir hábitos diários que acabam refletindo em seu consumo e curva de carga resultante, assim, há de se esperar uma certa componente sazonal, mesmo que possuindo maior variabilidade quando comparada com o mesmo método aplicado a um conjunto de consumidores.

Vetores de entrada e saída aplicados à decomposição temporal tem como origem a série temporal registrada. Adotam-se duas janelas temporais, uma contendo m dados passados e outra contendo n dados futuros, assim, são criados pares de vetores contendo m dados de entrada e n dados de saída utilizados durante o processo de treinamento e teste.

A Figura 55 apresenta um exemplo de janelas contendo 48 pontos passados para entrada e 24 pontos futuros para a saída.

Segmentos da série temporal											
-48	-47	...	-2	-1	0	1	2	...	...	22	23
Dados de entrada para a Decomposição					Dados de saída – Período de previsão						

Figura 55 – Entrada de dados para a decomposição temporal

Fonte: Elaborada pelo autor

O método pode ser modificado com base em dois parâmetros, números de ciclos passados utilizados na decomposição, origem dos dados passados. O número de ciclos tem relação com a quantidade de ciclos anteriores ao período da previsão. Neste caso cada ciclo possui 24 pontos referente as 24 horas do dia.

A origem dos dados tem relação de como eles foram separados dentro da série temporal. Foram testadas duas abordagens, onde, na primeira foram considerados os ciclos sequenciais anteriores, por exemplo, considerando 2 ciclos, para a previsão dos 24 pontos de uma segunda-feira serão necessários os dois ciclos diretamente anteriores, ou seja, os pontos referentes ao domingo e ao sábado. Na segunda abordagem os dados de origem estão relacionados ao mesmo dia da semana que será realizado a previsão. Por exemplos, considerando 2 ciclos, para previsão dos 24 pontos de uma segunda-feira serão necessários os dois ciclos anteriores, neste caso, referentes às curvas das duas últimas segundas-feiras.

O teste destas duas variações buscou avaliar sobre diferenças significativas com relação à quantidade de ciclos, além da origem dos dados utilizados na decomposição temporal.

A comparação entre os resultados se baseou na análise das métricas de erro médio percentual (MPE) e erro absoluto médio percentual (MAPE). Suas formulações são apresentadas respectivamente em (16) e (17).

$$MPE = \frac{100\%}{n} \sum_{t=1}^n \frac{A_t - F_t}{A_t} \quad (16)$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (17)$$

Onde:

A_t é o valor atual
 F_t é o valor da previsão
 n é o número de amostras

Após os testes constatou-se que os melhores resultados foram obtidos com 2 ciclos sequenciais. Os testes utilizando a abordagem de mesmo dia não tiveram resultados uniformes, com cada dia tendo um resultado diferente. Os resultados obtidos estão divididos entre a Tabela XXI e Tabela XXII.

Tabela XXI – Configuração da rede selecionada e resultados – Parte 1

Ciclos	Sequencial		Mesmo dia - Dia 0		Mesmo dia - Dia 1	
	MPE	MAPE	MPE	MAPE	MPE	MAPE
2	-3,66%	6,35%	-6,80%	9,69%	-6,87%	7,81%
3	-9,86%	9,86%	-9,84%	10,80%	-9,59%	9,59%
4	-7,86%	8,99%	-7,16%	8,73%	-8,96%	8,96%
5	-11,39%	11,39%	-5,93%	10,67%	-12,10%	10,09%

Tabela XXII – Configuração da rede selecionada e resultados – Parte 2

Ciclos	Mesmo dia - Dia 2		Mesmo dia - Dia 3		Mesmo dia - Dia 4		Mesmo dia - Dia 5		Mesmo dia - Dia 6	
	MPE	MAPE	MPE	MAPE	MPE	MAPE	MPE	MAPE	MPE	MAPE
2	-13,15%	13,15%	-1,13%	10,29%	-6,17%	6,45%	-5,96%	7,02%	-6,02%	6,02%
3	-14,20%	14,43%	-6,12%	9,64%	-4,40%	6,60%	-5,58%	5,58%	-6,13%	7,11%
4	-17,16%	17,16%	-6,93%	7,49%	-1,77%	4,08%	-3,45%	4,13%	-4,43%	4,43%
5	-18,94%	15,78%	-10,26%	6,84%	-2,93%	3,19%	-2,86%	3,40%	-4,52%	4,75%

Assim, adotou-se 2 ciclos sequenciais como configuração para a decomposição temporal, sendo que, para cada teste foram utilizados períodos diferentes. A seguir são apresentados os resultados de previsões utilizando essa configuração.

Figura 56 – Teste 01 – Previsão com decomposição temporal – Ciclos sequenciais
Fonte: Elaborada pelo autor

Figura 57 – Teste 02 – Previsão com decomposição temporal – Ciclos sequenciais
Fonte: Elaborada pelo autor

Figura 58 – Teste 03 – Previsão com decomposição temporal – Ciclos sequenciais
Fonte: Elaborada pelo autor

Figura 59 – Teste 04 – Previsão com decomposição temporal – Ciclos sequenciais
Fonte: Elaborada pelo autor

Nas Figuras 54 a 57, os trechos em vermelho representam os dados utilizados na técnica de decomposição temporal, neste caso, os últimos 2 ciclos completos, totalizando 48 pontos. Os trechos em amarelo apresentam o resultado real do próximo dia, contendo 24 pontos referentes às 24 horas. A curva em verde apresenta o resultado da previsão usando a decomposição temporal, com os pontos em verde destacando a previsão em cada intervalo.

Observando as formas de onda é possível notar que o método permitiu acompanhar a envoltória da onda real, inclusive prevendo picos de consumo de acordo com os comportamentos passados.

Nessa técnica, a previsão do próximo período está diretamente relacionada com as características dos ciclos utilizados na decomposição, assim, caso exista uma variação brusca de comportamento entre os ciclos passados e o próximo (previsão), é esperado que a curva não acompanhe perfeitamente a curva real, em especial nos períodos em que houver variações de picos muito grandes.

3.4.5 Nova Abordagem Proposta para Previsão do Consumo

Diversos estudos abordam a utilização das redes neurais [46], [47] para previsão do consumo, seja uma etapa à frente ou n etapas à frente. Entretanto, erros na previsão são esperados, sendo influenciados pela quantidade de dados e padrões utilizados durante o processo de treinamento, ou mesmo devido a diferenças de comportamento entre os dados utilizados no treinamento e os dados utilizados como entrada para a previsão.

Com o objetivo de minimizar esse tipo de erro buscou-se combinar o resultado da previsão realizada com uma rede neural e o resultado da previsão através da decomposição temporal. Foram testadas duas possibilidades de combinação:

- (i) Na primeira calculou-se a média de cada um dos pontos das previsões, o que resulta em uma nova curva, na qual os pontos fora do padrão tendem a serem atenuados; e
- (ii) Na segunda adotou-se uma abordagem adaptativa, na qual, o erro entre as previsões de consumo obtidas pelos dois métodos de previsão, foi utilizado para atualizar multiplicadores para correções das curvas. Esses multiplicadores foram associados a cada um dos métodos de previsão, sendo utilizados para compensar o erro de cada método de previsão individualmente, uma abordagem semelhante aos ajustes realizados por um controlador proporcional. Após a correção as duas curvas foram combinadas novamente através da média ponto a ponto, gerando uma nova curva.

Para avaliar os resultados foram utilizadas as métricas de erro médio absoluto percentual (MAPE) e erro médio percentual (MPE). Após um determinado número de previsões sequenciais foram verificados os resultados.

No próximo capítulo serão apresentados os resultados do sistema desenvolvido utilizando a metodologia proposta.

4 RESULTADOS EXPERIMENTAIS

A metodologia proposta consistiu na utilização dos dados referente às curvas de carga (Watt) de uma residência para a previsão da curva de carga de um período futuro. Como o objetivo é prever a energia consumida ao final do período, a curva resultante da previsão é integrada no período, fornecendo como resultado, a curva de consumo acumulado para o período referente às 24 horas da previsão. O mesmo procedimento foi realizado com a curva de carga real, permitindo o cálculo das métricas de erro entre a previsão e o dado real.

Os dados utilizados no processo de treinamento e teste foram obtidos de períodos diferentes da coleta de dados, permitindo verificar o real comportamento do sistema com relação à previsão, tanto individual de cada método, como da metodologia proposta combinando as duas previsões. Em uma aplicação real, os dados passados foram utilizados para previsão das curvas futuras, assim, a utilização de dados referentes e séries com períodos de coleta diferentes permitiu avaliar os resultados mais próximos da realidade. Ressalta-se que é importante que os dados utilizados tenham relação com os anteriores, ou seja, tenham a mesma origem.

Na previsão com a rede neural os dados de entrada e saída para o treinamento foram criados através de janelas na série temporal. Os dados de entrada englobaram uma janela de 12 pontos passados (anteriores) da curva de carga e os dados de saída englobaram uma janela de 24 pontos da janela temporal que estiveram atrasados em relação à entrada, ou seja, o primeiro ponto seguinte ao final da janela de entrada é o início da janela de saída.

Para a decomposição temporal foram utilizados como referência os dois ciclos anteriores a curva que será prevista, ou seja, são necessários 48 pontos passados para previsão dos próximos 24 pontos (horas), conforme ilustrado na Figura 60.

Segmentos da série temporal											
-12	-11	...	-2	-1	0	1	2	...	...	22	23
Dados de entrada para a ANN					Dados de saída para a ANN – Período de previsão						
-48	-47	...	-2	-1	0	1	2	...	...	22	23
Dados de entrada para a Decomposição					Dados de saída – Período de previsão						

Figura 60 – Entrada de dados para os métodos de previsão

Fonte: Elaborada pelo autor

Os testes foram conduzidos através de dois arranjos:

- (i) No primeiro considerou-se que os dados anteriores seriam referentes aos dias sequenciais; e
- (ii) No segundo utilizou-se os dados anteriores, neste caso, referentes ao mesmo dia da semana que será feita a previsão.

Os testes foram idealizados para determinar a existência de diferenças significativas entre as abordagens, e os melhores resultados.

Nos testes foram avaliados os resultados referentes as duas possibilidades de combinação, a decomposição temporal e a rede neural. Na primeira combinação calculou-se a média ponto-a-ponto das curvas resultantes, ou seja, sem considerar os multiplicadores de correção. Na segunda combinação, considerou-se os multiplicadores de correção, o que permitiu avaliar o impacto sobre a previsão ao longo do tempo.

4.1 Teste Utilizando Dias Sequenciais

Conforme apresentado, os testes envolveram a previsão pelo método de decomposição temporal, bem como através da rede neural. Ao final tendo seus resultados combinados para formar uma única série.

No teste com arranjo de dias sequenciais considerou-se os resultados de erro MAPE e MPE para 100 previsões.

Para a primeira combinação, calculou-se a média entre os pontos dos mesmos períodos, os resultados obtidos foram:

- MAPE: 9,9718%; e
- MPE: -2,6853%.

Para a segunda combinação, considerou-se a correção dos multiplicadores em função do erro, os resultados obtidos foram:

- MAPE: 9,8899%; e
- MPE: -2,8400%.

As figuras a seguir apresentam os resultados da previsão através da metodologia proposta. As curvas referentes aos dois métodos de previsão são apresentadas isoladamente para discussão do comportamento.

Figura 61 – Dados utilizados na previsão e resultados das previsões – Teste 01 – Dias sequenciais
Fonte: Elaborada pelo autor

Figura 62 – Destaque para o período da previsão – Teste 01 – Dias sequenciais
Fonte: Elaborada pelo autor

Na Figura 61 é ilustrado o resultado geral da previsão, contendo os dados utilizados como entrada para os métodos de previsão juntamente com os resultados obtidos pelos métodos individualmente.

A Figura 62 apresenta apenas o período da previsão, permitindo analisar melhor as curvas obtidas. A curva real é apresentada na cor amarela, a previsão através da decomposição temporal na cor verde e a previsão através da rede neural na cor preta.

Analisando as curvas é possível notar que a previsão realizada pela rede neural (preto) acompanha a tendência da curva real, entretanto existe um certo *offset* entre as duas curvas. O resultado da decomposição temporal possui um *offset* menor se comparada ao resultado da rede neural, no entanto, é evidente a previsão de um pico de consumo que não ocorre na prática, sendo este resultado decorrente das características dos ciclos anteriores aplicados ao método.

Figura 63 – Curva real vs. curva da previsão combinada – Teste 01 – Dias sequenciais
Fonte: Elaborada pelo autor

A Figura 63 indica o resultado combinado das previsões. A tendência geral de consumo é acompanhada, apesar do *offset* entre as curvas. Apesar do pico de consumo ainda estar presente, seu valor é menor devido à combinação dos métodos.

Figura 64 – Dados utilizados na previsão e resultados das previsões – Teste 02 – Dias sequenciais
Fonte: Elaborada pelo autor

Figura 65 – Destaque para o período da previsão – Teste 02 – Dias sequenciais
Fonte: Elaborada pelo autor

Na Figura 64 é ilustrado o resultado geral da previsão, contendo os dados utilizados como entrada para os métodos de previsão juntamente com os resultados obtidos pelos métodos individualmente.

A Figura 65 indica apenas o período da previsão, permitindo analisar melhor as curvas obtidas. A curva real é apresentada na cor amarela, a previsão através da decomposição temporal na cor verde e a previsão através da rede neural na cor preta.

Observando as curvas é possível notar que a previsão realizada pela rede neural (preto) acompanha a tendência da curva real, todavia, está um pouco defasada com relação ao principal pico de consumo, neste caso, não existindo *offset*

notável entre as curvas. O resultado da decomposição temporal por sua vez acompanhou a envoltória da curva, inclusive o pico de consumo, não possuindo *offset* aparente.

Figura 66 – Curva real vs. curva da previsão combinada – Teste 02 – Dias sequenciais
Fonte: Elaborada pelo autor

Na Figura 66 é mostrado o resultado combinado das previsões. A curva resultante acompanha tanto a tendência da curva real como sua envoltória, mesmo com a defasagem associada à previsão realizada pela rede neural.

A tendência geral de consumo é acompanhada, além de não existir um *offset* aparente entre as curvas.

4.2 Teste Utilizando Dias Iguais

A mesma sequência de testes foi repetida, mas agora utilizando dias iguais para o processo de previsão. Essa abordagem diminuiu bastante a quantidade dos dados utilizados no treinamento e nos testes, já que a série de origem é a mesma do teste anterior limitando os dados de acordo com o dia. O teste com arranjo de mesmos dias considerou os resultados de erro MAPE e MPE para 14 previsões.

Para a primeira combinação, considerou-se a média entre os pontos dos mesmos períodos, e os resultados obtidos foram:

- MAPE: 13,34%; e
- MPE: 1,5688%.

Para a segunda combinação, adotou-se a correção dos multiplicadores em função do erro, os resultados obtidos foram:

- MAPE: 13,2988%; e
- MPE: 1,1562%.

As figuras a seguir ilustram os resultados da previsão através da metodologia proposta. As curvas referentes aos dois métodos de previsão são apresentadas isoladamente para discussão do comportamento.

Figura 67 – Dados utilizados na previsão e resultados das previsões – Teste 01 – Dias iguais
Fonte: Elaborada pelo autor

Figura 68 – Destaque para o período da previsão – Teste 01 – Dias iguais
Fonte: Elaborada pelo autor

Na Figura 67 é mostrado o resultado geral da previsão, contendo os dados utilizados como entrada para os métodos de previsão juntamente com os resultados obtidos pelos métodos individualmente.

A Figura 68 indica apenas o período da previsão, permitindo analisar melhor as curvas obtidas. A curva real é apresentada na cor amarela, a previsão através da decomposição temporal na cor verde e a previsão através da rede neural na cor preta.

Analisando as curvas é possível verificar que a previsão realizada pela rede neural (preto) acompanha a tendência da curva real em todo o período, não existindo *offset* aparente entre as duas curvas. O resultado da decomposição temporal também acompanha a envoltória da curva real, no entanto, percebe-se a existência de um pico de consumo devidos aos padrões e consumo dos dias passados, resultado semelhante ao ocorrido no teste com dias sequenciais.

Figura 69 – Curva real vs. curva da previsão combinada – Teste 01 – Dias iguais
Fonte: Elaborada pelo autor

Na Figura 69 é verificado o resultado combinado das previsões. A tendência geral de consumo é acompanhada em quase toda a extensão da curva. O maior erro ocorreu devido ao pico de consumo previsto pela decomposição temporal, todavia, seu efeito é atenuado devido a combinação dos métodos.

Figura 70 – Dados utilizados na previsão e resultados das previsões – Teste 02 – Dias iguais
Fonte: Elaborada pelo autor

Figura 71 - Destaque para o período da previsão – Teste 02 – Dias iguais
Fonte: Elaborada pelo autor

Na Figura 70 é ilustrado um resultado geral da previsão, contendo os dados utilizados como entrada para os métodos de previsão juntamente com os resultados obtidos pelos métodos individualmente.

A Figura 71 apresenta apenas o período da previsão, permitindo analisar melhor as curvas obtidas. A curva real é apresentada na cor amarela, a previsão através da decomposição temporal na cor verde e a previsão através da rede neural na cor preta.

Verificando as curvas é possível observar que ambos os métodos isoladamente não conseguiram acompanhar totalmente a curva real. No caso da rede neural este fator pode ser explicado pelo comportamento da previsão ser muito

diferente dos padrões utilizados durante o treinamento. Para a decomposição temporal, devido a mudanças de comportamento dos últimos ciclos utilizados.

Figura 72 – Curva real vs. curva da previsão combinada – Teste 02 – Dias iguais
Fonte: Elaborada pelo autor

A partir da Figura 72 é observado o resultado combinado das previsões. Apesar de ambos os métodos não terem resultado em curvas semelhantes a real, a combinação entre os métodos originou uma curva que, em parte do período acompanha a curva real (nas primeiras 12 horas), não conseguindo acompanhar a diminuição no consumo nas 12 horas restantes.

4.3 Análise da Função ReLU e do Dropout nos Resultados

A função de ativação utilizada em todas as camadas foi a *Rectified Linear Unit* (ReLU), que tem como característica retornar um comportamento linear para valores de entrada positivos e zero para valores negativos. Essa função de ativação tem se disseminado com o avanço do *deep-learning*, devido às suas características, dentre elas, a facilidade de implementação e o baixíssimo custo computacional. Complementarmente, sua característica linear para valores positivos facilita processos de otimização baseados em gradiente de erro [48].

Em testes anteriormente desenvolvidos foi possível verificar algumas características interessantes com relação à utilização da função de ativação ReLU, entre elas o fato de inibir resultados negativos, o que pode acontecer com a função linear ou mesmo com a função tangente hiperbólica, além disso, a função ReLU

permite que os padrões de entrada estejam em escalas diferentes sem que a rede saturar o seu resultado, como acontece com a função sigmoide que limita o range de saída entre zero e um.

Apesar da série temporal original possuir muitas amostras, após o *downsampling* a quantidade de dados resultantes diminui bastante. Além disso, de acordo com a quantidade de épocas escolhidas para o treinamento pode ocorrer da rede neural entrar em *overfitting*, o que leva a resultados significativos para o treinamento e resultados inferiores quando a previsão é aplicada a dados exógenos. Para minimizar este efeito utilizou-se a técnica de regularização por *dropout*.

A regularização por *dropout* é uma técnica desenvolvida em [49], que busca evitar o *overfitting* através da desconexão aleatória de um percentual dos neurônios a cada iteração do processo de treinamento, forçando que a cada mudança a rede neural aprenda um novo caminho. Um processo similar a treinar muitas redes neurais com diferentes configurações em paralelo. Complementarmente os pesos são ajustados de acordo com a taxa de *dropout* já que a diminuição momentânea do número de neurônios tende a elevar o peso dos remanescentes [50].

Na implementação, a camada de *dropout* foi adicionada entre a camada escondida e a camada de saída. É importante salientar que a camada de *dropout* está ativa apenas durante o treinamento. Durante as previsões a camada de *dropout* se torna inativa, não interferindo no comportamento, que será igual a uma rede Perceptron multicamadas (*Multilayer Perceptron* – MLP) convencional, também referenciada como *fully-connected* na biblioteca *Keras*.

No próximo capítulo são apresentadas as conclusões relacionadas a esta tese.

5 CONCLUSÕES

5.1 Aspectos Gerais

Grande parte dos dispositivos e equipamentos atualmente ofertados aos consumidores residenciais são desenvolvidos utilizando tecnologias intrinsecamente eficientes. Assim, futuras atualizações tendem a não apresentar melhorias significativas com relação à eficiência energética. Em uma realidade onde a maioria dos dispositivos e equipamentos eletrônicos de uma residência possuem alta eficiência energética, a busca pela melhoria contínua torna-se mais efetiva através das alterações nos padrões de consumo dos consumidores, e não somente considerando a substituição dos equipamentos.

Esta tese apresentou o desenvolvimento de uma nova metodologia para previsão do consumo de energia elétrica de curto prazo utilizando redes neurais artificiais e decomposição de séries temporais, com os dados utilizados para o seu desenvolvimento sendo obtidos através de um medidor inteligente instalado em uma residência real.

Atualmente a previsão do consumo de energia elétrica já é utilizada no setor elétrico, entretanto, as abordagens apresentadas geralmente visam grandes edificações ou mesmo os grandes *players* do sistema, deixando de lado o consumidor residencial. Apesar deste não representar a maior fatia dos consumidores de eletricidade, são um dos grupos com maior perspectiva de crescimento em consumo, bem como na possibilidade de melhorias nos índices de eficiência energética.

No contexto apresentado, a previsão de consumo de energia de curto prazo se torna uma importante ferramenta de conscientização através do fornecimento de informações úteis ao consumidor residencial, auxiliando na busca por padrões mais eficientes de consumo e economia de energia.

Durante o desenvolvimento dessa tese foram encontrados diversos desafios, dentre os principais podem-se elencar: (i) a quantidade de dados necessários para o processo de treinamento e validação da abordagem neural, um ponto resolvido com o desenvolvimento de um protótipo para coleta de longas séries temporais; (ii) questões relacionadas ao salvamento dos dados coletados pelo protótipo devido às condições adversas, como interrupção no fornecimento de energia ou mesmo quedas de conexão com a internet; (iii) desafios relacionados ao tratamento dos dados disponibilizados pelo protótipo, visando corrigir qualquer inconsistência que pudessem afetar as etapas seguintes; e (iv) a busca pelo melhor arranjo para os métodos utilizados.

Através da análise das curvas apresentadas é possível verificar que a metodologia proposta conseguiu acompanhar a tendência dos dados reais, ou mesmo, a envoltória da curva de carga. No caso da rede neural, dados de testes com comportamento muito diferentes dos dados de treinamento podem originar maiores erros na curva prevista. Para a decomposição temporal, erros mais elevados ocorrem quando os períodos anteriores apresentam muita variabilidade entre si, não apresentando uma sazonalidade bem definida, afetando o comportamento da previsão por este método.

Ainda para a decomposição temporal, as características de consumo afetam na escolha entre as variações de dia sequencial ou mesmo dia. A hipótese inicial era que a variação de mesmo dia apresentaria melhor resultado na maioria dos testes, entretanto, como o padrão de consumo da residência onde os dados foram coletados são semelhantes durante toda a semana, ocorreu de a variação com padrão dias sequenciais retornar melhores resultados. No caso de uma residência com padrão de consumo diferente nos finais de semana, é de se esperar que o erro seja mais elevado, principalmente para previsão do consumo dos dois dias subsequentes ao final de semana bem como para previsão do consumo durante o final de semana.

Para ambos os testes ocorreu uma diminuição no MAPE após a implementação da estratégia de correção do erro. Este fato pode ser explicado pelas correções aplicadas aos multiplicadores de cada método que a cada nova previsão tendem a minimizar o erro geral.

Ambos os resultados demonstram que a metodologia utilizada permite boas previsões através da combinação da previsão utilizando redes neurais artificiais com a previsão utilizando decomposição temporal, e que ao longo do tempo os valores de MAPE e MPE tendem a serem minimizados. Mesmo nos testes com dados do mesmo dia da semana os resultados foram bastante promissores, visto que alcançou um erro médio de apenas 1,1562% após 14 períodos.

Durante o desenvolvimento do trabalho ficou evidente que o volume de dados disponíveis para o treinamento da rede neural afeta diretamente a capacidade de previsão da metodologia proposta. Na presença de maior volume de dados coletados, em diferentes residências com padrão de consumo semelhante, é de se esperar uma melhoria na previsão.

Grande parte dos artigos relacionados à previsão do consumo de energia elétrica não deixam totalmente claro quais as configurações e parâmetros utilizados no desenvolvimento das soluções propostas, em muitos casos tornando impossível saber se os resultados dos gráficos apresentados são referentes à previsão de um ponto à frente ou múltiplos pontos à frente, uma fator que interfere na análise dos resultados ou mesmo em sua aplicabilidade.

Devido à utilização da linguagem *Python* juntamente com a biblioteca *Keras*, além da implementação discreta do algoritmo para decomposição temporal, torna-se possível a implementação desta metodologia em um sistema embarcado baseado em uma *single board computer*, como a *Raspberry Pi*, que pode ser integrada ao sistema de uma *smart home*.

Com base nos microcontroladores lançados recentemente com mais memória *RAM* e *Flash*, além do surgimento de bibliotecas específicas para aplicações de inferência em *deep-learning*, voltadas aos sistemas embarcados e suas restrições, pode-se esperar que em breve será possível realizar a previsão em sistemas embarcados bastante reduzidos e de baixo custo.

5.2 Desdobramentos da Pesquisa

Como desdobramento desta tese vislumbra-se a possibilidade de integração da abordagem proposta em um sistema embarcado baseado em um

microcontrolador. O que tornaria possível sua integração no próprio medidor inteligente. Dentre os modelos disponíveis atualmente, o modelo *ESP32* da *Espressif* se mostra como uma possível alternativa devido à quantidade de memória e à capacidade de processamento, além da conectividade *wireless*.

Além disso, torna-se possível a implementação em um sistema de gerenciamento de uma casa inteligente, fornecendo os dados para o consumidor através de uma interface *web*, ou aplicativo para o *smartphone*. A partir deste tipo de implementação surgem outras derivações, como a utilização dos dados relativos à previsão para auxiliar o controle de carga ou mesmo aplicar a mesma metodologia a outros tipos de dados, como os dados relacionados à irradiação solar, permitindo prever qual o melhor momento para armazenar energia nas baterias, ou qual o melhor momento para realizar injeção de energia na rede, no caso de um sistema fotovoltaico *on-grid*.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] MME - Ministério de Minas e Energia, “Plano Decenal de Expansão de Energia - 2027”, Brasília/DF, 2018.
- [2] H. L. M. do Amaral, A. N. de Souza, D. S. Gastaldello, F. Fernandes, e Z. Vale, “Smart meters as a tool for energy efficiency”, in *2014 11th IEEE/IAS International Conference on Industry Applications*, 2014, p. 1–6.
- [3] H. L. M. do Amaral, “DESENVOLVIMENTO DE UM SMART METER: um estudo sobre eficiência energética através das redes inteligentes”, Universidade Estadual Paulista (UNESP), 2014.
- [4] H. L. M. do Amaral, A. N. de Souza, D. S. Gastaldello, T. X. da S. Palma, A. D. S. Maranhão, e J. P. Papa, “Use of virtual load curves for the training of neural networks for residential electricity consumption forecasting applications”, in *2018 13th IEEE International Conference on Industry Applications (INDUSCON)*, 2018, p. 85–90.
- [5] Y. Zhang, H. Li, Z. Wang, W. Zhang, e J. Li, “A preliminary study on time series forecast of fair-weather atmospheric electric field with WT-LSSVM method”, *J. Electrostat.*, vol. 75, p. 85–89, jun. 2015.
- [6] A. Tarsauliya, R. Kala, R. Tiwari, e A. Shukla, “Financial Time Series Volatility Forecast Using Evolutionary Hybrid Artificial Neural Network”, in *Communications in Computer and Information Science*, vol. 196 CCIS, 2011, p. 463–471.
- [7] L. Yunpeng, H. Di, B. Junpeng, e Q. Yong, “Multi-step Ahead Time Series Forecasting for Different Data Patterns Based on LSTM Recurrent Neural Network”, in *2017 14th Web Information Systems and Applications Conference (WISA)*, 2017, vol. 2018-Janua, p. 305–310.
- [8] S. A. Fatemi, A. Kuh, e M. Fripp, “Parametric methods for probabilistic

- forecasting of solar irradiance”, *Renew. Energy*, vol. 129, p. 666–676, dez. 2018.
- [9] P. Damrongkulkamjorn e P. Churueang, “Monthly energy forecasting using decomposition method with application of seasonal ARIMA”, in *2005 International Power Engineering Conference*, 2005, vol. 2005, n° December, p. 1–229.
- [10] H. Zhao e F. Magoulès, “A review on the prediction of building energy consumption”, *Renew. Sustain. Energy Rev.*, vol. 16, n° 6, p. 3586–3592, ago. 2012.
- [11] M. A. R. Biswas, M. D. Robinson, e N. Fumo, “Prediction of residential building energy consumption: A neural network approach”, *Energy*, vol. 117, p. 84–92, 2016.
- [12] H. M. Sri, P. Rao, P. K. Kammardi, S. S. Shekar, S. Kathavate, e K. H. Gowranga, “A smart adaptive LSTM technique for electrical load forecasting at source”, in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 2017, vol. 2018-Janua, p. 1717–1721.
- [13] W. He, “Load Forecasting via Deep Neural Networks”, *Procedia Comput. Sci.*, vol. 122, p. 308–314, 2017.
- [14] K. P. Amber, R. Ahmad, M. W. Aslam, A. Kousar, M. Usman, e M. S. Khan, “Intelligent techniques for forecasting electricity consumption of buildings”, *Energy*, vol. 157, p. 886–893, ago. 2018.
- [15] S. B. Shinde e P. K. Katti, “Smart measurement techniques for efficient operation of smart grid”, in *2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]*, 2015, p. 1–6.
- [16] X. Fang, S. Misra, G. Xue, e D. Yang, “Smart Grid - The New and Improved Power Grid: A Survey”, *IEEE Commun. Surv. Tutorials*, vol. 14, n° 4, p. 944–980, 2012.
- [17] M. V. da Cunha, D. P. Bernardon, e D. B. Ramos, “Técnicas de GLD aplicadas ao Consumidor de BT considerando a Tarifa Branca e a Geração Distribuída”, in *Anais do VI Simpósio Brasileiro de Sistemas Elétricos*, 2016.

- [18] A. Zanella, N. Bui, A. Castellani, L. Vangelista, e M. Zorzi, “Internet of Things for Smart Cities”, *IEEE Internet Things J.*, vol. 1, nº 1, p. 22–32, fev. 2014.
- [19] R. Bikmetov, M. Y. A. Raja, e T. U. Sane, “Infrastructure and applications of Internet of Things in smart grids: A survey”, in *2017 North American Power Symposium (NAPS)*, 2017, p. 1–6.
- [20] EPE - Empresa De Pesquisa Energética, “Estudos da Demanda de Energia - Demanda da Energia 2050”, Rio de Janeiro/RJ, 2016.
- [21] EPE - Empresa De Pesquisa Energética, “Anuário Estatístico de Energia Elétrica 2015 ano base 2014”, Rio de Janeiro/RJ, 2015.
- [22] EPE - Empresa De Pesquisa Energética, “Anuário Estatístico de Energia Elétrica 2018 ano base 2017”, Rio de Janeiro/RJ, 2018.
- [23] EPE - Empresa De Pesquisa Energética, “Balanço Energético Nacional - 2015 ano 2014”, 2015.
- [24] T. C. Soares, L. S. Lopes, e D. A. da Cunha, “A eficiência do consumo residencial de energia elétrica no Brasil”, *Econ. Apl.*, vol. 21, nº 3, p. 503–523, 2017.
- [25] R. Yesudas e R. Clarke, “Identifying consumer requirements as an antidote to resistance to smart meters”, in *IEEE PES Innovative Smart Grid Technologies, Europe*, 2014, vol. 2015-Janua, nº January, p. 1–6.
- [26] A. Podgornik, B. Sucic, P. Bevk, e D. Stanicic, “The Impact of Smart Metering on Energy Efficiency in Low-Income Housing in Mediterranean”, 2013, p. 597–614.
- [27] L. Gomes *et al.*, “Dynamic Approach and Testbed for Small and Medium Players Simulation in Smart Grid Environments”, *IFAC Proc. Vol.*, vol. 47, nº 3, p. 31–36, 2014.
- [28] T. S. Choi, K. R. Ko, S. C. Park, Y. S. Jang, Y. T. Yoon, e S. K. Im, “Analysis of energy savings using smart metering system and IHD (In-Home Display)”, *Transm. Distrib. Conf. Expo. Asia Pacific, T D Asia 2009*, p. 1–4, 2009.
- [29] S. M. Basnet, H. Aburub, e W. Jewell, “An artificial neural network-based peak demand and system loss forecasting system and its effect on demand

- response programs”, in *2016 Clemson University Power Systems Conference (PSC)*, 2016, p. 1–5.
- [30] A. Marinescu, C. Harris, I. Dusparic, S. Clarke, e V. Cahill, “Residential electrical demand forecasting in very small scale: An evaluation of forecasting methods”, in *2013 2nd International Workshop on Software Engineering Challenges for the Smart Grid (SE4SG)*, 2013, p. 25–32.
- [31] F. H. Saito e T. Y. Fujii, “MONITORAMENTO DO CONSUMO DE ENERGIA ELÉTRICA EM RESIDÊNCIAS CONECTADAS”, Universidade de São Paulo - USP, 2019.
- [32] F. Fernandes, T. Sousa, P. Faria, M. Silva, H. Morais, e Z. A. Vale, “Intelligent SCADA for Load control”, in *2010 IEEE International Conference on Systems, Man and Cybernetics*, 2010, vol. 440016, nº 1, p. 2987–2994.
- [33] H. S. Hippert, C. E. Pedreira, e R. C. Souza, “Neural networks for short-term load forecasting: a review and evaluation”, *IEEE Trans. Power Syst.*, vol. 16, nº 1, p. 44–55, 2001.
- [34] J. P. Papa e G. H. de Rosa, “APOSTILA DE TÓPICOS ESPECIAIS: APRENDIZADO DE MÁQUINA”, 2019.
- [35] J. Brownlee, *Deep Learning with Python*, V1.15. Vermont-Australia: Machine Learning Martery, 2018.
- [36] K. G. Sheela e S. N. Deepa, “Review on methods to fix number of hidden neurons in neural networks”, *Math. Probl. Eng.*, vol. 2013, 2013.
- [37] J. Deng e P. Jirutitijaroen, “Short-term load forecasting using time series analysis: A case study for Singapore”, *2010 IEEE Conf. Cybern. Intell. Syst. CIS 2010*, p. 231–236, 2010.
- [38] G. Omkar e S. V. Kumar, “Time series decomposition model for traffic flow forecasting in urban midblock sections”, in *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, 2017, p. 720–723.
- [39] H. L. M. do Amaral, J. A. G. Maginador, e D. S. Gastaldello, “Integration of consumption forecasting in smart meters and smart home management systems”, in *2018 Simposio Brasileiro de Sistemas Eletricos (SBSE)*, 2018, p. 1–6.

- [40] F. Ertam e G. Aydin, "Data classification with deep learning using Tensorflow", in *2017 International Conference on Computer Science and Engineering (UBMK)*, 2017, p. 755–758.
- [41] G. Hebrail e A. Berard, "UCI Machine Learning Repository: Individual household electric power consumption Data Set". [Online]. Available at: <https://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption>. [Acessado: 01-mar-2017].
- [42] S. Lee, "Smart - UMass Trace Repository". [Online]. Available at: <http://traces.cs.umass.edu/index.php/Smart/Smart>. [Acessado: 01-mar-2017].
- [43] D. Murray e L. Stankovic, "REFIT: Electrical Load Measurements", *Scientific Data*, 05-dez-2017. [Online]. Available at: <https://pureportal.strath.ac.uk/en/datasets/refit-electrical-load-measurements>. [Acessado: 01-mar-2017].
- [44] J. Brownlee, *Deep Learning for Time Series Forecasting: Predicting the Future with MLPs, CNNs and LSTMs in python*, V1.4. Vermont-Australia: Machine Learning Martery, 2018.
- [45] J. Brownlee, "How to use Learning Curves to Diagnose Machine Learning Model Performance", *Better Deep Learning*, 2019. [Online]. Available at: <https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/>. [Acessado: 05-mar-2019].
- [46] S. M. Sulaiman, P. A. Jeyanthi, e D. Devaraj, "Artificial neural network based day ahead load forecasting using Smart Meter data", in *2016 Biennial International Conference on Power and Energy Systems: Towards Sustainable Energy (PESTSE)*, 2016, p. 1–6.
- [47] B. Asare-Bediako, W. L. Kling, e P. F. Ribeiro, "Day-ahead residential load forecasting with artificial neural networks using smart meter data", in *2013 IEEE Grenoble Conference*, 2013, p. 1–6.
- [48] J. Brownlee, "A Gentle Introduction to the Rectified Linear Unit (ReLU) for Deep Learning Neural Networks", *Machine Learning Mastery*, 2019. [Online]. Available at: <https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/>. [Acessado: 01-ago-2019].

- [49] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, e R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”, *J. Mach. Learn. Res.*, vol. 15, p. 1929–1958, jan. 2014.
- [50] J. Brownlee, “A Gentle Introduction to Dropout for Regularizing Deep Neural Networks”, 2018. [Online]. Available at: <https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks/>. [Acessado: 10-ago-2019].

GLOSSÁRIO

Abordagem “ <i>reshape</i> ”:	Cria vetores que estão temporalmente deslocados por n amostras em relação a janela anterior, onde n é o número de entradas aplicadas simultaneamente a rede neural.
Abordagem “ <i>sequential</i> ”:	Cria vetores que estão unitariamente deslocados em relação ao tempo, ou seja, deslocados por uma amostra em relação a janela temporal anterior.
AFE:	Acrônimo para <i>Analog Front End</i> , que é o conjunto de circuitos analógicos associados a etapa de condicionamento dos sinais, podendo englobar também o conversor analógico-digital (ADC).
<i>Big Data</i> :	Área relacionada a tecnologia da informação que estuda como tratar, analisar e obter informações através de grandes volumes de dados.
<i>Buffer</i> :	Espaço na memória volátil utilizado para armazenar um conjunto de dados temporários entre etapas do processamento.
<i>Deep-Learning</i> :	É uma subárea do aprendizado de máquina que utiliza redes neurais profundas para abstrair e generalizar informações de um grande volume de dados.
<i>Downsampling</i> :	Diminuição no volume dos dados. Redução na quantidade de amostras.
<i>Early Adopters</i> :	Termo geralmente associados a entusiastas de tecnologia que procuram acesso antecipado a novos dispositivos e serviços.
<i>Edge Device</i> :	Em tradução direta, são dispositivos de borda. Dispositivos na camada mais externa de uma rede IoT, geralmente sem grande capacidade de processamento ou conexão direta com a nuvem.
ESP32:	Faz parte de uma série de microcontroladores de baixo custo de 32-bit desenvolvido pela <i>Espressif</i> . Integra no

	mesmo SoC (<i>System on a Chip</i> – Sistema em um chip) o microcontrolador e o <i>hardware</i> necessário para comunicação <i>Wi-Fi</i> e <i>Bluetooth</i> .
Hiperparâmetros:	Variáveis que controlar o próprio processo de treinamento e estrutura da rede neural artificial.
<i>Levenberg-Marquardt</i> :	Método de otimização utilizado para resolução do problema de mínimos quadrados não linear
MAPE:	Acrônimo para <i>Mean Absolute Percentage Error</i> , uma das métricas de erro utilizadas para avaliação das previsões realizadas.
MPE:	Acrônimo para <i>Mean Percentage Error</i> , outra métrica de erro utilizada para avaliação das previsões realizadas.
Nuvem:	Conceito que envolve a disponibilidade sob demanda de recursos, em especial, a capacidade de armazenamento e a capacidade computacional. Tecnologia que torna possível armazenar dados na internet através de um servidor <i>online</i> (sempre disponível).
<i>On-Grid</i> :	Sistema de geração distribuído com capacidade de injeção de potência na rede elétrica da concessionária.
<i>Overfitting</i> :	Termo utilizado para descrever um comportamento da rede neural onde seu comportamento se ajusta bem para os dados de treinamento, entretanto se mostra ineficaz para conjuntos de dados diferentes, ou seja, demonstra que a rede neural não conseguiu generalizar o comportamento dos dados de treinamento.
<i>Raspberry Pi</i> :	Uma série de pequenos computadores de placa única (<i>single board computer</i>) amplamente difundida entre os entusiastas de sistemas embarcados e <i>Linux</i> embarcado.
ReLU:	Acrônimo para <i>Rectified Linear Unit</i> . Um tipo de função de ativação com característica linear para valores positivos e não linear para valores negativo, retornando o próprio valor na sua entrada caso positivo, ou zero, caso negativo. Uma função de ativação que vem se popularizando bastante com o crescimento dos estudos relacionados ao <i>deep-learning</i> .

Single Board Computer.

Em uma tradução direta, computador de uma única placa. Termo geralmente associado a placas de desenvolvimento que possuem um processador com maior capacidade de processamento, tradicionalmente utilizando alguma variação embarcada do sistema operacional *Linux*.