



Ricardo Augusto Manfredini
Geraldo Nunes Corrêa
Bruno Rodrigues de Oliveira
Suellen Teixeira Zavadzki de Pauli
Organizadores

Aplicações de Machine Learning



2021

Ricardo Augusto Manfredini
Geraldo Nunes Corrêa
Bruno Rodrigues de Oliveira
Suellen Teixeira Zavadzki de Pauli
Organizadores

Aplicações de Machine Learning



Pantanal Editora

2021

Copyright© Pantanal Editora

Editor Chefe: Prof. Dr. Alan Mario Zuffo

Editores Executivos: Prof. Dr. Jorge González Aguilera e Prof. Dr. Bruno Rodrigues de Oliveira

Diagramação: A editora. **Diagramação e Arte:** A editora. **Imagens de capa e contracapa:** Canva.com. **Revisão:** O(s) autor(es), organizador(es) e a editora.

Conselho Editorial

Grau acadêmico e Nome

Prof. Dr. Adaylson Wagner Sousa de Vasconcelos
Prof. Msc. Adriana Flávia Neu
Prof. Dra. Albys Ferrer Dubois
Prof. Dr. Antonio Gasparetto Júnior
Prof. Msc. Aris Verdecia Peña
Prof. Arisleidis Chapman Verdecia
Prof. Dr. Arinaldo Pereira da Silva
Prof. Dr. Bruno Gomes de Araújo
Prof. Dr. Caio Cesar Enside de Abreu
Prof. Dr. Carlos Nick
Prof. Dr. Claudio Silveira Maia
Prof. Dr. Cleberton Correia Santos
Prof. Dr. Cristiano Pereira da Silva
Prof. Ma. Dayse Rodrigues dos Santos
Prof. Msc. David Chacon Alvarez
Prof. Dr. Denis Silva Nogueira
Prof. Dra. Denise Silva Nogueira
Prof. Dra. Dennyura Oliveira Galvão
Prof. Dr. Elias Rocha Gonçalves
Prof. Me. Ernane Rosa Martins
Prof. Dr. Fábio Steiner
Prof. Dr. Fabiano dos Santos Souza
Prof. Dr. Gabriel Andres Tafur Gomez
Prof. Dr. Hebert Hernán Soto Gonzáles
Prof. Dr. Hudson do Vale de Oliveira
Prof. Msc. Javier Revilla Armesto
Prof. Msc. João Camilo Sevilla
Prof. Dr. José Luis Soto Gonzales
Prof. Dr. Julio Cezar Uzinski
Prof. Msc. Lucas R. Oliveira
Prof. Dra. Keyla Christina Almeida Portela
Prof. Dr. Leandris Argentele-Martínez
Prof. Msc. Lidiene Jaqueline de Souza Costa Marchesan
Prof. Dr. Marco Aurélio Kistemann
Prof. Msc. Marcos Pisarski Júnior
Prof. Dr. Marcos Pereira dos Santos
Prof. Dr. Mario Rodrigo Esparza Mantilla
Prof. Msc. Mary Jose Almeida Pereira
Prof. Msc. Núbia Flávia Oliveira Mendes
Prof. Msc. Nila Luciana Vilhena Madureira
Prof. Dra. Patrícia Maurer
Prof. Msc. Queila Pahim da Silva
Prof. Dr. Rafael Chapman Auty
Prof. Dr. Rafael Felipe Ratke
Prof. Dr. Raphael Reis da Silva
Prof. Dr. Renato Jaqueto Goes
Prof. Dr. Ricardo Alves de Araújo
Prof. Dra. Sylvana Karla da Silva de Lemos Santos
Prof. Dr. Wéverson Lima Fonseca
Prof. Msc. Wesclen Vilar Nogueira
Prof. Dra. Yilan Fung Boix
Prof. Dr. Willian Douglas Guilherme

Instituição

OAB/PB
Mun. Faxinal Soturno e Tupanciretã
UO (Cuba)
IF SUDESTE MG
Facultad de Medicina (Cuba)
ISCM (Cuba)
UFESSPA
UEA
UNEMAT
UFV
AJES
UFGD
UEMS
IFPA
UNICENTRO
IFMT
UFMG
URCA
ISEPAM-FAETEC
IFG
UEMS
UFF
(Colômbia)
UNAM (Peru)
IFRR
UCG (México)
Mun. Rio de Janeiro
UNMSM (Peru)
UFMT
Mun. de Chap. do Sul
IFPR
Tec-NM (México)
Consultório em Santa Maria
UFJF
UEG
FAQ
UNAM (Peru)
SEDUC/PA
IFB
IFPA
UNIPAMPA
IFB
UO (Cuba)
UFMS
UFPI
UFG
UEMA
IFB
UFPI
FURG
UO (Cuba)
UFT

Conselho Técnico Científico

- Esp. Joacir Mário Zuffo Júnior

- Esp. Maurício Amormino Júnior
- Esp. Tayronne de Almeida Rodrigues
- Lda. Rosalina Eufrausino Lustosa Zuffo

Ficha Catalográfica

Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)

A642 Aplicações de machine learning [livro eletrônico] / Organizadores Ricardo Augusto Manfredini... [et al.]. – Nova Xavantina, MT: Pantanal, 2021. 55 p. : il.

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

ISBN 978-65-81460-20-4

DOI <https://doi.org/10.46420/9786581460204>

1. Inteligência artificial. 2. Redes neurais. 3. Aprendizado de máquina. I. Manfredini, Ricardo Augusto. II. Corrêa, Geraldo Nunes. III. Oliveira, Bruno Rodrigues de. IV. Pauli, Suellen Teixeira Zavadzki de.

CDD 006.3

Elaborado por Maurício Amormino Júnior – CRB6/2422



Nossos e-books são de acesso público e gratuito e seu download e compartilhamento são permitidos, mas solicitamos que sejam dados os devidos créditos à Pantanal Editora e também aos organizadores e autores. Entretanto, não é permitida a utilização dos e-books para fins comerciais, exceto com autorização expressa dos autores com a concordância da Pantanal Editora.

Pantanal Editora

Rua Abaete, 83, Sala B, Centro. CEP: 78690-000.
Nova Xavantina – Mato Grosso – Brasil.
Telefone (66) 99682-4165 (Whatsapp).
<https://www.editorapantanal.com.br>
contato@editorapantanal.com.br

APRESENTAÇÃO

Este livro aborda cinco diferentes contextos em que as técnicas de aprendizado de máquina podem ser utilizadas, servindo como referência prática em diferentes abordagens, tais como: previsão de consumo de energia elétrica, previsão do valor do preço do petróleo, classificação de arritmias cardíacas e método para a seleção automática de artigos.

Estas aplicações são explanadas pelos autores e diferentes técnicas de aprendizado de máquina são utilizadas, dentre elas: redes neurais (Dense, convolucional, recorrentes, Perceptron multicamadas, Elman e Jordan), Naïve Bayes e mineração de textos. Para a aplicação os autores empregam os softwares (línguas de programação) Python e R. Com o objetivo de apresentar aplicações de algumas das técnicas com destaque na atualidade, primeiramente há um descritivo a respeito de cada abordagem e então, são reportados os treinamentos dos modelos de aprendizado de máquina e os resultados. As técnicas incluem classificação, regressão e também mineração de texto. As possíveis abordagens para os problemas relatados não se restringem às técnicas utilizadas, mas a intenção é motivar o leitor a explorar aplicações na área de aprendizado de máquina.

O livro é destinado a profissionais, estudantes, pesquisadores e demais interessados no tema aprendizado de máquina, estatística e áreas relacionadas a fim de colaborar com a explanação de possibilidades de aplicações destas técnicas em contextos diversos. Presume-se que o leitor esteja familiarizado com os conceitos básicos de machine learning, álgebra linear, probabilidade, e análise de algoritmos. A intenção com esta obra é, primordialmente, explicar as possibilidades de aplicação dos algoritmos elencados.

Nos capítulos 1 e 2 as técnicas de redes neurais artificiais são aplicadas para a previsão de preço do valor de petróleo e consumo de energia elétrica, respectivamente. Tais técnicas tratam de modelos computacionais inspirados no sistema nervoso central de um animal. Elas são apresentadas como um sistema de neurônios interconectados, que podem computar valores de entradas, simulando o comportamento de redes neurais biológicas. Na primeira abordagem a autora utiliza as redes Perceptron multicamadas, Elman e Jordan, já na segunda o autor faz uso das redes híbridas Dense, convolucional e recorrente.

No capítulo 3 é tratada a classificação de arritmias cardíacas e, além da técnica de Naïve Bayes também é utilizada a Transformada Wavelet, que é uma transformada integral que utiliza função wavelets que são capazes de decompor determinado sinal (dado) em diferentes escalas. Além desta, também é empregada uma técnica de Ensemble, que encapsula os modelos obtidos por vários algoritmos de aprendizagem a fim de obter uma única previsão global. Por fim, no capítulo 4 é abordado um método para a seleção automática de artigos. Para isto, é utilizado mineração de texto, que trata do processo de obtenção de informações importantes de um texto.

Os organizadores

SUMÁRIO

Apresentação	4
Capítulo I.....	6
Predição diária do preço de petróleo WTI	6
Capítulo II	15
Redes Neurais Artificiais Híbridas Para a Predição de Consumo de Energia Elétrica	15
Capítulo III.....	32
Reconhecimento de padrões de arritmias cardíacas no Eletrocardiograma (ECG) empregando Transformada Wavelet e o classificador Naïve Bayes	32
Capítulo IV	45
Uso da mineração de textos na análise exploratória de artigos científicos.....	45
Índice Remissivo	54
Sobre os organizadores.....	55

Redes Neurais Artificiais Híbridas Para a Predição de Consumo de Energia Elétrica

 10.46420/9786581460204cap2

Ricardo Augusto Manfredini^{1*} 

INTRODUÇÃO

Nas últimas décadas, a população mundial está aumentando rapidamente e, em função desse aumento, a energia global demandada e consumida também está crescendo cada vez mais (LoBrano et al., 2014). Em relação a isso, os prédios residenciais ou comerciais, são identificados como grandes consumidores de energia em todo o mundo, representando cerca de 30% da demanda global de eletricidade relacionada ao consumo de energia no setor residencial (Lusis et al., 2017). Os prédios são responsáveis por uma significativa parte do desperdício de energia também. O desperdício de energia e as mudanças climáticas representam um desafio para a sustentabilidade, sendo crucial tornar os prédios mais eficientes (Marino et al., 2016). Portanto, o desenvolvimento e o uso de produtos limpos e energia renovável em prédios ganhou amplo interesse (Lo Brano et al., 2014). No setor residencial e comercial, os sistemas fotovoltaicos (PV) são a geração distribuída mais comum, minimizando a dependência da demanda das tradicionais usinas de energia e maximizando a autossuficiência das famílias (Lui et al., 2017).

Devido à dependência do PV das condições meteorológicas, a natureza intermitente da energia gerada traz alguma incerteza (Theocharides et al., 2017). Da mesma forma, o consumo de energia elétrica destes prédios também possuem incertezas inerentes à sazonalidade. A maneira mais fácil para gerenciar o risco da energia solar e aproveitar esse poder é prever a quantidade de energia a gerar (Mosaico; Saviozzi, 2019) bem como o consumo. Uma previsão confiável é a chave para várias aplicações de rede inteligente, como despacho, resposta ativa à demanda, regulação da rede e gestão inteligente de energia (Massucco et al., 2019).

O consumo energético de um prédio e a geração PV podem ser representados por uma série temporal com tendências e sazonalidade (Moretti; Tolo, 2006). Existem numerosos estudos de predição sobre séries temporais, desde as clássicas regressões lineares a trabalhos mais recentes utilizando algoritmos de aprendizado de máquina, que são poderosas ferramentas na previsão do consumo de energia elétrica e geração de energia fotovoltaica (Su et al., 2019). Recentemente, muitas técnicas de

¹ IFRS – Instituto Federal de Ciências e Tecnologia do Rio Grande do Sul

* Autor correspondente: ricardo.manfredini@farroupilha.ifrs.edu.br

previsão de energia fotovoltaica foram desenvolvidas, mas ainda não há uma unidade completamodelo de previsão versal e metodologia para garantir a acurácia das predições. Em relação a isso, Redes Neurais Artificiais (RNA) são algoritmos de aprendizado de máquina muito populares para previsão e classificação objetos e são baseados na abordagem clássica da rede neural *feed-forward* (Theocharides et al., 2017). RNAs são sistemas de computação inspirados nas redes neurais biológicas do cérebro, como os neurônios funcionam, passam e armazenam informações (Massucco et al., 2019; Theocharides et al., 2017).

Devido ao desenvolvimento acelerado da tecnologia de computação, a RNA forneceu uma estrutura poderosa para a aprendizagem supervisionada (Liu et al., 2017). O aprendizado profundo permite modelos compostos de várias camadas aprender representações de dados (Marino et al., 2016). Redes Neurais Profundas (DNN²) são inspiradas na estrutura dos sistemas visuais dos mamíferos e elas também são uma importante ferramenta de aprendizado de máquina que tem sido amplamente usado em muitos campos (Yi et al., 2017). DNN emprega uma arquitetura de múltiplas camadas de neurônios em uma RNA e podem representar funções com maior complexidade (Liu et al., 2017).

Este trabalho teve como objetivo a predição do consumo de energia elétrica de um prédio comercial utilização de RNA, nas suas diversas arquiteturas. Foram utilizadas e testadas diversas arquiteturas de RNAs e selecionada uma arquitetura híbrida (Densa, Convolucional e Recorrente), descrita originalmente por Lai et al. (2018) e adaptada para este estudo de caso.

As seções deste capítulo estão organizadas da seguinte forma: inicialmente é feita uma fundamentação teórica dos conceitos, técnicas e ferramentas utilizadas. Posteriormente são definidos os materiais e métodos que serão utilizados para a validação do modelo proposto. Finalmente, são apresentados os resultados do modelo proposto comparando-o com outros modelos.

FUNDAMENTAÇÃO TEÓRICA

Séries Temporais

Séries temporais são conjuntos de observações ordenadas no tempo (Moretti; Toloí, 2006). Uma série temporal pode ser definida como uma classe de fenômenos cujo processo observacional e consequente quantificação numérica geram uma sequência de observações distribuídas ao longo do tempo.

Históricos de consumo de energia elétrica ao longo do tempo são basicamente séries temporais univariadas (Spiegel, 1974) com tendências, ciclos, sazonalidades e aleatoriedades. Tendências são características de longo prazo relacionado com um intervalo de tempo. Ciclos são oscilações a longo prazo, mais ou menos regulares, em torno de uma linha ou curva de tendência. Sazonalidades são padrões

² DNN – do inglês *Deep Neural Network*

regulares observados de tempos em tempos. Finalmente, aleatoriedade são, basicamente, efeitos que ocorrem aleatoriamente e que não podem ser captados pelos ciclos, tendências e sazonalidades.

Desta forma, os modelos de predições de séries temporais mais utilizados na literatura são os de regressões linear e polinomiais. Dentre os modelos de regressões podemos citar o método SARIMAX (SARIMAX, 2021). Este modelo estatístico é uma variante do modelo autorregressivo de médias móveis (ARMA), adicionando derivações para tornar o modelo estacionário (I), adicionando sazonalidade (S) e finalmente adiciona-se o efeito de variáveis exógenas (X) ou aleatórias ao longo do tempo. Neste trabalho, utilizou-se o modelo SARIMAX como linha base de comparação dos seus resultados, da sua aplicação sobre o caso teste e os resultados obtidos de outros modelos de predição.

Redes Neurais Artificiais Convulsionais

Redes Neurais Artificiais Convulsionais (CNN³) são um tipo de DNN que, comumente, é aplicado para analisar imagens. Um dos principais atributos da CNN é conduzir diferentes camadas de processamento que geram uma representação eficaz das características das extremidades das imagens. A arquitetura da CNN permite múltiplas camadas dessas unidades de processamento a serem empilhadas, este modelo de aprendizagem profunda pode enfatizar a relevância de características em diferentes escalas (Yang et al., 2015).

A figura 1 demonstra uma arquitetura típica de uma CNN, composta de pelo menos, uma camada de convolução, uma camada *pooling*, uma camada de *flattening* e camadas densas.

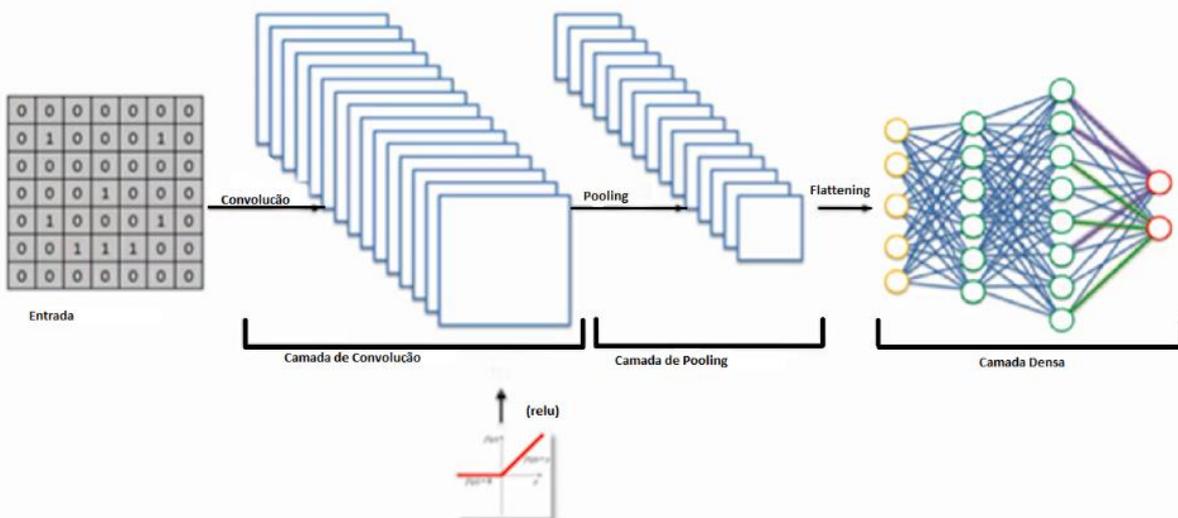


Figura 1. CNN básica. Fonte: O autor.

Na camada de covolução é aplicada um filtro (*kernel*, que também é uma matriz) à matriz de entrada visando a sua redução com a manutenção de suas caratcterística mais importantes. A figura 2

³ CNN – do inglês *Convolutional Neural Network*

representa, passo a passo, a aplicação da função de convolução $g(x, y) = \omega * f(x, y) = \sum_{dx=-a}^a \sum_{dy=-b}^b \omega(dx, dy)f(x + dx, y + dy)$, onde $g(x, y)$ representa o elemento da matriz de convolução, que é o produto matricial da matriz colorido na figura pelo *kernel*, a cada passo desloca-se uma posição à direita até a última coluna da matriz de entrada, após desloca-se uma linha a baixo e continua-se o processo até percorrer toda a matriz de entrada. No exemplo da figura 2 uma matriz de entrada de 7×7 foi reduzida para uma matriz de convolução de 5×5 . Todo o processo representado na figura é repetido para cada um dos *kernels* utilizados, gerando várias matrizes de convolução.

Para a camada de *pooling*, é usual aplicar a função de ativação *relu* $f(x) = \max(0, x)$, por exemplo, gerando uma nova matriz reduzida como demonstra a figura 3.

Finalmente a camada de *flattening* nada mais é que a transformação das matrizes das camadas de *pooling* em vetores, os quais serão as entradas da camada densa.

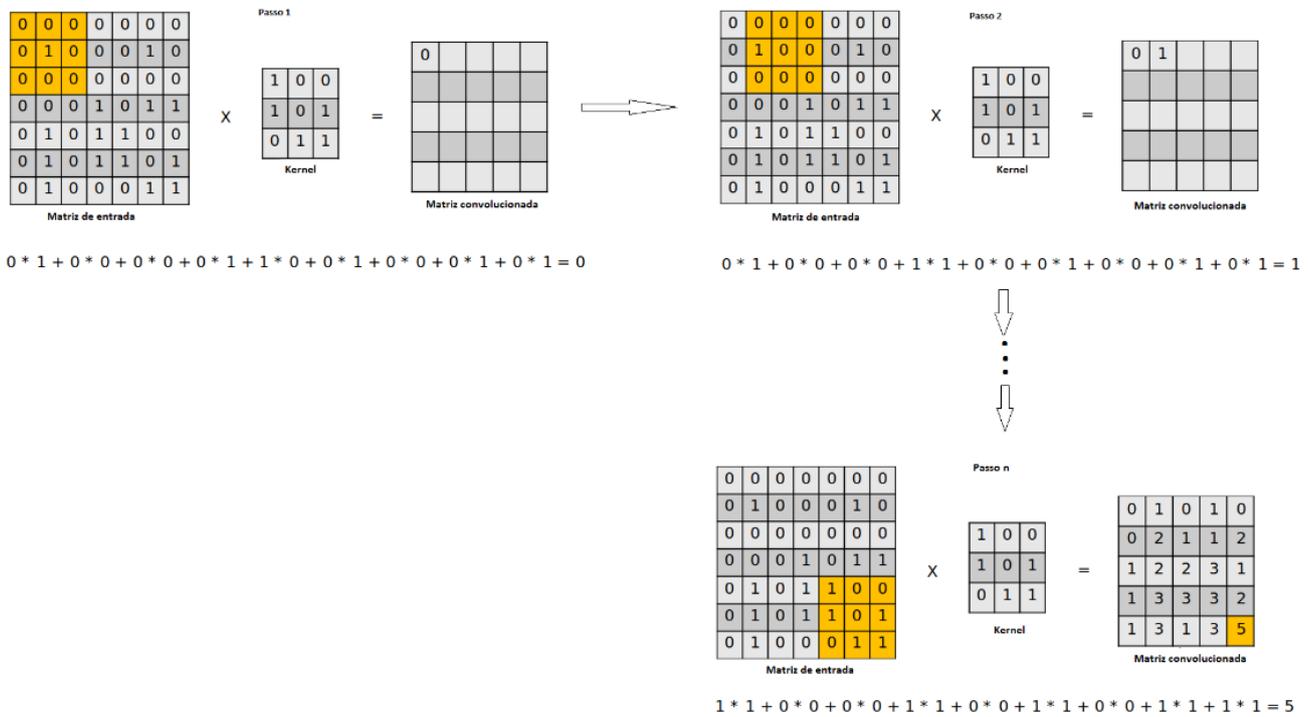


Figura 2. Processo de Convolução. Fonte: O autor.

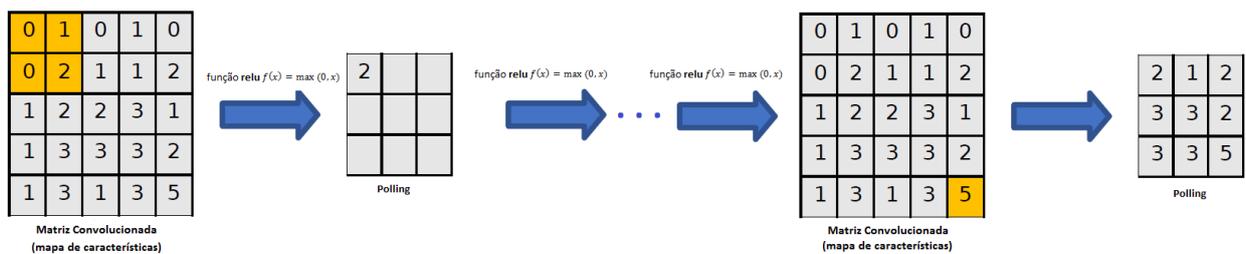


Figura 3. Processo de Pooling. Fonte o Autor.

Redes Neurais Artificiais Recorrentes

Nas RNAs tradicionais, as entradas (e saídas) são independentes umas das outras, dificultando sua utilização, por exemplos, no processamento de linguagem natural onde uma palavra na frase depende de palavras anteriores da mesma frase, ou ainda, em séries temporais que precisamos conhecer os valores ao longo do tempo para melhores projeções.

Em contraposição, as redes neurais artificiais recorrentes (RNN⁴) (Hammer, 2007) armazenam seu estado anterior e o utilizam também como entrada do estado atual para cálculos de novas saídas. Outra forma de pensar sobre as RNNs é que elas possuem uma “memória” que captura informações sobre o que foi calculado até agora. Em teoria, os RNNs podem fazer uso de informações em sequências arbitrariamente longas, mas, na prática, elas se limitam a olhar para trás apenas algumas etapas. A Figura 4 é uma representação típica de uma RNN.

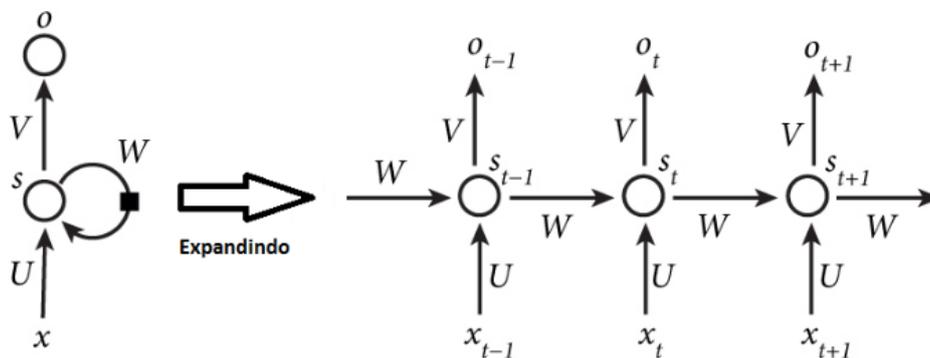


Figura 4. RNN básica. Fonte: O Autor.

A figura 4 mostra uma RNN sendo expandida em uma rede completa. Onde x_t é a entrada na etapa de tempo t . Por exemplo, x_1 poderia ser um vetor *one-hot* correspondente a segunda palavra de uma frase, s_t é o estado oculto na etapa de tempo t . É a “memória” da rede. s_t é calculado com base no estado oculto anterior e a entrada na etapa atual: $s_t = f(Ux_t + Ws_{t-1})$. A função f geralmente é uma não linearidade, como *tanh* ou *relu*. s_{-1} , que é necessário para calcular o primeiro estado oculto, normalmente é inicializado com zeros. o_t é a saída na etapa t . Por exemplo, se quiséssemos prever a próxima palavra em uma frase, seria um vetor de probabilidades em nosso vocabulário. $o_t = \text{softmax}(Vs_t)$. Por expandindo, queremos dizer simplesmente que escrevemos a rede para a sequência completa. Por exemplo, se a sequência que nos interessa é uma frase de 5 palavras, a rede seria desdobrada em uma rede neural de 5 camadas, uma camada para cada palavra.

⁴ RNN – do inglês *Recurrent Neural Network*

MATERIAL E MÉTODOS

Este trabalho foi realizado no Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD5), centro de pesquisa localizado no Instituto Superior de Engenharia do Porto do Instituto Politécnico do Porto ISEP/IPP, Porto, Portugal. Da mesma forma que o modelo HyFIS2 (Josi et al., 2016), o modelo proposto utiliza os dados reais de consumo elétrico de setores do Prédio N do ISEP/IPP onde está localizado o GECAD. O prédio possui cinco medidores de energia que armazenam os dados de consumo de energia elétrica de setores específicos do prédio, com intervalo de tempo de 10 segundos. Estas informações, bem como dados meteorológicos são armazenados num servidor SQL automaticamente, através de agentes desenvolvidos em Java.

Para validação do modelo descrito a seguir, foram realizados testes utilizando os mesmos dados de consumo aplicados ao modelo SARIMAX e o HyFIS2. O setor dos laboratórios do Prédio N não foi computado pois possui grande variação de consumo em razão dos experimentos lá conduzidos que geram muitos *outliers* no histórico de consumo. Para os testes dos experimentos foi realizada uma média horária dos consumos armazenados a cada dez segundo, devido a necessidade de predição da próxima hora de consumo.

O Modelo Long and Short Time series Network Adapted (LSTNetA)

O modelo desenvolvido para a predição de consumo energético utilizou com base o modelo proposto por Lai et al. (2018), representado na Figura 4, constitui-se de uma RNA híbrida, com três camadas distintas, inicialmente possui uma camada convolucional para a extração de padrões de curto prazo da série temporal, tem como entrada a série temporal, a saída desta camada é a entrada da camada recorrente que memoriza informações históricas da série temporal, que por sua vez sua saída é a entrada da camada densa altamente conectada. Finalmente a saída da camada altamente conectada é combinada com a saída da regressão linear autorregressiva (ARMA) (Zhang, 2003) garantindo que a saída terá a mesma escala da entrada, compondo assim a predição.

⁵ <http://www.gecad.isep.ipp.pt/GECAD/Pages/Pubs/PublicationsPES.aspx>

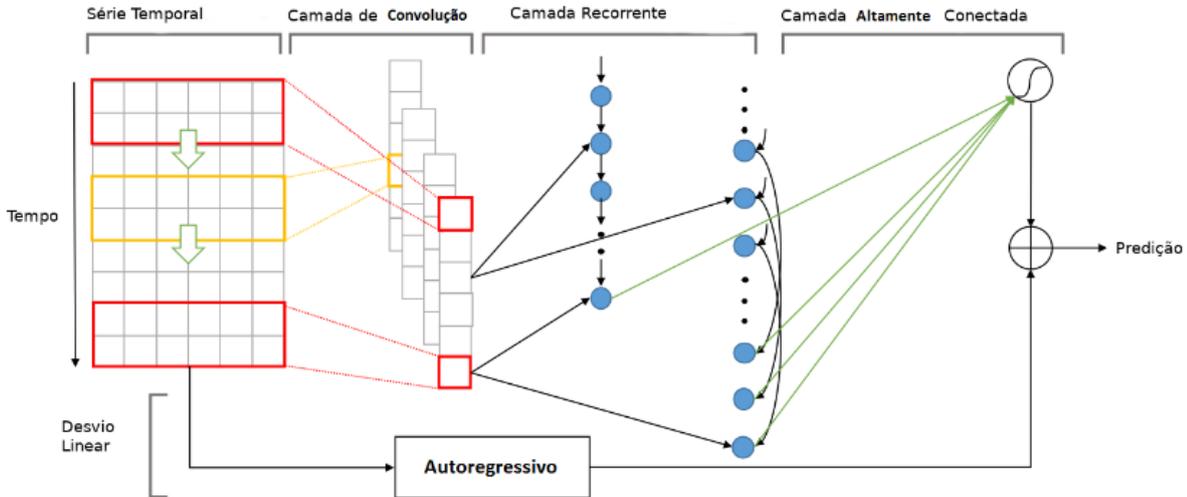


Figura 5. Arquitetura do modelo LSTNetA. Fonte: Adaptado de Lai (Lai et al., 2018).

A Figura 6 sumariza a implementação da rede LSTNetA. A camada de convolução está representada pela classe **Conv2D**, a camada recorrente está representada pelas classes **GRU**, a camada densa está representada pelas classes **Dense**, a autorregressão está representada na classe **PostARTrans**.

É importante salientar que a camada recorrente usa uma das variantes de RNN a GRU (*Gated Recurrent Unit*) (Chung et al., 2014), esse modelo de RNA assim como as LSTM (*Long Short-Term Memory*) visa resolver o problema da memória de curto prazo das RNN que, em séries longas, têm dificuldade de transportar os resultados de etapas anteriores para as posteriores.

This may be caused by multiline strings or comments not indented at the same level as the code.
Model: "model"

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 24, 24)]	0	
reshape (Reshape)	(None, 24, 24, 1)	0	input_1[0][0]
conv2d (Conv2D)	(None, 19, 1, 100)	14500	reshape[0][0]
dropout (Dropout)	(None, 19, 1, 100)	0	conv2d[0][0]
reshape_1 (Reshape)	(None, 19, 100)	0	dropout[0][0]
pre_skip_trans (PreSkipTrans)	(None, 1, 100)	0	reshape_1[0][0]
gru (GRU)	[(None, 100), (None, 60600		reshape_1[0][0]
gru_1 (GRU)	[(None, 5), (None, 5 1605		pre_skip_trans[0][0]
dropout_1 (Dropout)	(None, 100)	0	gru[0][1]
post_skip_trans (PostSkipTrans)	(None, 95)	0	gru_1[0][1] input_1[0][0]
pre_ar_trans (PreARTrans)	(None, 24)	0	input_1[0][0]
concatenate (Concatenate)	(None, 195)	0	dropout_1[0][0] post_skip_trans[0][0]
flatten_1 (Flatten)	(None, 24)	0	pre_ar_trans[0][0]
flatten (Flatten)	(None, 195)	0	concatenate[0][0]
dense_1 (Dense)	(None, 1)	25	flatten_1[0][0]
dense (Dense)	(None, 24)	4704	flatten[0][0]
post_ar_trans (PostARTrans)	(None, 24)	0	dense_1[0][0] input_1[0][0]
add (Add)	(None, 24)	0	dense[0][0] post_ar_trans[0][0]

=====
Total params: 81,434
Trainable params: 81,434
Non-trainable params: 0
=====

Figura 6. Sumário da implementação do LSTNet. Fonte: O Autor.

Na etapa de *backpropagation*, processo de aprendizagem das RNAs, as RNNs sofrem com o problema da dissipação do gradiente (*The Vanishing Gradient Problem*). Gradientes são valores usados para atualizar os pesos das redes neurais. O problema da dissipação do gradiente é quando esses propagados durante o treinamento de uma rede, vão sofrendo multiplicações por valores menores que 1 a cada camada da rede atravessada, chegando nas camadas iniciais da rede com valores ínfimos. Isso faz com que o ajuste dos pesos, calculados a cada iteração do treinamento da rede, sejam também ínfimos, isto onera o treinamento da rede.

Desta forma, nas RNNs as camadas que recebem uma pequena atualização do gradiente param de aprender, com isso as RNNs podem esquecer o que foi visto em sequências mais longas, tendo assim uma memória de curto prazo.

A Figura 7 mostra uma arquitetura típica de uma GRU. Basicamente o que a difere de uma RNN padrão são as portas de descarte (*reset gate*) e de atualização (*update gate*), que através da aplicação das funções da ativação *Sigmoid* e *tanh*, é definido se a saída anterior h_{t-1} será considerada ou descartada para o cálculo da nova saída.

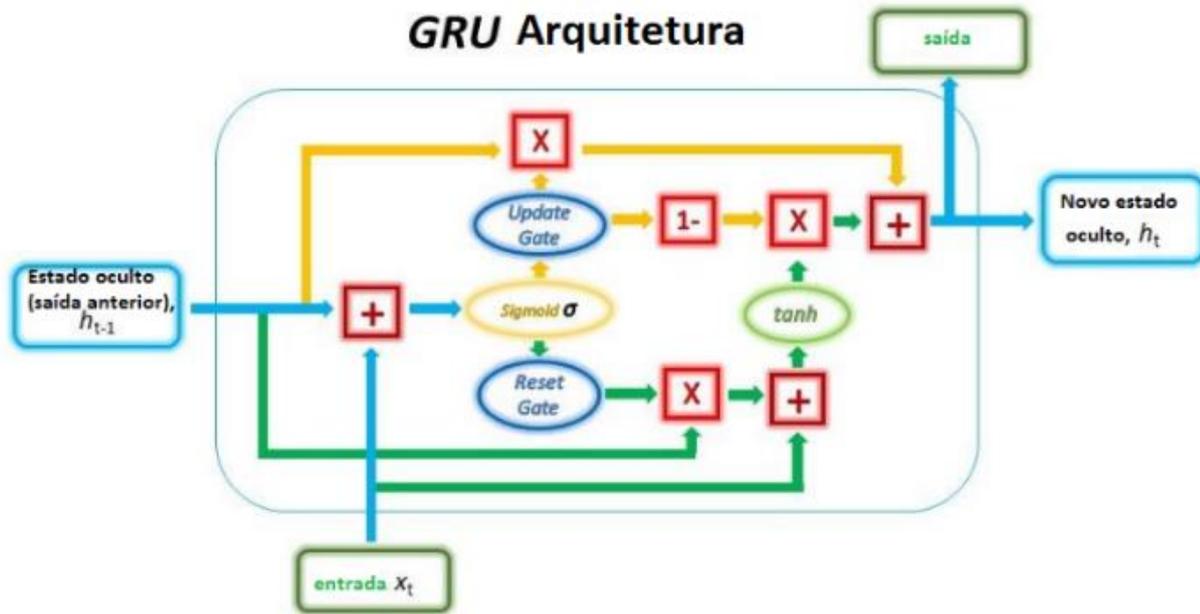


Figura 7. Arquitetura típica de uma GRU. Fonte: O Autor

O modelo LSTMNetA foi desenvolvido na linguagem de programação Python versão 3.7 (Python, 2021) utilizando a biblioteca de aprendizagem de máquina, desenvolvida pelo Google, TensorFlow versão 2.0.

TRABALHOS RELACIONADOS

A Figura 8, representa a série temporal do consumo de energia elétrica utilizada pelo modelo SARIMAX, para treinar e testar o modelo LSTMNetA e o HyFIS2. O gráfico superior representa série histórica do consumo em *watts*, que inicia às zero horas de 08/04/2019 às oito horas de 20/12/2019. O gráfico ao centro mostra a tendência calculada da série e o inferior à sua sazonalidade.

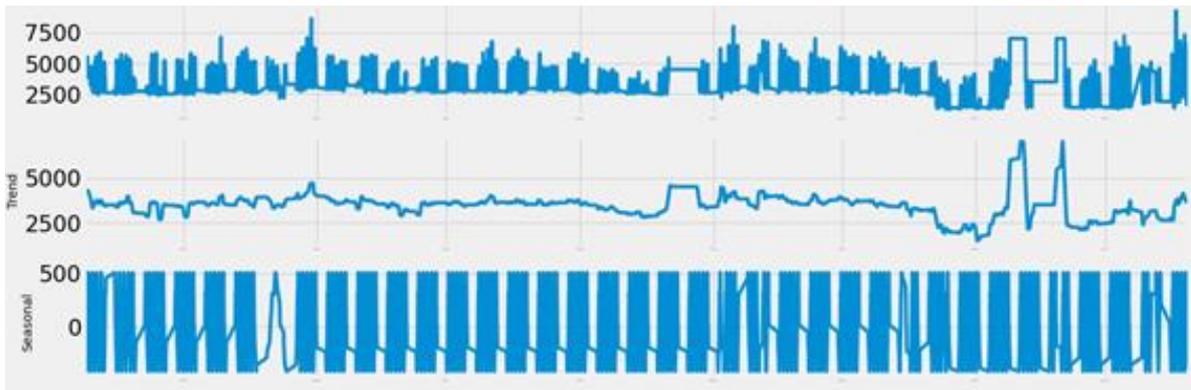


Figura 8. Série histórica de consumo. Fonte: O Autor.

SARIMAX

Como visto anteriormente, o método SARIMAX é um método estatístico de análise de séries temporais, possibilitando a predição através de regressões lineares. Desta forma, não se pode caracterizá-lo como um algoritmo de aprendizagem de máquina. No escopo deste trabalho foi aplicado para obter dados de predição de um modelo amplamente utilizado, obtendo-se resultados para comparação com o modelo proposto e com o modelo HyFIS2.

Para verificar a acurácia de todos os modelos abordados, neste trabalho, foram utilizados os últimos 120 registros, correspondentes a cinco dias de consumo, para comparação entre o consumo real e o predito, demonstrado na Figura 9. Para cálculo do erro utilizado para a verificação dos resultados deste trabalho, em todos os modelos, foi utilizado a raiz do erro quadrático médio (*Root Mean Square Error* – RMSE – Descrito no capítulo 01), demonstrado na Figura 10. A aplicação deste modelo resultou num RMSE médio de 604,72 que foi considerado como acurácia deste modelo, neste trabalho.

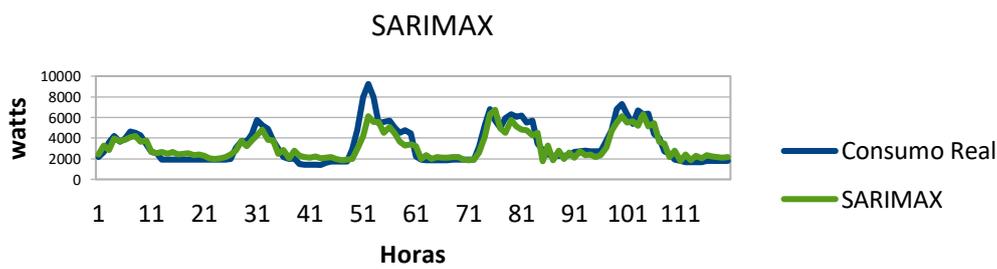


Figura 9. Comparativo Consumo Real X Sarimax. Fonte: O Autor.

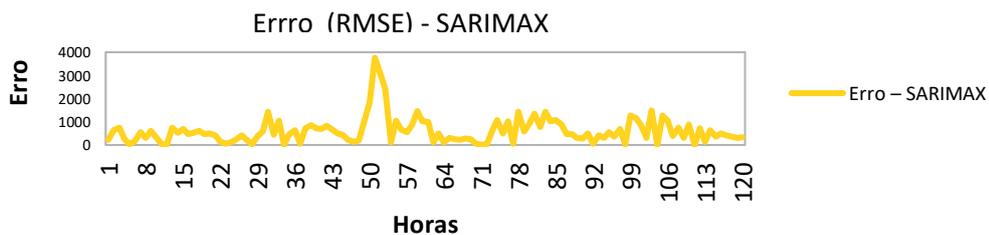


Figura 10. Erros verificados do método SARIMAX. Fonte: O Autor.

Modelo HyFIS2

O modelo HyFIS2 (*Hybrid neural Fuzzy Inference System*) utiliza uma abordagem híbrida com a combinação de RNA densa e lógica difusa (*fuzzy logic*). O sistema inclui cinco camadas, conforme mostrado na Figura 11. Na primeira camada, os nós são as entradas que transmitem os sinais para a próxima camada. Na segunda e na quarta camadas, os nós atuam como funções de pertinência para expressar as variáveis linguísticas difusas de entrada-saída. Nessas camadas, os conjuntos *fuzzy* definidos para as variáveis de entrada-saída são representados como: grande (L), médio (M) e pequeno (S). No entanto, para algumas aplicações, estes podem ser mais específicos e representados como, por exemplo, positivo grande (LP), positivo pequeno (SP), zero (ZE), negativo pequeno (SN) e negativo grande (LN). Na terceira camada, cada nó é um nó de regra e representa uma regra difusa. Os pesos de conexão entre a terceira e a quarta camada representam fatores de certeza das regras associadas, ou seja, cada regra é ativada e controlada pelos valores de peso. Por fim, a quinta camada contém o nó que representa a saída do sistema.

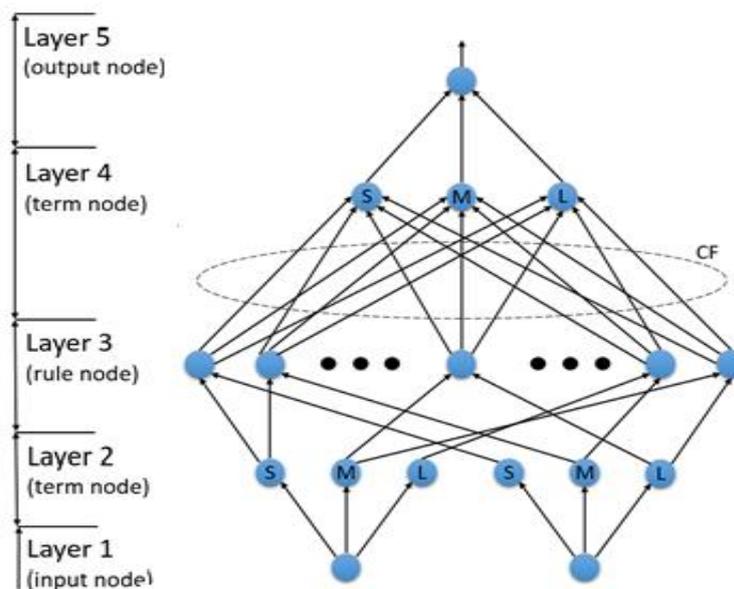


Figura 11. Estrutura Neuro-Fuzzy do modelo HyFIS2. Fonte: Jozi et al. (2016).

Para predição de consumo de eletricidade, como em todos os modelos testados, foram utilizados os últimos 120 registros dos históricos, correspondente a cinco dias de consumo. A comparação entre o consumo real e o predito, é demonstrado na Figura 12. A Figura 13 mostra os erros RMSE apurados. A aplicação deste modelo resultou num RMSE médio de 602,71 que foi considerado como acurácia deste modelo, neste trabalho.

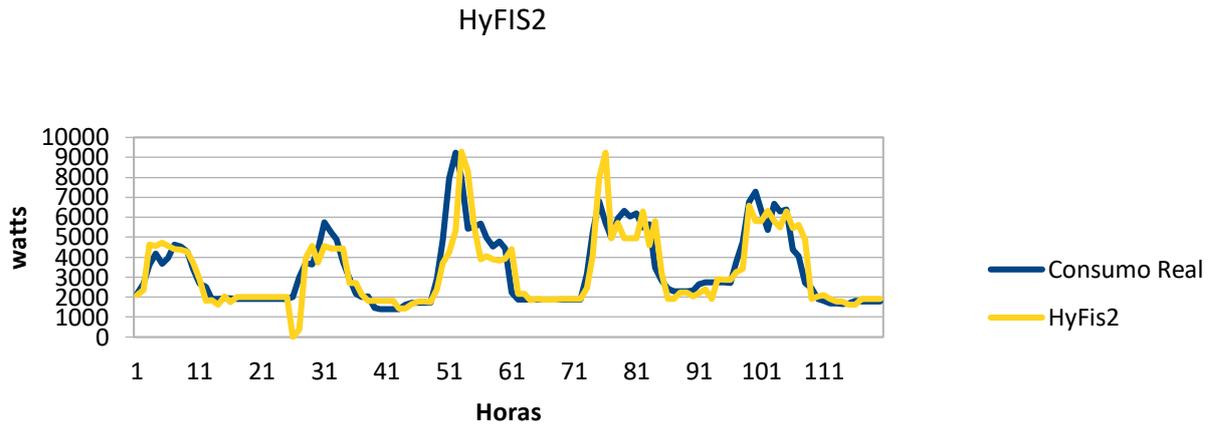


Figura 12. Comparativo Consumo Real X HyFis2. Fonte: O Autor.

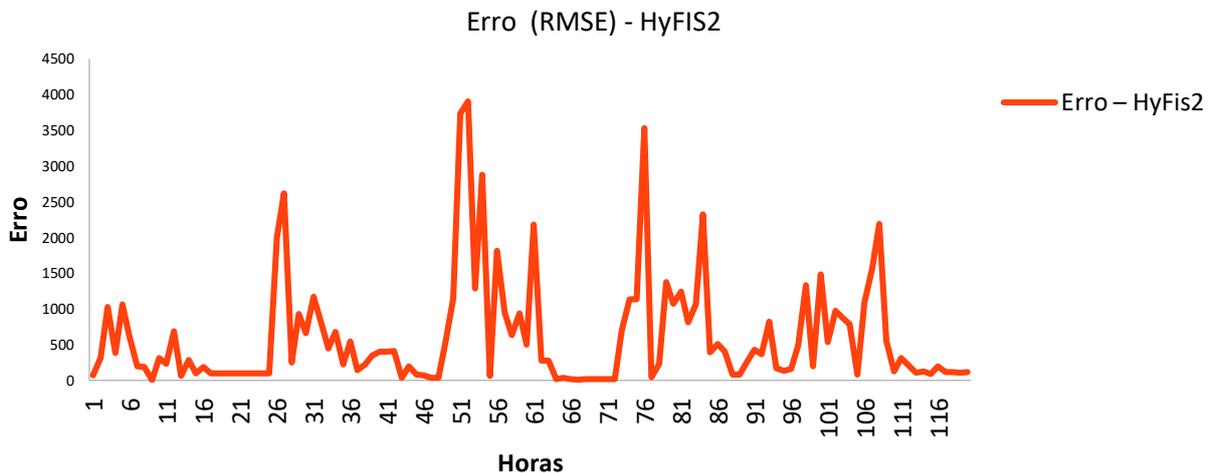


Figura 13. Erros verificados do modelo HyFIS2. Fonte: O Autor.

APLICAÇÃO DO MODELO LSTNETA

O treinamento da RNA LSTNetA foi realizado conforme descrito anteriormente, utilizando-se os dados do consumo real de energia elétrica do prédio N do ISEP/IPP onde está localizado o GECAD, excetuando-se o setor de laboratórios. A série histórica analisada foi das zero horas de 08/04/2019 às oito horas de 20/12/2019, com medições a cada dez segundos, totalizadas a cada hora, resultando 4186 registros, contendo hora e consumo. O treinamento foi realizado com taxa de aprendizagem de 0.0003, utilizando o método estocástico Adam (Kingma; Ba, 2015) de otimização da descida do gradiente para atualização dos pesos no processo de *backpropagation*. Para os pesos iniciais da RNA foi utilizado o algoritmo *VarianceScaling* (He et al., 2010) que gera pesos iniciais com valores na mesma escala das entradas. O kernel de convolução utilizado foi uma matriz identidade 6×6 e foi realizado um loop de

treinamento com 1000 épocas. Todos estes parâmetros foram obtidos de forma experimental e os de melhor resultado, selecionados.

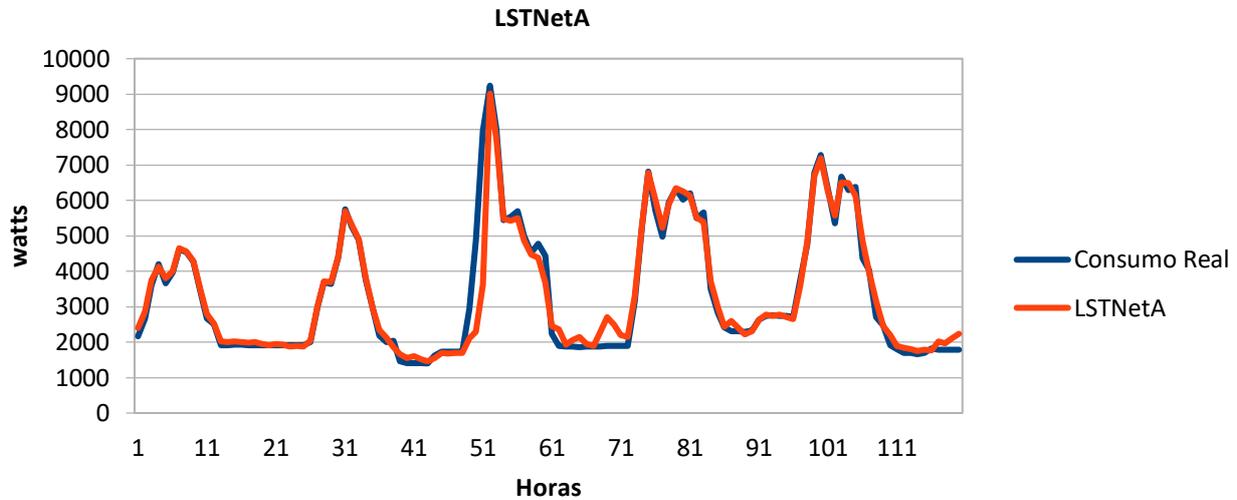


Figura 14. Comparativo Consumo Real X LSTNetA. Fonte: O Autor.

Para predição de consumo de eletricidade, como em todos os modelos testados, foram utilizados os últimos 120 registros dos históricos, correspondente a cinco dias de consumo. A comparação entre o consumo real e o predito, é demonstrado na Figura 14. A Figura 15 mostra os erros RMSE apurados. A aplicação deste modelo resultou num RMSE médio de 198,44 que foi considerado como acurácia deste modelo, neste trabalho.

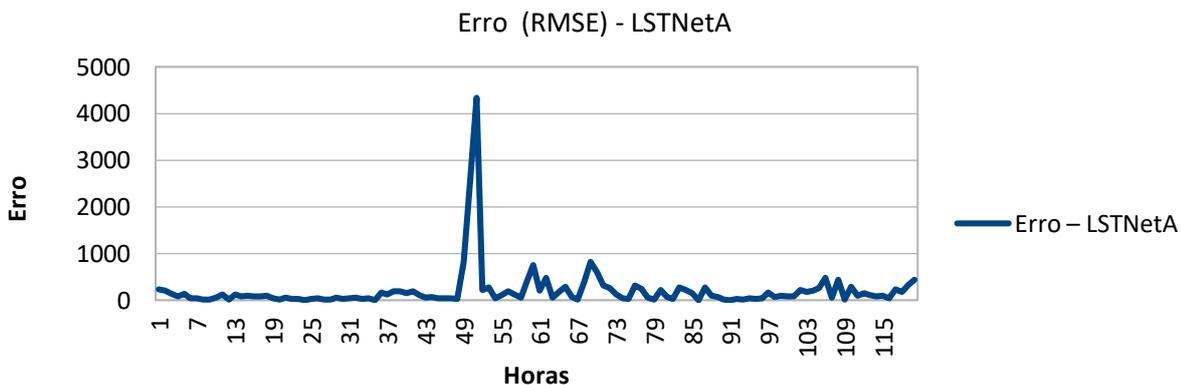


Figura 15. Erros verificados do modelo LSTNetA. Fonte: O Autor.

RESULTADOS E CONSIDERAÇÕES FINAIS

A Tabela 1 mostra um fragmento dos resultados dos três modelos, a coluna *Data e Hora*, a coluna *Real* mostrando o consumo real de eletricidade em *watts* naquela data e hora, a coluna *LSTNetA* a predição deste modelo na data e hora, a coluna *Erro – LSTNetA* o erro absoluto deste modelo na predição, a coluna *HyFIS2* a predição deste modelo na data e hora, a coluna *Erro – HyFIS2* o erro absoluto deste

modelo na predição, finalmente as colunas *SARIMAX* e *Erro – SARIMAX*, representando a predição e o erro absoluto, respectivamente, no modelo SARIMAX.

Comparando os resultados dos modelos SARIMAX, HyFIS2 e LSTNetA, pode-se observar, como demonstrado na Figura 16, que o método LSTNetA, com os dados utilizados para teste, foi o que apresentou as predições mais próximas do consumo real de energia elétrica, onde a linha vermelha, que representa as predições do modelo LSTNetA, em grande parte do período sobrepôs a linha azul que representa o consumo real. Isso demonstra uma predição muito próxima do valor real de consumo, com erros baixos.

Tabela 1. Fragmento de Predições e Erros dos 3 Modelos.

Data e Hora	Consumo Real	LSTNetA	Erro – LSTNetA	HyFis2	Erro – HyFis2	SARIMAX	Erro – SARIMAX
19/12/2019 09:00	4759,38	4824,27	64,8900	3427,13	1332,2500	4721,76	37,6190
19/12/2019 10:00	6781,51	6685,28	96,2346	6583,38	198,1300	5516,26	1265,2476
19/12/2019 11:00	7279,1	7194,26	84,8373	5798,56	1480,5400	6124,20	1154,8976
19/12/2019 12:00	6332,88	6247,08	85,8038	5798,38	534,5000	5497,10	835,7849
19/12/2019 13:00	5350,34	5569,95	219,6063	6322,98	972,6400	5653,27	302,9276
19/12/2019 14:00	6677,56	6499,50	178,0639	5798,37	879,1900	5197,56	1479,9983

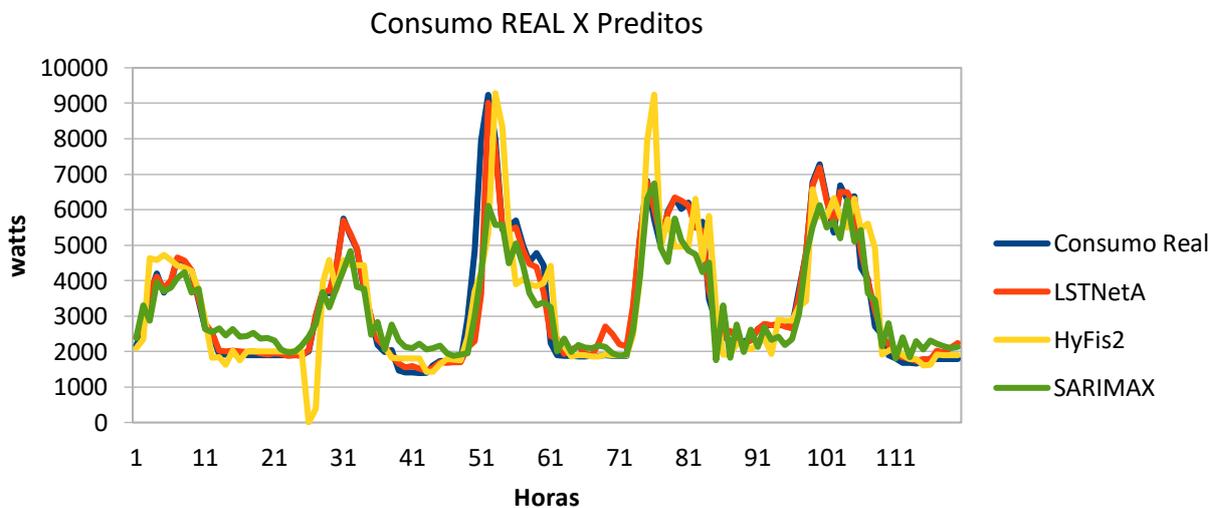


Figura 16. Comparativo Consumo Real X Modelos de Predição. Fonte: O Autor.

A Figura 17 representa os erros (RSME) dos três modelos, permitindo comparar a assertividade das predições de cada um dos métodos e, ainda, concluir que o método LSTNetA apresentou uma melhor eficácia em suas predições em comparação aos métodos SARIMAX e HyFIS2. Esta afirmação pode ser corroborada com os dados apresentados na Tabela 2, onde o erro total médio do modelo LSTNetA é significativamente menor que os demais modelos.

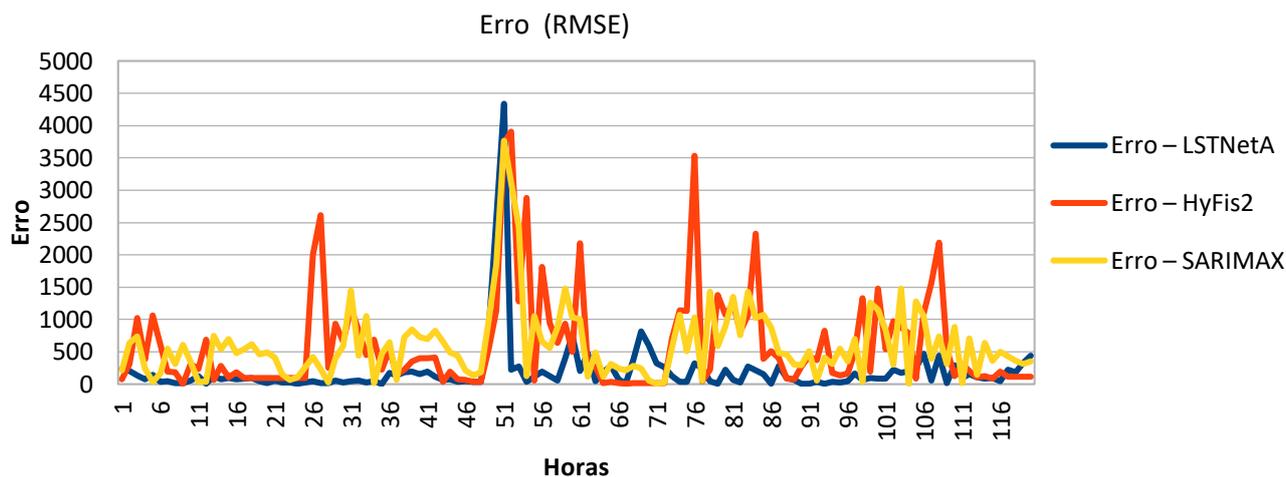


Figura 17. Comparativos de erros verificados em todos os modelos. Fonte: O Autor.

Tabela 2. RSME dos 3 Modelos Testados.

	Erro – LSTNetA	Erro – HyFis2	Erro – SARIMAX
RSME	198,4496	602,7109	604,5810

REFERÊNCIAS BIBLIOGRÁFICAS

- Chung J et al. (2014). Empirical evaluation of gated recurrent neural network on sequence modeling. in NIPS 2014 Workshop on Deep Learning, December 2014.
- Das UK et al. (2018). Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, 81: 912–928.
- He K et al. (2015). Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification.
- Lai G et al. (2018). Modeling long- and short-term temporal patterns with deep neural networks. 41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018, 95–104.
- Liu W et al. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234: 11–26.
- Lo Brano V et al. (2014). Artificial neural networks to predict the power output of a PV panel. *International Journal of Photoenergy*.
- Lusis P et al. (2017). Short-term residential load forecasting: Impact of calendar effects and forecast granularity. *Applied Energy*, 205: 654–669.
- Hammer B (2007). *Learning with Recurrent Neural Networks*. London: Springer.

- Jozi A et al. (2016). Energy Consumption Forecasting based on Hybrid Neural Fuzzy Inference System. 2016 IEEE Symposium Series on Computational Intelligence
- Kingma DP, Ba J (2016). Adam: A Method for Stochastic Optimization. Computer Science, Mathematics ICLR 2015. (SSCI). 6-9. Disponível em: <https://arxiv.org/abs/1412.6980.pdf>. Acessado em: 01/03/2021.
- Marino DL et al. (2016). Building energy load forecasting using Deep Neural Networks. IECON Proceedings (Industrial Electronics Conference), 7046–7051.
- Massucco S et al. (2019). A hybrid technique for day-ahead pv generation forecasting using clear-sky models or ensemble of artificial neural networks according to a decision tree approach. Energies, 12(7).
- Montavon G et al. (2018). Methods for interpreting and understanding deep neural networks. Digital Signal Processing: A Review Journal, 73: 1–15.
- Morete PA, Toloi CMC (2006). Análise de séries temporais. 2. ed. São Paulo: Egard Blucher.
- Mosaico G, Saviozzi M (2019). A hybrid methodology for the day-ahead pv forecasting exploiting a clear sky model or artificial neural networks. In IEEE EUROCON 2019 -18th International Conference on Smart Technologies, 1–6.
- Pelletier C et al. (2021). Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. MDPI - Remote Sensinf - Open Access Journal. Disponível em: <https://arxiv.org/pdf/1811.10166.pdf> Acessado em: 01/03/2021.
- Python (2021). Python is a programming language that lets you work quickly and integrate systems more effectively. Disponível em: <https://www.python.org>. Acessado em: 01/03/2021
- Reddy KS, Ranjan M (2003). Solar resource estimation using artificial neural networks and comparison with other correlation models. Energy Conversion and Management, 44(15): 2519–2530.
- SARIMAX (2021). SARIMAX: Introduction. Disponível em: https://www.statsmodels.org/dev/examples/notebooks/generated/statespace_sarimax_stata.html. Acessado em: 01/03/2021
- Spiegel MR (1974). Estatística. Brasília: McGraw-Hill do Brasil.
- Su D et al. (2019). Machine learning algorithms in forecasting of photovoltaic power generation. In 2019 International Conference on Smart Energy Systems and Technologies (SEST), 1–6.
- TensorFlow (2021). Uma plataforma completa de código aberto para machine learning. Disponível em: <https://www.tensorflow.org>. Acessado em: 01/03/2021
- Theocharides S et al. (2017). Pv Production Forecasting Model Based on Artificial Neural Networks (Ann). 33rd European Photovoltaic Solar Energy Conference, 1830 – 1894.
- Yang J et al. (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. in Ijcai, 15: 3995–4001.

Yi H et al. (2017). A study on Deep Neural Networks framework. Proceedings of 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, IMCEC 2016, 1519–1522.

Zhang GP (2003). Time series forecasting using a hybrid arima and neural network model. Neurocomputing, 50:159–175.

ÍNDICE REMISSIVO

A

algoritmo, 13, 25, 27, 42, 51
aprendizado de máquina, 4, 9, 13, 16, 17, 35,
37, 43
arritmia, 41

C

classificação, 4, 17, 37, 39
conhecimento, 46, 47, 50, 53
convolucional
rede neural, 4, 21

D

Dense
rede neural, 4, 22

E

Elman
rede neural, 4, 8, 9, 10, 12, 14, 15
energia elétrica, 4, 8, 16, 17, 21, 24, 27, 29
ensemble, 31

J

Jordan
rede neural, 4, 8, 9, 10, 12, 14, 15

M

mineração de texto, 4, 48

P

padrão, 24, 33, 37, 38, 41, 42
petróleo, 4, 7, 8, 10
predição, 4, 16, 17, 18, 21, 25, 26, 28, 29, 38, 43
pré-processamento, 43, 51, 52
Pubmed, 48, 49, 51

R

recorrente
rede neural, 4, 8, 21, 22
rede neural, 7, 8, 9, 17, 20

S

série temporal, 7, 8, 16, 17, 21, 24

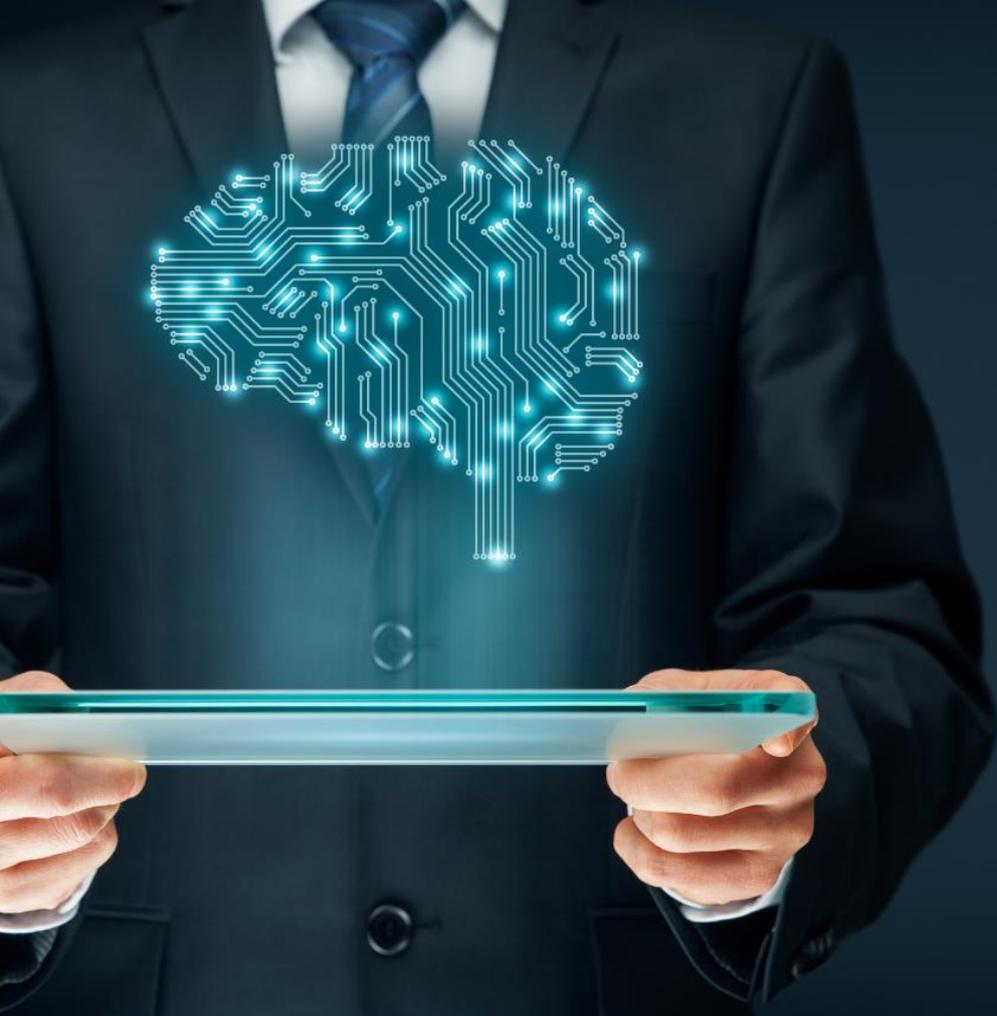
SOBRE OS ORGANIZADORES

 **Ricardo Augusto Manfredini** Possui graduação em Bacharel em Ciências da Computação pela Universidade de Caxias do Sul (1990), mestrado em Ciências da Computação pela Universidade Federal do Rio Grande do Sul (2001) , doutorado pelo Instituto de Biotecnologia da Universidade de Caxias do Sul(2015) e pós-doutorado no GECAD do Instituto Superior de Engenharia do Porto em Inteligência Artificial e IoT, também é professor do Instituto Federal de Ciências e Tecnologia do Rio Grande do Sul - campus Farroupilha. Tem experiência na área de Ciência da Computação, com ênfase em Programação, Engenharia de Software e Tolerância a Falhas, atuando principalmente nos seguintes temas: IoT & IA (2020), bioinformática, injeção de falhas, sistemas distribuídos, tolerância a falhas, metodologias de desenvolvimento de sistemas e linguagens de programação.

 **Geraldo Nunes Corrêa** Possui graduação em Ciência da Computação pela Universidade de São Paulo (1991), mestrado em Ciências da Computação (Área de Inteligência Artificial, Banco de Dados e Manufatura) pela Universidade de São Paulo (1994) e doutorado em Engenharia Mecânica (Departamento de Engenharia de Produção) pela Universidade de São Paulo (1999). Pós doutorado em Mineração de Textos no Instituto de Ciências Matemáticas e Computação da Universidade de São Paulo (2013). Consultor em soluções educacionais. Mentor de Startups. Empreendedor Digital.

 **Bruno Rodrigues de Oliveira** Graduado em Matemática pela Universidade Estadual de Mato Grosso do Sul (UEMS, 2008). Mestrado (2015) e Doutorado (2020) em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP, 2015). Atualmente é Analista Judiciário no Tribunal de Justiça de Mato Grosso do Sul e professor de Matemática no Colégio Maper e Editor na Pantanal Editora. Tem experiência nos temas: Matemática, Processamento de Sinais via Transformada Wavelet, Análise Hierárquica de Processos, Teoria de Aprendizagem de Máquina e Inteligência Artificial.

 **Suellen Teixeira Zavadzki de Pauli** Doutoranda em Métodos Numéricos na Universidade Federal do Paraná, possui mestrado em Engenharia da Produção pela Universidade Federal do Paraná (2020), pós-graduação em Engenharia da Confiabilidade pela Universidade Tecnológica Federal do Paraná (2016) e graduação em Estatística pela Universidade Federal do Paraná (2013). Possui experiência com monitoramento de modelos estatísticos de crédito e análise de indicadores. Atuação no planejamento e estratégia de campanhas de marketing através da análise de produtos e de perfil de clientes, dashboards analíticos e KPIs. Geração e implementação de simulador de oferta ideal para o cliente (next best offer). Experiência na avaliação de dados e cenários, utilização de técnicas estatísticas (análises descritivas, árvores de decisão, testes estatísticos). Experiência como professor substituto na Universidade Tecnológica Federal do Paraná (2020). Experiência com pesquisa em redes neurais artificiais aplicadas em dados de séries temporais (preço de petróleo e valor de ações), modelos mistos, análise de variância, entre outras técnicas. Familiaridade com linguagem SQL e ferramentas de análises como SAS, SPSS, R, Phyton e pacotes office. Coorganizadora do RLadies Curitiba.



Pantanal Editora

Rua Abaete, 83, Sala B, Centro. CEP: 78690-000

Nova Xavantina – Mato Grosso – Brasil

Telefone (66) 99682-4165 (Whatsapp)

<https://www.editorapantanal.com.br>

contato@editorapantanal.com.br