



Aplicação de Técnicas de Ciência de Dados na Previsão de Consumos Energéticos

Mestrado em Ciência de Dados

José Eduardo Figueiredo Costa Simões de Sá

Leiria, setembro de 2023



Aplicação de Técnicas de Ciência de Dados na Previsão de Consumos Energéticos

Mestrado em Ciência de Dados

José Eduardo Figueiredo Costa Simões de Sá

Projeto de Mestrado realizado sob a orientação do Professor Doutor Carlos Grilo, do Professor Doutor Fernando Sebastião e do Professor Doutor Rolando Miragaia.

Leiria, setembro de 2023

Originalidade e Direitos de Autor

O presente relatório de projeto é original, elaborado unicamente para este fim, tendo sido devidamente citados todos os autores cujos estudos e publicações contribuíram para o elaborar.

Reproduções parciais deste documento serão autorizadas na condição de que seja mencionado o Autor e feita referência ao ciclo de estudos no âmbito do qual o mesmo foi realizado, a saber, Curso de Mestrado em Ciência de Dados, no ano letivo 2022/2023, da Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria, Portugal, e, bem assim, à data das provas públicas que visaram a avaliação destes trabalhos.

Agradecimentos

O projeto aqui apresentado não teria sido possível sem o precioso apoio e envolvimento de algumas pessoas que deram o seu contributo para que este trabalho fosse concretizado.

Deixo aqui o meu agradecimento.

Agradeço aos meus orientadores, o Professor Doutor Carlos Grilo, o Professor Doutor Fernando Sebastião e o Professor Doutor Rolando Miragaia pela ajuda e pelos conhecimentos transmitidos, não só durante a realização do projeto, como também no decorrer do ano letivo. A total colaboração na transmissão do saber, na solução de dúvidas e problemas que foram surgindo, conduziu a um crescendo de conhecimentos imprescindíveis à minha formação.

Quero agradecer a toda a comunidade que constitui o CENFIM, especialmente ao Dr. Ângelo Ribeiro, à Dra. Lurdes Alves e ao Eng. Jorge Rodrigues, por terem tornado possível a realização deste projeto através da cedência da informação necessária e por todo o apoio prestado durante este ano letivo.

Agradeço ao IPMA – Instituto Português do Mar e da Atmosfera pela cedência dos dados climáticos da região.

Agradeço aos meus pais por me terem proporcionado a possibilidade de frequentar este mestrado e pelo apoio e incentivo incondicionais que sempre demonstraram.

Por último, aos meus restantes familiares, aos meus amigos e à Rita estou grato pela ajuda e paciência incondicionais demonstradas nos períodos mais conturbados deste projeto.

A todos, muito obrigado!

Resumo

A adoção das melhores práticas que visem a utilização eficiente de recursos promove o desenvolvimento sustentável o que, por sua vez, se traduz em benefícios sociais e económicos. O presente projeto foca-se nos consumos energéticos de um edifício de serviços, nomeadamente o Núcleo de formação do CENFIM da Marinha Grande, com o objetivo de utilizar ferramentas de ciência de dados para realizar previsões diárias de consumos energéticos. O edifício em estudo promove cursos nas áreas da metalurgia e da metalomecânica, estando equipado com máquinas adequadas para o efeito. Tendo em conta que os consumos energéticos medidos ao longo do tempo podem ser estudados através de séries temporais, este projeto reporta as metodologias e os modelos utilizados usualmente para a sua análise, dividindo-os em dois grupos, os modelos de base estatística, onde se destacam os modelos SARIMA e SARIMAX, e os modelos de aprendizagem computacional, com maior ênfase em redes neuronais do tipo MLP, RNN, LSTM e GRU. O problema foi abordado de várias formas, nomeadamente na definição da quantidade e forma de dias anteriores para realizar a previsão do consumo do dia seguinte, assim como na utilização de variáveis exógenas para melhorar o resultado da previsão, como é o caso da ocupação do edifício. Foi possível obter uma previsão com um MAPE de cerca de 12,5% utilizando o modelo SARIMAX com a ocupação total do edifício por dia como variável exógena. Com as redes neuronais foi apenas utilizada a variável referente ao consumo energético, contudo, várias arquiteturas, tipologias e hiperparâmetros foram testados para cada rede. A rede que apresenta melhor resultado para o problema em questão é a GRU com um MAPE de cerca de 14,5%, embora a MLP apresente um resultado bastante semelhante.

Palavras-chave: consumos energéticos, séries temporais, redes neuronais, previsão, ciência de dados, análise de dados

Abstract

The adoption of best practices aimed at the efficient use of resources promotes sustainable development which, in turn, translates into social and economic benefits. This project focuses on the energy consumption of a service building, namely the CENFIM training center in Marinha Grande, with the aim of using data science tools to make daily forecasts of energy consumption. The building under study promotes courses in the areas of metallurgy and metalworking, and is equipped with suitable machines for this purpose. Taking into account that energy consumption measured over time can be studied through time series, this project reports the methodologies and models usually used for its analysis, dividing them into two groups, the statistical-based models, where highlight the SARIMA and SARIMAX models, and computational learning models, with greater emphasis on neural networks such as MLP, RNN, LSTM and GRU. The problem was approached in several ways, namely in defining the quantity and form of previous days to forecast the next day's consumption, as well as in the use of exogenous variables to improve the forecast result, as is the case with building occupancy. It was possible to obtain a forecast with a MAPE of around 12,5% using the SARIMAX model with the total occupancy of the building per day as an exogenous variable. With neural networks, only the variable referring to energy consumption was used, however, several architectures, typologies and hyperparameters were tested for each network. The network that presents the best result for the problem in question is GRU with a MAPE of around 14,5%, although MLP presents a very similar result.

Keywords: energy consumption, time series, neural networks, forecasting, data science, data analysis

Índice

Originalidade e Direitos de Autor	iii
Agradecimentos	iv
Resumo	v
Abstract	vi
Lista de Figuras	ix
Lista de Tabelas	xi
Lista de Siglas e Acrónimos	xii
1. Introdução	1
1.1. Enquadramento	1
1.2. Breve Descrição da Entidade.....	3
1.3. Objetivos.....	4
1.4. Estrutura do Trabalho	4
2. Matéria Relacionada	7
2.1. Conceitos Gerais de Séries Temporais.....	7
2.2. Conceitos Gerais de Aprendizagem Computacional.....	14
3. Estado da Arte	23
4. Metodologia, Modelos e Métricas	35
4.1. Metodologia.....	35
4.2. Modelos de Base Estatística	36
4.2.1. SARIMA.....	36
4.2.2. SARIMAX.....	37
4.3. Modelos de Aprendizagem Computacional	38
4.3.1. RNN.....	39

4.3.2.	LSTM	40
4.3.3.	GRU	41
4.4.	Métricas de Erro.....	42
5.	Descrição e Tratamento dos Dados.....	45
5.1.	Descrição dos Dados.....	45
5.1.1.	Origem e Formato dos Dados.....	45
5.2.	Tratamento dos Dados.....	46
5.2.1.	Dados de Energia	48
5.2.2.	Dados da Ocupação	51
5.2.3.	Dados da Temperatura	52
6.	Criação dos Modelos e Análise dos Resultados	55
6.1.	Modelos de Base Estatística.....	55
6.1.1.	Modelo SARIMA	57
6.1.2.	Modelo SARIMAX – com uma variável exógena.....	59
6.1.3.	Modelo SARIMAX – com quatro variáveis exógenas.....	61
6.2.	Modelos de Aprendizagem Computacional	63
6.2.1.	Rede Neuronal – MLP.....	67
6.2.2.	Rede Neuronal – RNN	69
6.2.3.	Rede Neuronal – LSTM	70
6.2.4.	Rede Neuronal – GRU	72
6.3.	Análise de Resultados.....	74
6.4.	Caso de Retroanálise	75
7.	Conclusão	79
7.1.	Trabalho Futuro.....	80
Bibliografia		82
Anexos		87

Lista de Figuras

Figura 1 – Evolução do consumo total de energia primária (Esquerda), Evolução do consumo de energia final por setor de atividade (Direita), ambos referentes a Portugal. Fonte: DGEG	2
Figura 2 – Produção trimestral de tijolos na Austrália	9
Figura 3 – Esquerda: Gráfico de autocorrelação. Direita: Gráfico de autocorrelação parcial. Ambos referentes à série temporal de produção de tijolos na Austrália	12
Figura 4 – Mapa de algoritmos de Aprendizagem Computacional [17].....	15
Figura 5 – Exemplo de uma rede neuronal artificial [20].....	18
Figura 6 – Entradas e saídas de um neurónio artificial de uma rede neuronal	19
Figura 7 – Exemplo de uma rede neuronal para determinar o género com base no peso e na altura	20
Figura 8 – Tipos de relação entre entradas e saídas de uma rede neuronal	21
Figura 9 – Artigos publicados para previsão de consumos energéticos, 2005-2019, [32].....	25
Figura 10 - Metodologia no tratamento e análise dos dados	35
Figura 11 – Representação das diferentes ordens de um modelo SARIMA [50]	37
Figura 12 – Exemplo de uma rede neuronal artificial [51].....	38
Figura 13 – Esquerda: Representação das entradas e saídas dos neurónios numa RNN. Direita: Representação das entradas e saídas dos neurónios numa MLP [52].....	39
Figura 14 – Exemplo de um neurónio numa RNN [53]	40
Figura 15 - Exemplo de um neurónio numa LSTM [54].....	41
Figura 16 - Exemplo de um neurónio numa GRU [54].....	42
Figura 17 – Consumo de energia ativa durante o período de 01/10/2015 a 30/09/2022	48
Figura 18 – Consumo de energia ativa diário e agregado por semana e por mês.....	49
Figura 19 - Consumo de energia ativa por ano letivo.....	49
Figura 20 – Diagramas de extremos e quartis da energia ativa por ano e por mês.....	50
Figura 21 – Ocupação total do edifício durante o período de 01/10/2015 a 30/09/2022.....	51
Figura 22 - Ocupação total do edifício por ano letivo	52
Figura 23 – Valores de temperatura máxima e mínima total durante o período de 01/10/2015 a 30/09/2022	53
Figura 24 -Mapa de correlação das variáveis	53
Figura 25 - Série temporal e médias móveis	55
Figura 26 - Decomposição de série temporal através do modelo aditivo e com um período de 7.....	56

Figura 27 - Gráfico de autocorrelação e de autocorrelação parcial.....	57
Figura 28 - Valores reais vs previsões para o modelo SARIMA	59
Figura 29 - Valores reais vs previsões para o modelo SARIMAX – com uma variável exógena.....	60
Figura 30 - Valores reais vs previsões para o modelo SARIMAX – com quatro variáveis exógenas	62
Figura 31 – Divisão conjunto treino, validação e teste	64
Figura 32 - Progresso de treino da MLP	68
Figura 33 - Previsões vs valores reais da MLP	69
Figura 34 - Progresso de treino da RNN	70
Figura 35 - Previsões vs valores reais da RNN.....	70
Figura 36 - Progresso de treino da LSTM.....	71
Figura 37 - Previsões vs valores reais da LSTM.....	72
Figura 38 - Progresso de treino da GRU	73
Figura 39 - Previsões vs valores reais da GRU	73
Figura 40 – Previsões vs valores reais da RNN com os valores horários agregados por dia	77

Lista de Tabelas

Tabela 1 – Modelos vs Autocorrelação e Autocorrelação parcial	13
Tabela 2 - Variáveis do <i>Dataset</i>	48
Tabela 3 - Teste de hipóteses para averiguar a estacionariedade.....	56
Tabela 4 - Critérios de Informação para o modelo SARIMA.....	57
Tabela 5 – Métricas de Erro para o modelo SARIMA	59
Tabela 6 - Critérios de Informação para o modelo SARIMAX – com uma variável exógena	59
Tabela 7 - Métricas de Erro para o modelo SARIMAX – com uma variável exógena.....	60
Tabela 8 - Critérios de Informação para o modelo SARIMAX – com quatro variáveis exógenas.....	61
Tabela 9 - Métricas de Erro para o modelo SARIMAX – com quatro variáveis exógenas.....	61
Tabela 10 - Arquiteturas e hiperparâmetros testados com o Optuna	65
Tabela 11 - Resultados <i>Optuna</i> - MLP.....	67
Tabela 12 - Resultados MLP	67
Tabela 13 - Resultados <i>Optuna</i> - RNN.....	69
Tabela 14 - Resultados RNN.....	69
Tabela 15 - Resultados <i>Optuna</i> - LSTM.....	71
Tabela 16 - Resultados LSTM.....	71
Tabela 17 - Resultados <i>Optuna</i> - GRU.....	72
Tabela 18 - Resultados GRU	72
Tabela 19 – Resumo dos resultados dos modelos utilizados para efetuar previsões	74
Tabela 20 - Resultados RNN para valores horários.....	76

Lista de Siglas e Acrónimos

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller Test
AIC	Akaike Information Criterion
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
CAD	Computer-Aided Design
CAM	Computer-Aided Manufacturing
CENFIM	Centro de Formação Profissional da Indústria Metalúrgica e Metalomecânica
CNC	Computer Numerical Control
CSV	Comma-Separated Values
GPU	Graphics Processing Unit
GRU	Gated Recurrent Unit
IoT	Internet of Things
IPMA	Instituto Português do Mar e da Atmosfera
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLP	Multilayer Perceptron
MSE	Mean Squared Error

PACF	Partial Autocorrelation Function
PIB	Produto Interno Bruto
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SARIMA	Seasonal Autoregressive Integrated Moving-Average
SARIMAX	Seasonal Autoregressive Integrated Moving-Average with Exogenous Variables
STL	Seasonal-Trend decomposition using LOESS
VBA	Visual Basic for Applications

1. Introdução

A escassez de recursos e as preocupações ambientais são temas comuns hoje em dia. Naturalmente, a utilização de recursos energéticos está correlacionada com o meio ambiente, refletindo o facto de que, a utilização dos primeiros, de uma forma geral, prejudica o segundo. Em todo o caso, o progresso das sociedades encontra-se ligado à sua capacidade de produção, o que significa que, quanto maior o progresso, maior a quantidade de energia consumida. Contudo, a existência de políticas relacionadas com o meio ambiente e a energia têm vindo a ser atualizadas e adaptadas à realidade com que nos deparamos, num esforço de alcançarmos um desenvolvimento mais sustentável. O presente relatório foca-se na energia elétrica utilizada por um edifício e na aplicação de técnicas de Ciência de Dados para realizar previsões de consumos energéticos.

1.1. Enquadramento

A reflexão prática “O que não conseguimos medir, não conseguimos melhorar” de *William Thompson (Lorde Kelvin)* teve a sua origem na física, embora esta possa ser aplicada a várias áreas do conhecimento. Assim, e aplicando este conceito ao tema em estudo, ganhos energéticos só são alcançáveis após medições dos consumos energéticos, sendo que, atualmente a existência de sistemas de medição de energia de edifícios é comum, tendo estes a capacidade de fornecer os dados necessários para uma análise fina do consumo energético.

Hoje, face aos constantes aumentos dos custos associados aos mercados da energia, o tema da sustentabilidade ganha uma importância acrescida. Aliado a este tema, vem o da eficiência energética. O princípio da eficiência significa a capacidade de produzirmos o mesmo com menos recursos que outrora, o que representa ganhos em qualquer que seja o processo. Todavia, este princípio não basta por si só para alcançar uma sociedade desenvolvida, cabendo a cada um de nós, assim como aos governos e às empresas, estruturarmo-nos num outro, o princípio da suficiência, que objetiva a minimização do desperdício.

Em relação ao tema da energia, em Portugal, no decorrer dos anos, a capacidade instalada de fontes de energia renovável tem vindo a aumentar e o consumo energético tem vindo a diminuir, sendo este último aspeto visível na Figura 1 [1]. Contudo, sobre este aspeto

devemos ter em consideração dois pontos. O primeiro encontra-se relacionado com a pandemia COVID-19, tendo esta sido mais proeminente durante o ano de 2020 e 2021, fazendo com que alguns estabelecimentos de comércio e serviços, assim como edifícios industriais, embora estes últimos de forma menos rígida, fechassem ou diminuíssem substancialmente a sua atividade. Este facto resultou numa redução do consumo energético da tipologia de edifícios mencionada. Contudo, resultou também num aumento do consumo energético de edifícios habitacionais, embora estes possuam uma capacidade energética instalada inferior relativamente aos edifícios de comércio/serviços e industriais, logo um consumo também inferior. O segundo ponto refere-se a um aspeto já abordado: o consumo energético de um país encontra-se relacionado com a sua produção, o que por sua vez se encontra relacionado com o seu crescimento económico. Logo, embora os gráficos relacionados com o consumo energético permitam dar uma ideia da evolução do consumo, da diversificação das fontes de energia e da parcela específica de cada setor, estes devem ser analisados com cautela, uma vez que não espelham de uma forma clara os ganhos de energia provenientes do aumento de eficiência de processos ou tecnologia.

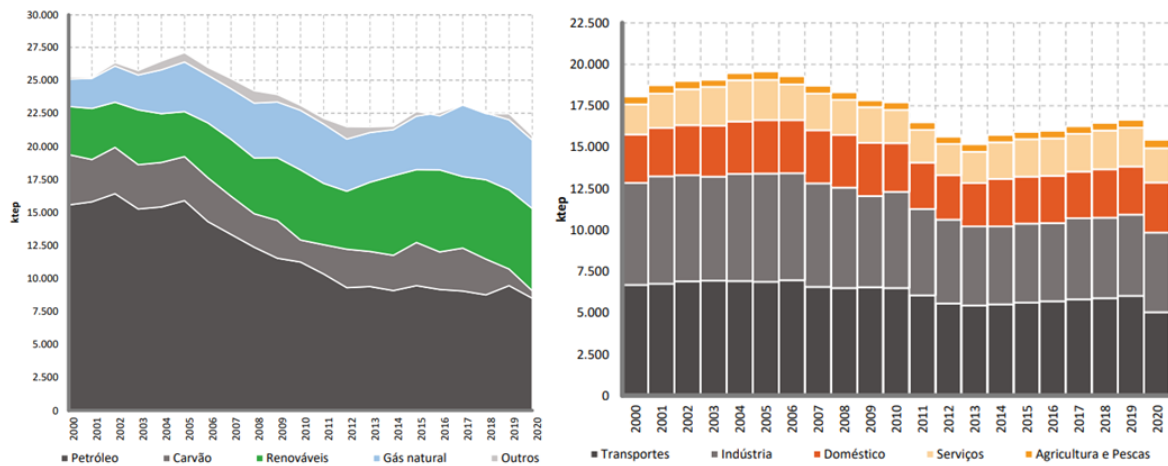


Figura 1 – Evolução do consumo total de energia primária (Esquerda), Evolução do consumo de energia final por setor de atividade (Direita), ambos referentes a Portugal. Fonte: DGEG

Em Portugal, uma das formas de medir os ganhos energéticos dos edifícios é através da emissão e análise do certificado energético, sendo que estes possuem a classificação global do desempenho energético de um edifício e medidas que promovem ganhos de eficiência e, conseqüentemente, uma classificação superior. Porém, a maioria destas medidas carecem de um investimento substancialmente elevado e possuem períodos de retorno também bastante elevados, o que as torna pouco atrativas.

Outra forma de medir o desempenho energético de um edifício é através da monitorização do seu consumo relativo à sua produção. O presente trabalho foca-se nos consumos energéticos de um edifício de comércio e serviços, nomeadamente, serviços de formação profissional da indústria metalúrgica e metalomecânica. Os consumos energéticos associados a este edifício são, naturalmente, os consumos relacionados com o seu funcionamento, onde se incluem os sistemas de climatização e iluminação relacionados com os ocupantes dos espaços e os sistemas de apoio à formação, nomeadamente, máquinas, ferramentas elétricas e sistemas de produção, distribuição e exaustão de gases. Esta análise promove uma maior consciencialização para a utilização eficiente dos recursos, pelo que a previsão de consumos energéticos e a deteção atempada de consumos anómalos configuram-se também como estratégias. De forma a executar esta atividade, devem ser criados modelos que consigam prever com precisão o consumo energético, pelo que diferentes modelos, assim como diferentes parametrizações de cada um destes, devem ser testados de forma a encontrar o que apresenta melhor desempenho, com a finalidade de o selecionar para posterior utilização.

1.2. Breve Descrição da Entidade

O presente estudo é realizado no CENFIM – Centro de Formação Profissional da Indústria Metalúrgica e Metalomecânica, o qual foi criado em 1985 e tem como propósito a promoção de atividades de formação para valorização dos recursos humanos no setor metalúrgico, metalomecânico e eletromecânico. Geograficamente, o CENFIM distribui os seus 13 Núcleos de Formação em Portugal Continental, sobretudo na faixa litoral, de acordo com a distribuição das empresas do setor. Neste momento, os Núcleos do CENFIM encontram-se em Amarante, Arcos de Valdevez, Caldas da Rainha, Ermesinde, Lisboa, Marinha Grande, Oliveira de Azeméis, Peniche, Porto, Santarém, Sines, Torres Vedras e Trofa. Tendo em conta as atividades industriais dos locais onde cada Núcleo do CENFIM está localizado, e tendo em conta que a maior parte do volume de formação desenvolvido é na área profissional da metalurgia e metalomecânica, podemos dizer que o CENFIM explora um conjunto de áreas de formação, tais como a fabricação de produtos metálicos, a fabricação mecânica, a soldadura, o projeto de desenho e fabrico assistido, assim como as áreas de energia, frio e climatização, e de eletricidade, eletrónica e automação.

1.3. Objetivos

O principal objetivo deste trabalho consiste em aplicar técnicas de Ciência de Dados para realizar previsões de consumos de energia elétrica de um edifício do CENFIM – Núcleo de Formação da Marinha Grande e, de forma concomitante, comparar as técnicas consideradas de base estatística com algumas das técnicas mais recentes de aprendizagem computacional (área designada em inglês por *Machine Learning*). De forma breve, a análise dos consumos energéticos pode ser representada como um problema de séries temporais, pelo que as técnicas de base estatística objetivam a análise das diferentes componentes da série temporal e a aplicação de modelos regressivos, enquanto que, as técnicas de aprendizagem computacional objetivam a captação das representações lineares e não lineares dos dados em estudo para criar um modelo representativo desses dados.

Para isso, este trabalho visa a recolha e o tratamento dos dados de consumo energético e outras variáveis (temperatura, ocupação) que se configurem pertinentes, a seleção dos modelos a utilizar e, por fim, a comparação, análise e apresentação dos resultados obtidos.

1.4. Estrutura do Trabalho

Os restantes capítulos deste documento estão organizados da seguinte forma:

- Capítulo 2, Matéria Relacionada – Através da matéria relacionada começamos por introduzir os conteúdos relacionados com os temas que são desenvolvidos neste relatório;
- Capítulo 3, Estado da Arte – Neste capítulo é analisado o estado da arte dos conteúdos desenvolvidos no decorrer deste trabalho;
- Capítulo 4, Metodologia, Modelos e Métricas – Capítulo onde é apresentada a metodologia que se adotou para o tratamento dos dados e a análise dos resultados. Neste capítulo ainda são descritos e explicados os algoritmos utilizados na realização das previsões dos consumos energéticos do edifício em estudo, assim como as métricas de erro para a análise dos resultados;
- Capítulo 5, Descrição e Tratamento dos Dados – Neste capítulo são descritos aspetos relacionados com a recolha e tratamento dos dados, as operações de limpeza realizadas e outros aspetos relacionados com o edifício de onde os dados foram originados;

- Capítulo 6, Criação dos Modelos e Análise dos Resultados – Nesta fase descrevem-se os modelos utilizados e as suas parametrizações, assim como a apresentação e discussão dos resultados obtidos;
- Capítulo 7, Conclusão – Por fim, é apresentada a conclusão do trabalho realizado e as perspetivas de conteúdos que possam originar trabalho futuro.

Esta página foi propositadamente deixada em branco

2. Matéria Relacionada

O estudo das séries temporais tem sido uma área de interesse em diversos campos do conhecimento. O objetivo principal deste estudo é analisar e compreender a dinâmica de uma variável registada ao longo do tempo, isto é, identificar padrões e tendências numa sequência de valores observados. Assim, a utilização de técnicas de base estatística, as quais foram melhoradas mais tarde com o apoio e modernização de sistemas computacionais, tem-se tornado cada vez mais relevante na análise de séries temporais, permitindo a criação de modelos preditivos cada vez mais precisos e eficazes. A aprendizagem computacional utiliza algoritmos que aprendem a partir dos dados e são capazes de identificar padrões complexos numa grande quantidade de informação.

Este capítulo tem como objetivo introduzir conceitos gerais de séries temporais e de aprendizagem computacional, destacando os principais conceitos e técnicas utilizados para a análise de dados relativos a séries temporais. A primeira secção aborda conceitos gerais de séries temporais, como estacionariedade, tendência, sazonalidade, autocorrelação, entre outros. São apresentadas as principais técnicas de análise de séries temporais, como as funções de autocorrelação e as técnicas de decomposição. A segunda secção explora os conceitos gerais de aprendizagem computacional, tais como as redes neuronais e a aprendizagem profunda (designada em inglês por *Deep Learning*).

2.1. Conceitos Gerais de Séries Temporais

A análise de séries temporais (também conhecidas por sucessões cronológicas) é uma área de estudo que conta com vários anos de história e que consiste em entender e prever fenómenos que variam ao longo do tempo. Esta técnica ganhou bastante interesse, especialmente para economistas e estatísticos nos seus primórdios, uma vez que a implementação de medidas e estratégias de uma forma atempada apenas se torna possível através da análise e da interpretação do passado [2]. As séries temporais são conjuntos de dados organizados por ordem cronológica, isto é, cada variável em estudo varia consoante a sua progressão no tempo. Assim, são consideradas séries temporais os conjuntos de dados que variam ao longo do tempo tais como a temperatura, o Produto Interno Produto (PIB), a quantidade e tipos de produtos vendidos ou a quantidade de energia consumida por edifícios ou sistemas destes. A aplicação do estudo das séries temporais é utilizada em diversas áreas

do conhecimento, tais como economia, finanças, meteorologia, saúde, engenharia, entre outras. Nesta área, quando se estuda apenas uma variável dizemos que estamos no caso univariado, enquanto que o estudo de mais de uma variável em simultâneo designa-se por análise multivariada.

O estudo das séries temporais é uma área de investigação que se encontra em constante evolução, sendo impulsionada pela disponibilidade de dados em grande escala e pela necessidade de tomar decisões em tempo real [3]. De facto, este estudo requer geralmente uma grande quantidade de dados para garantir consistência e fiabilidade durante o tratamento estatístico, permitindo uma análise que possa eliminar dados ruidosos e identificar corretamente os padrões associados a movimentos ou variações características de cada série, designados por componentes da série. A análise de séries temporais ganha um maior interesse quando utilizada para fazer previsões. Quando os dados são analisados em intervalos constantes é possível realizar previsões de probabilidade de eventos futuros, sendo esta considerada uma análise preditiva, ao mesmo tempo que possibilita a identificação de mudanças prováveis das componentes da série, o que por sua vez fornece uma melhor compreensão das variáveis em estudo, permitindo igualmente uma previsão de boa qualidade, que pode ser interpretada pelas medidas de precisão [4].

Enquanto que realizar previsões de ativos financeiros do mercado acionista é difícil devido ao elevado grau de complexidade associado ao valor de um ativo, isto é, existem muitas variáveis que afetam o preço de um ativo financeiro difíceis de quantificar e analisar, outras variáveis mais simples e que apenas dependam de alguns parâmetros possibilitam uma previsão com uma precisão considerável. Por exemplo, um hospital pode usar a análise de séries temporais para identificar oportunidades de melhoria no atendimento ao paciente e reduzir os custos para os pacientes e para o hospital. Foi desta forma que o hospital *Stamford Health* [5] usou o histórico de dados dos seus pacientes e identificou melhores horários para administrar medicamentos, reduzindo o tempo médio de internamento e os custos associados a estes tratamentos.

As componentes das séries temporais são mais facilmente explicadas quando visualizadas. Assim, apresenta-se na Figura 2 um exemplo de uma série temporal que possui as quatro componentes, consideradas nas abordagens clássicas. A série temporal é constituída por dados trimestrais recolhidos durante cerca de 50 anos relativamente à quantidade de tijolos produzidos na Austrália.

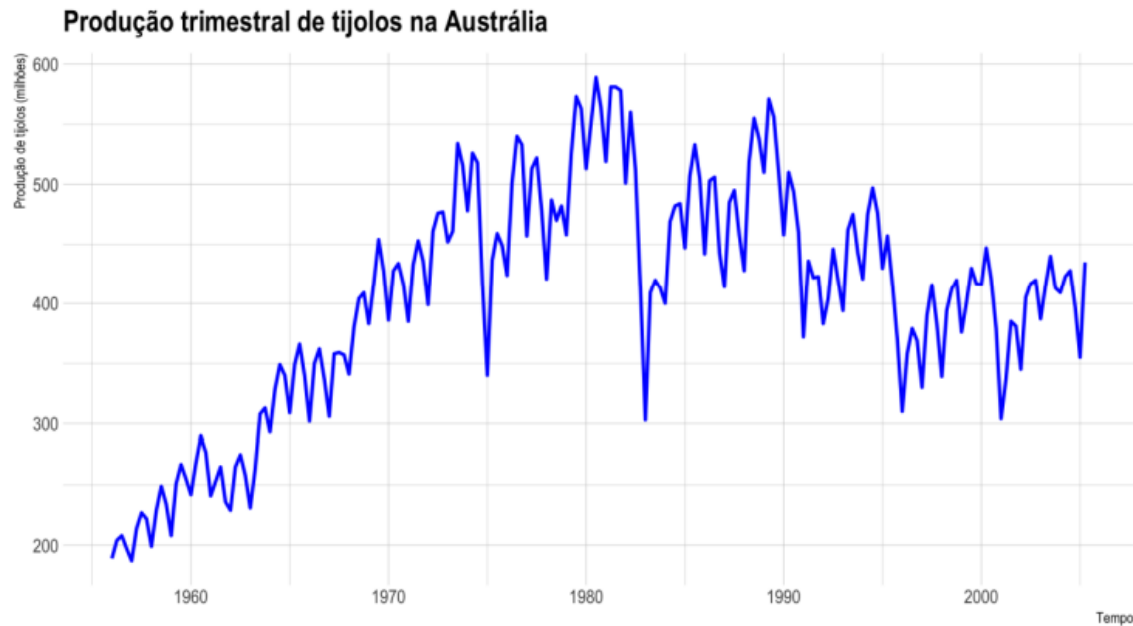


Figura 2 – Produção trimestral de tijolos na Austrália

Através da inspeção visual, sendo esta a primeira análise que se deve realizar no estudo de séries temporais, é possível identificar o seguinte [6]:

- **Tendência (T_t):** Esta componente identifica se os valores se encontram, de uma forma geral, a aumentar ou diminuir e de que forma o fazem ao longo do tempo. Esta é a componente mais simples de identificar numa série temporal. Neste caso, é facilmente visível que a produção de tijolos na Austrália teve um aumento até cerca de 1980 e, sensivelmente após essa data, a produção de tijolos até ao final da série apresenta uma tendência decrescente.
- **Sazonalidade (S_t):** A sazonalidade de uma série temporal caracteriza-se pela constante e regular variação de dados ao longo de um período fixo de tempo. Por exemplo, é expectável que o consumo de energia de um edifício de serviços, que esteja aberto durante o dia e fechado durante a noite, possua sazonalidade diária, isto é, num período de 24 horas, existe sempre um consumo mais elevado de energia durante o dia e uma diminuição de consumo durante a noite. Na série apresentada é visível uma componente sazonal anual, uma vez que em todos os anos a produção de tijolos é superior no 2º e 3º trimestre e inferior no 1º e 4º trimestre de cada ano.
- **Componente cíclica (C_t):** Esta componente é mais difícil de identificar e nem sempre tem sido considerada na literatura. Padrões cíclicos existem quando os dados apresentam subidas ou descidas repentinas em períodos inconstantes. O que diferencia os padrões cíclicos da sazonalidade é o facto de estes não apresentarem

um período fixo e conhecido de flutuações, enquanto que as variações sazonais acontecem num período possível de identificar. Na produção de tijolos na Austrália, esta componente é visível a partir do ano de 1975, onde se verifica a primeira queda de produção, sendo que as componentes sazonais e a tendência permanecem praticamente inalteradas. Este primeiro ciclo parece estender-se por oito anos até se identificar um segundo mínimo por volta do ano de 1983. Por esta altura começa um segundo ciclo que, por coincidência, também possui oito anos, estendendo-se até ao ano de 1991. Até ao final da série temporal são visíveis pelo menos mais dois ciclos com durações mais curtas e que coincidem com os mínimos observados no ano de 1996 e 2001. Estas quedas abruptas podem ser explicadas pelas mais variadas questões políticas, socioeconómicas ou por questões de oferta e procura.

- **Componente Irregular (I_t):** Esta componente, também designada por componente aleatória ou de ruído, à semelhança da anterior, deriva de vários fatores que acabam por influenciar o comportamento de uma série temporal, sendo estes imprevisíveis, como greves, eleições, epidemias. Todas as séries temporais reais apresentam componentes irregulares que podem afetar o movimento da série temporal e, por sua vez, originar movimentos cíclicos.

Quanto à componente irregular, esta é semelhante ao erro aleatório em modelos de regressão, como será apresentado mais à frente neste trabalho. Contudo, esta componente é visível apenas quando se realiza a decomposição ou a análise da série temporal e caracteriza-se como a componente que não pode ser explicada pelo modelo utilizado para ajustar as restantes componentes. Ao construir um modelo para uma série temporal, o objetivo é extrair os diferentes padrões subjacentes nos dados, por forma a efetuar previsões minimizando os respetivos erros de previsão. A componente aleatória é uma importante ferramenta de diagnóstico na avaliação do ajuste de um modelo de série temporal, pois permite avaliar a qualidade do ajustamento das previsões aos dados originais.

Do ponto de vista da análise ao longo do tempo, as séries temporais podem ser divididas em dois tipos, nomeadamente, as séries estacionárias e as não estacionárias, sendo que para efetuar esta análise normalmente recorreremos a testes de hipóteses, como o teste *Augmented Dickey-Fuller (ADF)* ou o teste *Kwiatkowski-Phillips-Schmidt-Shin (KPSS)* [7, 8]. As séries temporais estacionárias possuem duas características, a saber, a média e a variância praticamente constantes ao longo da série, ao contrário das séries não estacionárias onde estas medidas não assumem valores constantes. Na prática, a maior parte de séries com

tendência e sazonalidade não são estacionárias, assim como o exemplo apresentado anteriormente, sendo que, para a análise adequada das séries deverá ser efetuado um teste de hipóteses para averiguar esta condição. A modelação de séries temporais através de modelos estatísticos exige geralmente que a série seja estacionária, pelo que existem vários métodos, como a remoção da tendência, a aplicação de diferenciações ou a logaritmização dos dados para a tornar estacionária [9].

O estudo de séries temporais começa na maior parte das vezes com a averiguação da existência das diferentes componentes da série, sendo esta frequentemente classificada como a decomposição da série nos seus movimentos característicos. Nos modelos de decomposição de séries temporais inclui-se a decomposição clássica, o método de médias móveis, a suavização exponencial e o método *STL (Seasonal-trend Decomposition Procedure Based on Loess)* [10]. Quanto à decomposição clássica existem dois tipos de modelos principais, sendo estes o modelo aditivo e o modelo multiplicativo, representados pelas equações abaixo, onde X_t representa a série temporal ao longo do tempo t . Alguns autores não consideram a componente cíclica C_t , o que simplifica as expressões.

$$X_t = T_t + S_t + C_t + I_t$$

$$X_t = T_t * S_t * C_t * I_t$$

Os métodos estatísticos de análise de séries temporais incluem a decomposição da série temporal nas suas diferentes componentes, a aplicação dos mais variados modelos tais como médias móveis para suavizar a série ou os modelos *ARIMA (Autoregressive Integrated Moving Average)* e *SARIMA (Seasonal Autoregressive Integrated Moving Average)*, sendo estes tópicos abordados mais à frente. Independentemente disso, existem parâmetros que deverão ser definidos quando aplicados estes métodos, que surgem da análise de conceitos de séries temporais. Destes conceitos, devem destacar-se as funções de autocorrelação (*ACF*, do inglês *Auto-Correlation Function*) e de autocorrelação parcial (*PACF*, do inglês *Partial Auto-Correlation Function*) [11].

A autocorrelação é utilizada para indicar como uma observação de uma série temporal é semelhante a outra num instante temporal diferente e desfasado. Por outras palavras, os gráficos de autocorrelação indicam o grau de similaridade entre uma série temporal e a versão desfasada dessa série (*lags*) nos vários intervalos observados. A função de autocorrelação parcial é semelhante à função de autocorrelação, contudo é um pouco mais

difícil de entender. A autocorrelação parcial é um resumo das correlações das observações de uma série temporal com as observações da versão desfasada da série onde as correlações entre estas observações foram removidas. A autocorrelação parcial no *lag* 10, por exemplo, é a correlação que resulta após a remoção de quaisquer efeitos de correlação devido a *lags* anteriores. Resumindo, a função de autocorrelação, para uma observação e outra numa etapa anterior, é composta pelas correlações diretas e indiretas. As correlações indiretas são uma função linear da correlação entre a observação e as observações de *lags* intermédios. A função de autocorrelação parcial procura remover as correlações indiretas [12].

Na Figura 3 apresentam-se os gráficos de autocorrelação e autocorrelação parcial correspondentes à série temporal de produção de tijolos na Austrália. As observações que excedam o tracejado azul são consideradas significativas e podemos verificar que, em ambos os gráficos, existem vários *lags* significativos.

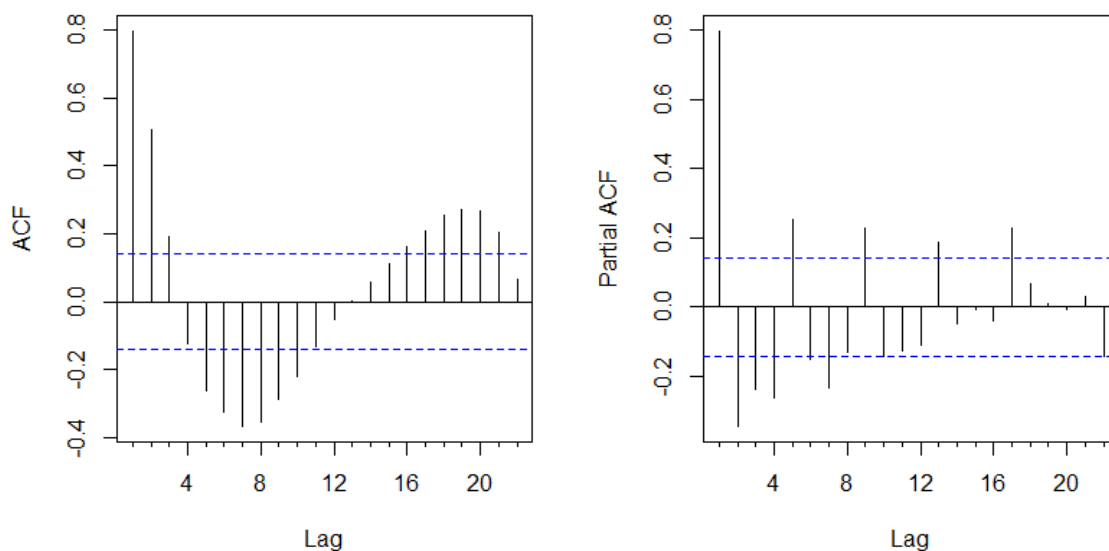


Figura 3 – Esquerda: Gráfico de autocorrelação. Direita: Gráfico de autocorrelação parcial. Ambos referentes à série temporal de produção de tijolos na Austrália

Estes gráficos não apresentam o *lag* 0, uma vez que os valores de autocorrelação e autocorrelação parcial são sempre iguais a 1 nesse ponto. Este valor representa a correlação da série temporal com ela própria, o que é esperado e não fornece informações úteis para identificar padrões de sazonalidade ou tendência na série. É também possível verificar que, o gráfico de autocorrelação aparenta uma função sinusoidal, o que é indicativo de existência de uma componente sazonal na série. Sem entrar em grande detalhe, os *lags* significativos identificados nestes gráficos ajudam a parametrizar modelos estatísticos adequados para esta série definindo as suas ordens, que poderão posteriormente ser utilizados para realizar

previsões. Nestes parâmetros incluem-se os autorregressivos, as ordens de integração ou de diferenciação e os relativos às médias móveis.

Os gráficos ACF e PACF podem configurar-se de várias formas, as quais dependem do tipo de série temporal que se está a analisar e que, por sua vez, ajudam a parametrizar o melhor modelo a aplicar para a série em questão. De seguida, apresenta-se uma tabela resumo com alguns tipos de gráficos que se podem obter e alguns modelos estatísticos que se identificam como mais adequados para cada uma das hipóteses [13]. Para o exemplo apresentado na Figura 2, como ambos os gráficos de autocorrelação e de autocorrelação parcial parecem cair gradualmente ou de forma exponencial, o mais adequado seria um modelo autorregressivo de médias móveis. Todavia, ainda seria necessário analisar a estacionariedade da série de forma a perceber se deveria ser aplicada alguma ordem de diferenciação. Através da aplicação de métodos para tornar uma série temporal estacionária e da análise dos gráficos ACF e PACF, onde analisamos os *lags* significativos e a forma como os gráficos se comportam, obtemos as ordens dos modelos autorregressivos, de médias móveis e as ordens de diferenciação para parametrizar os modelos ARIMA e SARIMA.

Tabela 1 – Modelos vs Autocorrelação e Autocorrelação parcial

ACF	PACF	Modelo
Gráfico cai gradualmente	Gráfico cai instantaneamente a partir do <i>lag p</i>	Autorregressivo (<i>p</i>)
Gráfico cai instantaneamente a partir do <i>lag q</i>	Gráfico cai gradualmente	Médias móveis (<i>q</i>)
Gráfico cai gradualmente a partir do <i>lag q</i>	Gráfico cai gradualmente a partir do <i>lag p</i>	Autorregressivo (<i>p</i>) de médias móveis (<i>q</i>)
Gráfico cai instantaneamente	Gráfico cai instantaneamente	Nenhum modelo deve ser aplicado

As principais desvantagens na aplicação destes modelos de previsão é lidar com a presença de ruído e com a variabilidade nos dados, sendo que para enfrentar estas dificuldades são aplicados vários métodos de tratamento de dados. Estes modelos possuem uma ampla variedade de aplicações com grande interesse e, com os recentes avanços em tecnologias de tratamento e análise de dados, como a aprendizagem computacional, o futuro (embora este já bastante presente) da previsão de séries temporais envolverá o uso de modelos mais complexos, como as redes neuronais, e uma maior integração de técnicas de aprendizagem

computacional e Inteligência Artificial num formato, possivelmente, híbrido com diferentes técnicas para diferentes casos e aplicações [14].

2.2. Conceitos Gerais de Aprendizagem Computacional

A utilização de ferramentas ao longo da história teve sempre o mesmo intuito, facilitar a execução de uma tarefa através da criação de mecanismos que a tornem mais simples. Assim, conseguimos mudar a forma como realizamos atividades nas mais variadas áreas, como a dos transportes, na indústria e na computação, melhorando a nossa qualidade de vida. A aprendizagem computacional é uma dessas ferramentas que, em várias áreas do conhecimento, tem revolucionado a forma como lidamos com os problemas.

Atualmente, várias aplicações que usamos numa base diária, como motores de busca, apresentam resultados que provêm da utilização desta tecnologia, a qual tem vindo a melhorar os resultados obtidos nas pesquisas que efetuamos. O que difere a aprendizagem computacional da programação clássica, é o facto de que o primeiro tem como objetivo dar a capacidade de aprendizagem a uma máquina sem que esta tenha sido explicitamente programada para a atividade em questão. Isto é, na programação clássica são utilizados como entrada dados e uma lógica e, através desta lógica, o programa dá uma resposta como saída, enquanto que, na aprendizagem computacional, o modelo é alimentado com entradas e as saídas correspondentes, sendo a lógica gerada pelo algoritmo de aprendizagem computacional e posteriormente aplicada a novos dados. Por exemplo, na atividade de deteção de *spam* de email, a programação clássica compara os emails com um conjunto de emails que são rotulados como *spam* e, caso identifique aspetos iguais, como o remetente ou outra informação característica, marca o email como *spam*. Para treinar algoritmos de aprendizagem computacional para esta atividade, seria criada uma base de dados que contivesse emails normais e emails de *spam* devidamente rotulados como tal, sendo o modelo treinado e testado com esse conjunto de dados e, posteriormente, posto em utilização. A programação clássica é, portanto, considerada uma técnica determinista, enquanto que a aprendizagem computacional é probabilística [15]. Embora nenhuma destas técnicas alcance uma precisão de 100%, a programação clássica exige um maior esforço, uma vez que necessita de atualizações manuais na lógica para se adaptar a mudanças, enquanto que nas técnicas de aprendizagem computacional, os algoritmos permitem uma aprendizagem contínua quando lhes forem cedidos novos dados de treino.

A área da aprendizagem computacional utiliza diferentes algoritmos para resolver problemas com dados. Numa altura onde a quantidade de dados é abundante, muitas vezes deparamo-nos com situações onde extrair informação útil destes é uma tarefa difícil e que, geralmente, representa mais valias consideráveis em diferentes áreas de negócio. Contudo, não existe um algoritmo único e melhor para resolver todos os problemas relacionados com dados. A seleção da melhor técnica, para além de depender fortemente do nível de conhecimento do problema a resolver e da sensibilidade do cientista de dados para o mesmo, exige sempre uma certa quantidade de experimentação. Isto é, independentemente do conhecimento especializado que tenhamos sobre um problema, existe sempre a necessidade de seleccionar vários algoritmos, e vários parâmetros para cada um deles, com o objetivo de testar cada uma das hipóteses e analisar qual apresenta o melhor modelo que pode ser posto em utilização. Para além disso, estes algoritmos dependem de vários fatores, tais como o tipo de problema a resolver e o número de variáveis do conjunto de dados [16]. Hoje em dia, existe uma enorme quantidade destes algoritmos à disposição dos investigadores em geral e dos cientistas de dados em particular. Na figura seguinte, podemos observar alguns desses algoritmos, os quais incluem redes neuronais artificiais, árvores de decisão, máquinas de suporte vetorial, entre outros.

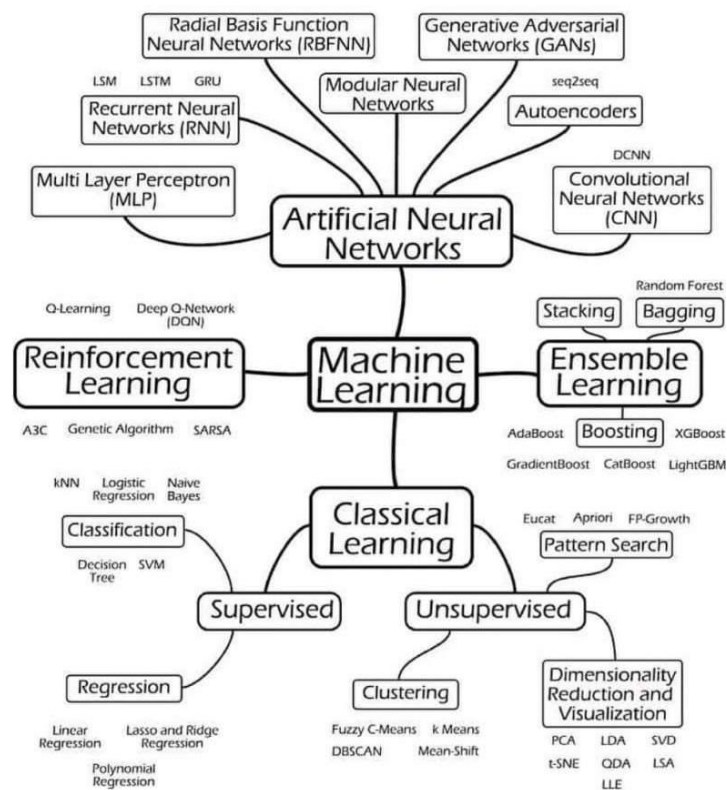


Figura 4 – Mapa de algoritmos de Aprendizagem Computacional [17]

Os problemas de aprendizagem computacional podem ser divididos em três áreas principais, nomeadamente [18]:

- **Problemas de Classificação:** Nestes problemas o objetivo é prever uma categoria de uma determinada observação. Isto é, procura-se estimar uma classificação qualitativa de um dado não observado com base em dados de entrada que serviram para treinar o modelo. Os dados de treino deste problema podem incluir, por exemplo, uma base de dados de fotografias de diferentes raças de cães, todas corretamente classificadas, e objetiva a classificação de uma nova imagem dentro de todas as categorias que foram utilizadas para o treino.
- **Problemas de Agrupamento:** Estes problemas são também conhecidos como *clustering* e têm como objetivo agrupar observações em grupos que apresentem aspetos semelhantes entre si e distintos dos outros grupos. Por exemplo, o agrupamento de diferentes raças de cães tendo em conta as características que os distinguem, como peso, estatura e outras. Nesta abordagem, ao contrário dos problemas de classificação, não se realiza uma rotulação em grupos, apenas um agrupamento tendo em conta as semelhanças de cada grupo.
- **Problemas de Regressão:** Estes problemas são semelhantes aos anteriores, contudo, a diferença encontra-se na previsão que se pretende obter. Nos problemas de regressão, em vez de classificar uma observação, pretende-se estimar um valor numérico com base em valores de entrada. Por exemplo, na análise de uma série temporal univariada, podemos pretender utilizar os valores passados para prever os valores futuros da série em estudo. Outro exemplo é a previsão dos valores de imóveis com base em várias características conhecidas. Isto é, utiliza-se como treino uma base de dados de preços de casas e outros dados relativos a estas como, tipologia, ano de construção, acessibilidades e utiliza-se o modelo para prever o preço de uma casa onde seja conhecida a informação referida.

Para além da divisão feita dos diferentes tipos de problemas, os algoritmos de aprendizagem computacional ainda podem ser distinguidos pelo processo de aprendizagem que utilizam, e que delimita o tipo de algoritmo a utilizar para o cenário a analisar. Os três métodos principais são: a aprendizagem supervisionada, a aprendizagem não supervisionada e a aprendizagem por reforço.

A aprendizagem supervisionada envolve o treino de um modelo com dados rotulados, onde o algoritmo aprende padrões e relações entre os dados de entrada e os rótulos correspondentes. Esta metodologia é utilizada em cenários como reconhecimento de imagem e classificação de texto, onde podem ser utilizados algoritmos como regressões lineares e árvores de decisão para fazer previsões e classificações. Ao contrário da aprendizagem supervisionada, a não supervisionada não utiliza dados rotulados e tem como objetivo descobrir padrões, estruturas ou relações inerentes nos dados em estudo. Algumas das técnicas mais utilizadas nesta metodologia incluem o agrupamento através da técnica do vizinho mais próximo e redução de dimensionalidade, permitindo revelar informações e relações nos dados que, de outra forma seria difícil de visualizar em dados não estruturados. Já a aprendizagem por reforço envolve um agente que, ao interagir com um ambiente e à medida que toma decisões sequenciais para maximizar uma recompensa, aprende um “caminho” ou um conjunto de decisões que permitem resolver o problema da forma mais eficiente possível. Esta metodologia utiliza conceitos de probabilidade de processos estocásticos, como o Processo de *Markov*, e sinais de reforço ou, por outras palavras, recompensas por decisões positivas. Esta metodologia permite que a máquina aprenda através da exploração do ambiente onde se encontra e abriu caminho para avanços em áreas como a automação, robótica e jogos tradicionais, como é o caso do xadrez ou do *go*.

Dentro das técnicas de aprendizagem computacional, destaca-se a aprendizagem profunda (já referida anteriormente como *Deep Learning*). Esta abordagem tem como objetivo simular o funcionamento do cérebro humano para realizar tarefas de processamento de informação complexas. Estes modelos são denominados por redes neuronais artificiais que são constituídos por camadas de neurónios artificiais, como visível na Figura 5. As redes neuronais possuem a capacidade de aprender representações hierárquicas dos dados em estudo com diferentes níveis de abstração, de acordo com a quantidade de camadas que constitui a arquitetura da rede [19]. Os algoritmos de *Deep Learning* permitiram melhorar significativamente o desempenho de modelos artificiais em problemas como o de reconhecimento de fala e o de reconhecimento visual de objetos, ao terem a capacidade de descobrir estruturas complexas num grande volume de dados.

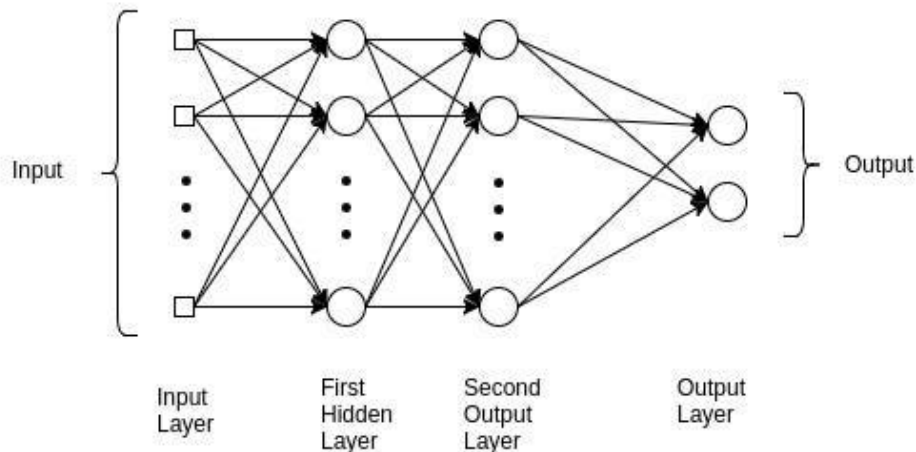


Figura 5 – Exemplo de uma rede neuronal artificial [20]

Estas técnicas têm sido desenvolvidas desde a década de 1960, contudo, foram consideradas de aplicação difícil no final da década de 1980. Só mais tarde, com a evolução das tecnologias de computação, como o caso dos GPUs, o desenvolvimento de algoritmos mais eficientes, o aumento da disponibilidade de dados e por se ter verificado que estas tecnologias superavam outras de aprendizagem computacional, atraíram uma atenção generalizada [21].

Uma rede neuronal utilizada frequentemente é a *MLP – Multilayer Perceptron*, que é composta por uma arquitetura de várias camadas constituídas por neurónios artificiais ligados entre si de uma camada para a camada seguinte, como representado na Figura 5. Para perceber como funciona uma *MLP*, temos que entender os diferentes componentes principais que a constituem, sendo estes os dados de entrada, os pesos, o *bias*, e o mecanismo de atualização dos pesos.

Os dados de entrada correspondem a vetores de valores que representam as diferentes características dos dados, por exemplo, em problemas de regressão e análise de séries temporais, estes vetores de entrada podem representar uma quantidade de dias definida para aplicar qualquer uma das técnicas.

Os pesos representam valores que são atualizados à medida que se treina a rede neuronal. No início, quando uma rede neuronal é treinada, os pesos são iniciados com valores aleatórios, sendo que, à medida que a rede é treinada, estes são atualizados de forma a melhorar o desempenho da rede. Os pesos determinam a força das conexões entre os neurónios artificiais da rede e controlam como a informação circula na rede. Para além dos

pesos, cada neurónio da rede neuronal possui um *bias*, sendo este um parâmetro que ajuda a rede neuronal a capturar a tendência geral dos dados. Este *bias* é também atualizado à medida que uma rede é treinada através de algoritmos como o de gradiente descendente. A diferença entre o *bias* e os pesos da rede é que o *bias* não é influenciado pelos resultados obtidos da camada anterior, ao contrário dos pesos [22].

A saída de uma *MLP* é gerada por meio de uma série de cálculos que envolvem a soma ponderada das entradas e uma função de ativação. Cada neurónio da rede realiza a soma ponderada das suas entradas com os pesos das conexões, adiciona o termo *bias* e aplica uma função de ativação para produzir uma saída, como exemplificado na Figura 6. As funções de ativação comuns incluem funções lineares, *sigmóides*, *ReLU* (unidade linear retificada) e *tanh* (tangente hiperbólica) [23]. Esta transmissão de informação de camada para camada permite que a rede aprenda representações hierárquicas de características nos dados.

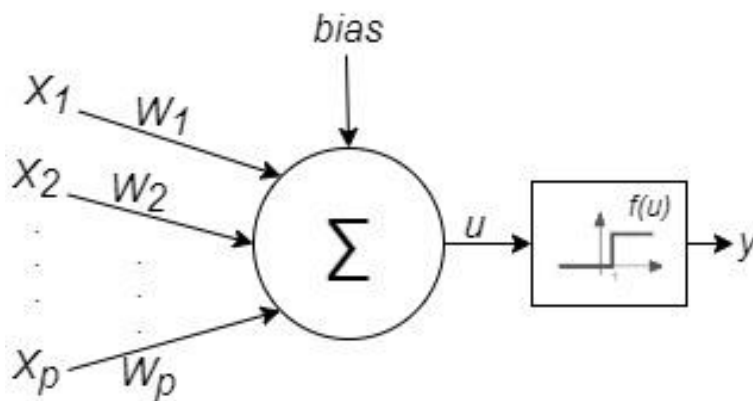


Figura 6 – Entradas e saídas de um neurónio artificial de uma rede neuronal

Na Figura 6 os cálculos efetuados no neurónio apresentado são os seguintes, onde u representa a soma ponderada das entradas (X_1 a X_p) multiplicadas pelos respetivos pesos (W_1 a W_p) mais o *bias*:

$$u = X_1 * W_1 + X_2 * W_2 + \dots + X_p * W_p + bias$$

O neurónio representado na figura possui uma função de ativação, y , bastante simples que assume apenas dois valores, sendo que a saída do neurónio é o resultado de:

$$y = f(u) = \begin{cases} y = 1, se u \geq 1 \\ y = 0, se u < 1 \end{cases}$$

Assim, o valor à saída do neurónio em análise assume apenas um de dois valores, ou 1 ou 0, sendo este facto determinado pela função de ativação.

Vejam os outros exemplos numa rede que contém dois dados de entrada, peso e altura, definidos pelas letras p e a , e uma camada escondida com dois neurónios, representada na figura seguinte.

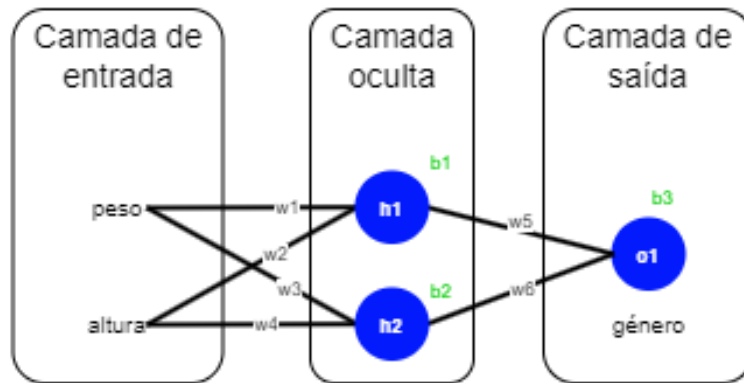


Figura 7 – Exemplo de uma rede neuronal para determinar o género com base no peso e na altura

O cálculo dos valores de saída dos neurónios seria dado pelas seguintes expressões, onde f representa a função de ativação utilizada na camada da rede, w os pesos da rede, b os *bias* de cada neurónio, $h1$, $h2$ e $o1$ o resultado de cada neurónio:

$$h1 = f(p * w_1 + a * w_2 + b1)$$

$$h2 = f(p * w_3 + a * w_4 + b2)$$

$$o1 = f(h1 * w_5 + h2 * w_6 + b3)$$

O valor obtido em $o1$ (neste caso, pretende-se prever o género de uma pessoa com base em dois parâmetros apenas, o peso e a altura) poderá ser resultado de uma função de ativação semelhante à apresentada anteriormente, onde, por exemplo, o valor 0 representa o género masculino e o valor 1 representa o género feminino.

O mecanismo de atualização em *MLPs* é baseado num algoritmo de otimização chamado *backpropagation* [24]. Durante o processo de treino, a saída do modelo é comparada com a saída desejada dos dados rotulados por meio de uma função de perda, que quantifica a diferença entre os valores previstos e os reais. O algoritmo de *backpropagation* calcula o gradiente da função de perda em relação aos pesos e aos *bias*, permitindo que o modelo ajuste estes parâmetros de forma a minimizar os erros.

O treino de um *MLP* envolve iterar o conjunto de dados várias vezes, conhecido como épocas (*epochs*). Em cada época, o modelo processa todo o conjunto de dados. Isto é, caso sejam

definidas 100 épocas para o processo de treino, o conjunto inteiro de dados será processado pela rede neuronal 100 vezes, não querendo dizer que os pesos e os *bias* sejam atualizados apenas 100 vezes. Outro parâmetro crítico no treino das *MLP* é o tamanho do lote (*batch size*). O tamanho do lote define a quantidade de observações processadas pela rede neuronal antes de os pesos e os *bias* serem atualizados. Assim, quanto menor for o tamanho do lote, mais frequente são as atualizações dos parâmetros da rede.

Todos os componentes das redes *MLP* e a forma como estas são construídas permitem que estes algoritmos aprendam padrões complexos em grandes conjuntos de dados e façam previsões precisas numa gama ampla de tarefas em várias áreas do conhecimento.

No campo do *Deep Learning*, é importante mencionar as diversas formas de relação entre os dados de entrada e saída, nomeadamente os conceitos de um-para-um (1:1/*one to one*), uma-para-muitos (1:N/*one to many*), muitos-para-um (N:1/*many to one*) e muitos-para-muitos (N:M/*many to many*), representados na Figura 8. Estas relações entre dados de entrada e de saída podem influenciar os resultados obtidos durante o treino de uma rede neuronal e ser aplicados consoante o tipo de ferramenta que se pretende criar.

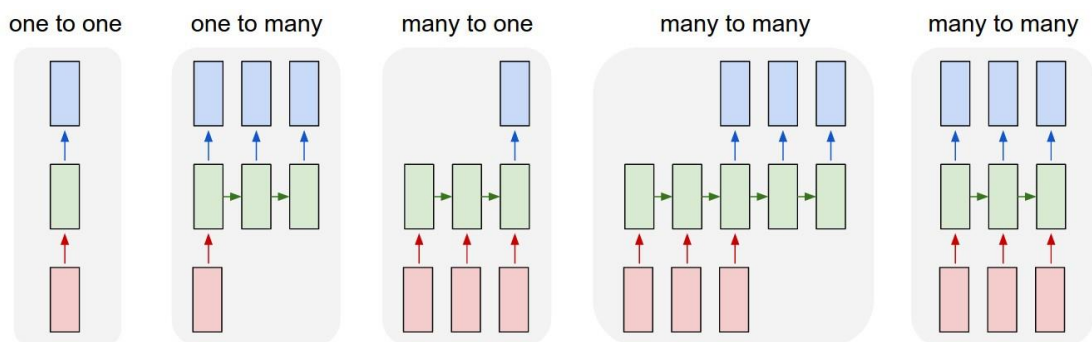


Figura 8 – Tipos de relação entre entradas e saídas de uma rede neuronal

Numa relação de um-para-um, cada dado de entrada corresponde a um dado de saída. Isto é, existe um mapeamento entre cada entrada e a correspondente saída. Por exemplo, na classificação de imagens, uma rede neuronal recebe uma imagem como entrada e faz corresponder uma previsão de classe onde a imagem se insere. Nas relações um-para-muitos, uma única entrada gera um conjunto de saídas. Este tipo de relação é comum em redes que geram legendas de imagem, onde a rede neuronal recebe uma imagem e tem como propósito gerar uma frase que a descreva. De forma inversa, as relações muitos-para-um, representam uma relação que recebe várias entradas e gera apenas uma saída, esta relação pode ser encontrada em problemas de regressão, onde são utilizados, por exemplo, os últimos dez

valores para prever o décimo primeiro. Outra aplicação deste tipo de relação entre entradas e saída é na análise de sentimento, onde um texto pode ser analisado para perceber qual o sentimento relacionado com o tipo e conteúdo do texto. Por fim, as relações muitos-para-muitos, recebem várias entradas que resultam em várias saídas. O comprimento destas pode diferir e é utilizada, por exemplo, em atividades de tradução.

3. Estado da Arte

Atualmente, o crescimento económico de países subdesenvolvidos e o aumento da população mundial estão a causar um aumento no consumo global de energia. Apesar dos ganhos de eficiência energética alcançados nas últimas décadas pelo desenvolvimento tecnológico dos sistemas, é estimado que entre 2017 e 2040 a procura global de energia aumente em 67% [25]. Além disso, outros fatores como as preocupações ambientais e os eventos geopolíticos, sendo o mais recente e relevante a guerra entre a Ucrânia e a Rússia, afetam a disponibilidade e a economia de utilização de recursos energéticos. O consumo energético de edifícios representa cerca de 40% do total de energia consumida no mundo, o que o torna responsável por uma parcela significativa do mercado global de energia [26]. Por estes motivos, a previsão de consumos energéticos para alcançar uma maior eficiência energética em edifícios tornou-se num campo de grande interesse, potenciando uma redução dos custos associados com o consumo e, de forma consequente, uma redução no impacto ambiental que esta parcela de energia consumida representa.

A criação de modelos de previsão de consumos energéticos é um tópico de estudo que tem vindo a ser analisado, existindo hoje uma ampla variedade de modelos que utilizam diferentes métodos de previsão. Estes modelos são abordados por vários autores aqui referenciados e, de uma maneira geral, podem ser divididos em três grupos principais, nomeadamente, os modelos físicos, ou de engenharia (“*white-box*”), os modelos baseados em dados, ou *data-driven* (“*black-box*”) e os modelos híbridos (“*grey-box*”) [27].

Os modelos físicos (“*white-box*”) preveem o comportamento térmico de um edifício através de equações que consideram as trocas de calor entre espaços, tendo em conta as propriedades físicas dos materiais de construção e outras características como a orientação solar e os elementos de sombreamento dos espaços exteriores. Existem já muitos sistemas de previsão baseados em modelos físicos, como o *EnergyPlus*, o *TRNSYS*, o *DOE-2* e o *eQUEST*, sendo estes utilizados, a título de exemplo, para efeitos de certificação energética de edifícios em Portugal. As principais vantagens destes modelos são a sua capacidade de descrever os mecanismos de transferência de calor e de apresentar uma visão computacional detalhada do edifício em estudo [28]. Contudo, quanto maior for a complexidade do edifício maior é a complexidade na criação destes modelos. Ao contrário dos edifícios habitacionais, os edifícios industriais e de serviços possuem outras dinâmicas como o perfil de ocupação e de

utilização de equipamentos, os sistemas de aquecimento, ventilação e ar condicionado (AVAC), redes de distribuição e outros equipamentos de produção que fazem com que a complexidade dos modelos aumente e se torne num trabalho exigente do ponto de vista técnico. Além disso, alguns destes sistemas possuem componentes que se degradam ao longo do tempo, o que dificulta a obtenção de um modelo preciso. Estes factos, aliados às dificuldades em fazer suposições apropriadas, o elevado tempo na definição de todas estas características e a incapacidade de adaptação às vicissitudes ambientais e socioeconómicas tornam o cálculo do consumo de energia por meio de modelação do edifício muito difícil [29].

Os modelos híbridos (“*grey-box*”) surgem como uma opção para diminuir a complexidade dos modelos físicos. Estes modelos são uma combinação dos modelos físicos e dos *data-driven* e têm vindo a ganhar cada vez mais interesse nos últimos anos, uma vez que possuem a capacidade de beneficiar das vantagens dos dois métodos, simultaneamente reduzindo as desvantagens de cada um. Exemplos destes modelos utilizam técnicas de aprendizagem computacional para aprender o impacto que as variáveis em estudo representam num edifício, como as condições ambientais exteriores e o perfil de ocupação, dados que podem ser recolhidos ou obtidos através de simulação de modelos físicos [30]. Estes dados são utilizados para a criação de modelos híbridos que depois são validados com dados novos, de forma a aferir se a previsão de consumo de energia elétrica é viável e se pode ser aplicada a outros edifícios. No entanto, os modelos híbridos continuam a enfrentar os problemas que geralmente se encontram nos modelos físicos, como suposições impróprias, exigência de especialização e um tempo considerável de execução [31].

Os modelos baseados em dados (“*black-box*”), uma vez que apenas dependem dos dados de consumo energético do passado, conseguem superar as limitações mencionadas dos modelos anteriores. A aplicação de métodos matemáticos para o estudo de séries temporais possui a capacidade de descobrir padrões estatísticos nos dados disponíveis e realizar previsões para o futuro. Hoje em dia, é possível melhorar o desempenho energético dos edifícios através da instalação de sistemas modernos de operação e gestão de energia, assim como de medição e monitorização, tornando possível uma recolha vasta de dados. Esta análise e controlo potencia o desenvolvimento de atividades de redução do consumo energético, fornecendo simultaneamente os dados necessários para a criação de modelos com a capacidade de prever necessidades futuras. Assim, estes modelos tornaram-se alvo de maior interesse nos últimos tempos. O principal objetivo dos modelos *data-driven* é fornecer uma estimativa precisa e

confiável do consumo de energia de edifícios, sem a necessidade de conhecer todos os detalhes técnicos ou as dinâmicas deste. Estes modelos utilizam várias técnicas de base estatística e de aprendizagem computacional, ou até combinações distintas de ambas. Desta forma, possuem a capacidade de aprender padrões complexos e não lineares nos dados de consumo de energia, podendo ser úteis em situações em que os modelos de engenharia são impraticáveis devido à falta de informação detalhada ou à elevada complexidade dos sistemas de energia em questão. A aplicação dos modelos físicos e híbridos também se torna impraticável em situações em que se pretende prever o consumo energético de cidades e grandes conjuntos de edifícios, uma vez que o tempo necessário para a sua modelação é demasiado elevado, assim como o risco de obter previsões pouco precisas.

Considerando os resultados obtidos e o potencial que a aplicação de modelos baseados em dados representa, tem-se verificado um aumento exponencial nas últimas décadas de publicações sobre estes, como se verifica na Figura 9 [32]. Na maior parte dos casos, os estudos são específicos do local e os resultados dependem fortemente do horizonte de tempo de previsão e das características dos dados e dos modelos utilizados, o que torna uma generalização dos resultados difícil.

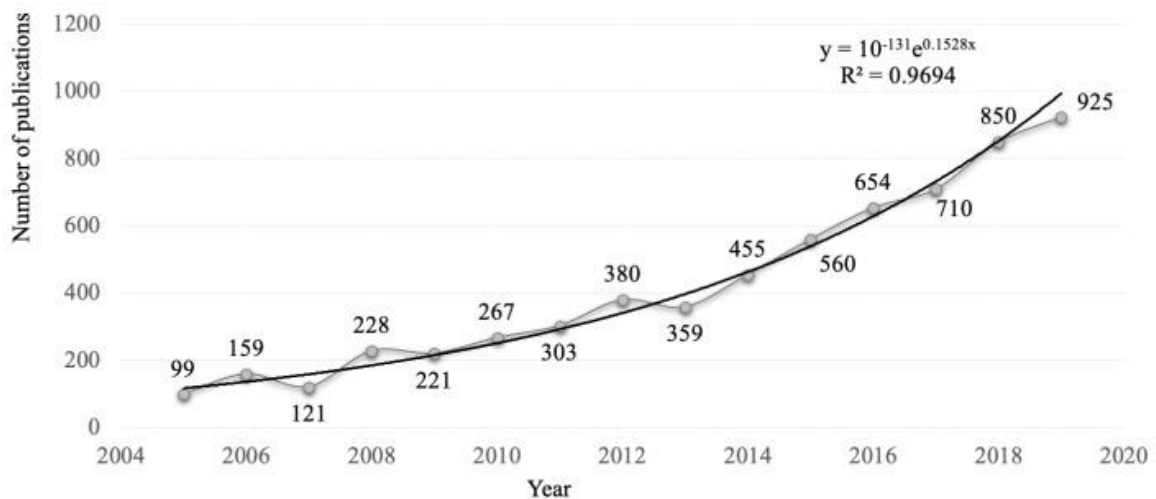


Figura 9 – Artigos publicados para previsão de consumos energéticos, 2005-2019, [32]

De acordo com o horizonte temporal, existem principalmente três categorias de estudos de previsão de consumos energéticos na literatura. A previsão de longo prazo considera um horizonte temporal de 5 a 20 anos e é utilizada para gestão de recursos e para investimentos em desenvolvimento. A previsão de médio prazo possui um horizonte temporal de um mês a cinco anos e é aplicada principalmente para planeamento de recursos e taxas de produção

de energia elétrica pelos distribuidores e produtores de energia. As previsões de curto prazo, de uma hora a uma semana, são utilizadas para estudos de análise de consumos, gestão de operações da rede de distribuição elétrica e estudos de modelos. Tendo em conta que a procura de energia varia consoante um conjunto variado de fatores, tais como as variáveis climáticas, as alturas do ano, os parâmetros socioeconómicos e os demográficos, prever o consumo de energia é uma tarefa importante e difícil [33]. Os modelos *black-box*, sendo alimentados com dados suficientes e parametrizados de uma maneira otimizada, permitem obter previsões precisas de consumo de energia. Todavia, o horizonte temporal de previsão é o que definirá o tipo de medida a implementar. Espera-se que políticas de economia de energia implementadas por governos surjam da aplicação de previsões de longo prazo, assim como medidas de gestão de recursos energéticos por parte dos produtores de energia surjam de previsões de médio prazo. Já as previsões de curto prazo potenciam um funcionamento otimizado dos sistemas de energia de edifícios através da previsão de necessidades, sendo também utilizadas para comparar diferentes modelos e configurações. A maioria dos artigos publicados sobre a previsão de consumos energéticos com recurso a abordagens baseadas em dados referem-se a previsões de curto prazo, considerando também que estas são as que atingem melhores resultados de previsão [34].

Os modelos baseados em dados podem ter duas abordagens distintas, uma de base estatística, onde se inserem os modelos de estatística clássica e outros mais recentes de análise e estudo de séries temporais, e outra de base de aprendizagem computacional onde se incluem as técnicas mais recentes de aprendizagem computacional. A abordagem de base estatística normalmente aplica uma função matemática pré-definida e tem mostrado bom desempenho para previsão de energia de médio e longo prazo, superando os métodos mais recentes quando aplicado a um grande número de edifícios [35]. Os métodos de aprendizagem computacional têm sido mais aplicados ultimamente e apresentam altos níveis de precisão na previsão de consumos energéticos em edifícios singulares, superando os métodos estatísticos tradicionais [36].

Dos métodos estatísticos podemos destacar os modelos ARIMA e o SARIMA, desenvolvidos por *Box e Jenkins* em 1976 [37]. A diferença entre estes dois modelos prende-se com o facto de o SARIMA ter em conta a sazonalidade da série temporal e o ARIMA apenas considerar a tendência da série. Estes dois modelos são frequentemente utilizados nos estudos de séries temporais univariadas. Ainda assim, existem variações destes modelos quando se adiciona uma variável exógena, nomeadamente os modelos ARIMAX e

SARIMAX, sendo que o X representa essa variável. A variável exógena pode ser, por exemplo, o nível de ocupação de um edifício ou a temperatura ao longo do tempo e tem como objetivo melhorar o desempenho dos modelos originais. De uma forma geral, o modelo ARIMA e as suas variações, ARIMA sazonal e com variáveis exógenas, são bastante utilizados em previsões de séries temporais uma vez que são relativamente fáceis de implementar e interpretar, e têm uma boa capacidade de lidar com diferentes tipos de variações e tendência nos dados, o que resulta em bons resultados de previsão em certas situações. Os modelos ARIMAX foram aplicados para várias aplicações de consumos energéticos em edifícios, como por exemplo, a modelação e previsão de temperaturas ambiente dos espaços, otimização de operações dos sistemas de aquecimento e refrigeração e previsão e controlo do consumo energético num edifício governamental em horas de pico. Uma vez que a ocupação de um edifício está correlacionada com o consumo energético do mesmo, as componentes da série temporal de consumo energético possuem implicitamente esse efeito. Contudo, a adição desta variável como variável exógena num modelo ARIMA apresenta melhorias ao modelo original que poderão ser justificáveis pelo facto de os dados de ocupação de edifícios serem geralmente fáceis de obter e quantificar [38].

Dentro dos métodos de base estatística apresentados, o modelo SARIMAX é a versão mais avançada destes modelos e tem a capacidade de minimizar os valores de erro e apresentar bons resultados de previsão mesmo para horizontes de longo prazo de previsão. Este modelo, aplicado a séries temporais não estacionárias possui a capacidade de diferenciar a série temporal e a variável exógena antes de proceder à estimação do mesmo [39].

O modelo SARIMAX foi aplicado a um caso de estudo da Arábia Saudita que teve como objetivo analisar o comportamento do sistema de energia a longo prazo [40]. Foram recolhidos dados históricos em intervalos trimestrais durante o período de 1980 a 2020 para prever o consumo de eletricidade, geração, potência de pico e capacidade instalada para um período de 30 anos (2021 a 2050). A pesquisa efetuada forneceu evidências de que esta abordagem de base estatística, apresentou resultados superiores aos métodos mais complexos de aprendizagem computacional, como as árvores de decisão e as redes neuronais, concretamente uma rede MLP – *Multilayer perceptron* e uma LSTM – *Long short-term memory* para este caso de estudo. Os autores realçam ainda a necessidade de testar ativamente várias combinações de modelos e métodos de preparação de dados para determinar qual o método mais eficaz, para além de definir e usar cuidadosamente uma estratégia de preparação de dados e de modelos. Os resultados deste estudo revelaram que o

modelo SARIMAX superou os outros modelos referenciados em termos de precisão de previsão, sobreajuste dos dados (*overfitting*), eliminação de redundância e tempo de treino do modelo e de execução dos testes, o que prova o seu desempenho notável. Contudo, os autores do estudo referem que este modelo pode ainda ser aprimorado através da otimização dos parâmetros e ao utilizar um conjunto de dados de intervalo inferior, como o histórico mensal.

Em 2016 foi realizado um estudo de previsão de potência de energia elétrica em Abu Dhabi nos Emirados Árabes Unidos, com o objetivo de realizar previsões de curto prazo para melhorar a gestão das operações e planejamento do sistema de energia da cidade com vista a atingir melhores níveis de estabilidade e rentabilidade do sistema [41]. Este estudo baseou-se na recolha de dados horários de eletricidade pelo sistema de serviços públicos da cidade, bem como dados meteorológicos, incluindo a temperatura, a humidade específica, a radiação solar e a velocidade do vento. Os dados recolhidos vão desde o segundo semestre de 2009 até ao final no primeiro semestre de 2011, desde 1 de julho de 2009 até 30 de junho de 2011, sendo os dados de eletricidade maioritariamente imputáveis a edifícios residenciais, comerciais e institucionais e os dados meteorológicos recolhidos através de estações meteorológicas. Foram utilizados e comparados três métodos, nomeadamente, um modelo ARIMA, uma rede neuronal artificial e um modelo que os autores denominaram de *transfer function*, o qual corresponde a um modelo ARIMAX. Dos modelos analisados, concluiu-se que a inclusão das variáveis exógenas no modelo ARIMAX melhorou o desempenho dos modelos. Quando comparados, o modelo ARIMAX revelou apresentar melhores resultados para todos os horizontes temporais testados, neste caso, previsões de um dia, dois dias e uma semana. Com um cenário mais realista, onde as variáveis exógenas não são conhecidas ao longo do horizonte de previsão e devem ser previstas antes de utilizadas, o modelo ARIMAX apresentou melhores resultados do que a rede neuronal para todos os horizontes de previsão testados. Todas as variáveis meteorológicas foram estimadas com base nos valores do passado de cada variável exógena, tendo sido posteriormente inseridos no modelo de previsão. A rede neuronal apresentou previsões superiores das variáveis meteorológicas exógenas quando comparado com o modelo ARIMAX, enquanto que este último apresentou um desempenho significativamente superior na previsão da eletricidade para todos os horizontes de previsão.

Para além dos métodos apresentados anteriormente, outros de base estatística, como as regressões lineares simples e múltiplas foram aplicadas para o estudo de séries temporais de

consumos energéticos em edifícios. Num caso de estudo onde foi comparada a utilização de uma rede *MLP – Multilayer Perceptron* com a utilização de um método de regressão linear múltipla observou-se que ambas as técnicas apresentaram resultados satisfatórios, tendo a rede neuronal apresentado um melhor desempenho [42]. Neste caso de estudo, a previsão de consumos energéticos refere-se a um edifício habitacional de três andares e seis frações, tendo os autores discriminado a energia utilizada pelos diferentes sistemas de cada fração, nomeadamente iluminação, aquecimento e arrefecimento dos espaços, aquecimento de água, ventilação, refrigeração e outros de uso diverso. Foi determinado um perfil de utilização de equipamentos e ocupação dos espaços onde o consumo de energia resultante foi depois comparado com o valor registado nas faturas de energia elétrica, sendo assim validada a desfragmentação do consumo entre as diferentes aplicações. Desta forma, os autores conseguiram criar um *dataset* de consumo energético horário para as diferentes aplicações, sendo que, a energia total consumida foi posteriormente validada com a fatura energética, permitindo assim uma previsão específica por sistema. Para além disso, foram ainda recolhidos dados de parâmetros físicos do edifício para serem utilizados nos modelos, nomeadamente, o ano de construção, a área da fração, materiais dos pavimentos e das paredes e valores de pé-direito. Todos os dados recolhidos e gerados foram utilizados para a criação dos modelos de regressão linear múltipla e para a rede neuronal com o objetivo de fazer previsões de consumos energéticos por aplicação e global para um horizonte temporal de médio prazo, nomeadamente, para o mês seguinte. Para a rede neuronal o *dataset* foi separado em 80% de treino e 20% de teste. Os dados de previsão foram depois comparados com os dados reais de consumo e verificou-se que a rede neuronal apresentou os resultados com variações mais baixas.

Um estudo semelhante foi realizado em edifícios de escritórios no Chile para a previsão do consumo global de energia, específico de aquecimento e arrefecimento e quantidade de emissões de dióxido de carbono, recorrendo a regressão linear e a redes neuronais [43]. Neste estudo foram recolhidas oito variáveis relacionadas com os parâmetros físicos dos edifícios, sendo que, as necessidades e os consumos energéticos, assim como as emissões de CO₂ foram determinadas através de fórmulas estabelecidas na norma ISO 13790:2008 que considera coeficientes de desempenho e outras rácios relacionados com emissões e eficiências dos sistemas. Deve referir-se que este estudo procurou inicialmente gerar uma base de dados de diferentes edifícios, isto é, edifícios tipicamente considerados edifícios de escritório localizados no Chile sujeitos às condições climáticas da região. Assim,

considerando vários casos possíveis de edifícios de escritórios e gerando todos os dados para cada um deles, o *dataset* deste estudo ficou composto por 77 000 casos em que cada um deles possui informação diferente relativa aos dados anteriormente referidos. Estes conjuntos de dados não são compostos por uma só variável, mas sim por uma complexa interação de vários parâmetros que suportam tanto relações lineares como não lineares. Os dados são posteriormente divididos em treino, validação e teste com uma distribuição respectiva de 50%, 25% e 25%. Aplicados os dois métodos a este conjunto de dados, nomeadamente, regressões lineares múltiplas e uma MLP, os autores concluíram que os modelos criados a partir da rede neuronal têm uma maior precisão para substituir os procedimentos de cálculo estabelecidos na norma ISO 13790:2008, o que leva a crer que a implementação deste método avançado de aprendizagem computacional poderá ser aplicada para o panorama legislativo Chileno. Os modelos de redes neuronais podem fornecer informações simplificadas e métodos de cálculo mais rápidos, assim como outras utilidades na definição de políticas e objetivos relacionados com a eficiência energética em edifícios. Para além disso, este estudo possui um potencial de replicação considerando condições climáticas distintas e outras tipologias de edifícios.

Os métodos apresentados anteriormente objetivaram uma comparação dos métodos de base estatística com os de aprendizagem computacional informática. Contudo, no que toca a estes últimos, foram realizados vários estudos que os analisam e comparam para diferentes situações de previsão de consumos energéticos em edifícios. Num desses estudos foram aplicadas três técnicas de aprendizagem computacional, nomeadamente uma rede MLP, uma SVM – *Support Vector Machine* – máquina de suporte vetorial regressivo e uma rede neuronal LSTM, para previsão de consumos energéticos dos sistemas de aquecimentos, arrefecimento e iluminação e para previsão de produção elétrica de um sistema de painéis solares fotovoltaicos [44]. Foi modelado um edifício de quatro andares com recurso a modelos físicos para gerar as potências elétricas necessárias durante um ano, tendo em conta as variações climáticas da região. Assim, os dados de entrada nos algoritmos são divididos em três grupos, nomeadamente, as variáveis relacionadas com o tempo, sendo estas a temperatura ambiente, a humidade relativa, a quantidade de radiação e o rácio de cobertura de nuvens; variáveis relacionadas com a operação do edifício onde se enquadram os perfis de ocupação e utilização de equipamentos e iluminação e, por fim, as variáveis relacionadas com a energia, nomeadamente, a potência e a temperatura do sistema solar fotovoltaico e os consumos energéticos dos sistemas de aquecimento, arrefecimento e iluminação. Quanto aos

dados de saída, pretendeu-se prever a quantidade de energia relacionada com os sistemas mencionados de energia. Neste estudo, o modelo que utiliza a rede MLP foi o que apresentou melhores resultados e um tempo computacional de execução mais baixo.

Similar ao anterior, outro estudo foi realizado com o objetivo de comparar três tipos de modelos de aprendizagem computacional para previsão de consumo de energia, nomeadamente uma rede MLP, os SVM e modelos híbridos [45]. Este estudo pretendeu analisar a complexidade dos modelos, a precisão dos resultados e outros parâmetros dos modelos estudados, analisando uma vasta gama de tipologias de edifícios, especificamente residenciais, comerciais, governamentais e educacionais e diferentes horizontes temporais como dados de entrada e saída dos modelos. Com este estudo, os autores concluíram que os métodos de aprendizagem computacional, quando comparados com outros métodos de base estatística ou os modelos físicos, tendem a apresentar resultados de previsão superiores em termos de precisão e com um tempo de parametrização de modelos inferior. Quando aplicados a edifícios residenciais, atendendo que estes edifícios possuem dinâmicas individuais que dependem de cada morador, torna-se muito difícil prever o consumo de energia de uma forma confiável. Os diferentes métodos aplicados possuem diferentes recursos que são utilizados em diferentes situações, possuindo assim respetivos pontos fortes e pontos fracos, o que leva a que, para diferentes casos, os modelos mais adequados precisam de ser selecionados e parametrizados para o caso concreto que se pretende analisar.

Num outro estudo, foi realizada uma revisão bibliográfica sobre técnicas de aprendizagem computacional para previsão de consumos energéticos em edifícios através da utilização de técnicas de *text mining* com o objetivo de entender os fatores mais relevantes que afetam o consumo de energia e as melhores técnicas utilizadas para classificar e prever o consumo de energia nos edifícios [46]. Foram analisados no total 822 publicações entre 2013 e 2020 que constam em algumas bibliotecas digitais, tais como a *IEEE Xplore*, *Science Direct*, *Springer*, *Scopus* e *Web of Science*. Deste estudo concluiu-se que alguns dos fatores predominantemente relacionados com o consumo energético são o perfil de ocupação, os consumos relacionados com o aquecimento e arrefecimento dos espaços e as condições climáticas. Quanto às técnicas de previsão, as mais utilizadas são os algoritmos de *Random Forest* (Florestas Aleatórias), redes neuronais, máquinas de suporte vetorial e modelos de regressão.

A aplicação de *Deep Learning* para previsão de consumos energéticos em edifícios tem apresentado resultados positivos em vários casos de estudo. Assim, com o objetivo de aumentar a precisão da previsão de consumo de energia de um edifício, foi proposta uma tecnologia de *Extreme Deep Learning* [47]. Esta abordagem combina *Stacked Autoencoders* com *Extreme Learning Machines* de forma a retirar partido das vantagens de cada uma destas técnicas. Os *Stacked Autoencoders* são utilizados para retirar as características principais do conjunto de dados, enquanto que as *Extreme Learning Machines* são utilizadas para fazer as previsões. Outras variáveis de entrada definidas neste estudo foram analisadas através dos gráficos de autocorrelação e de autocorrelação parcial e foi feita, posteriormente, uma comparação com tecnologias de aprendizagem computacional mais simples, tais como uma rede neuronal, SVM e regressões lineares múltiplas. Os dados deste estudo foram recolhidos em intervalos de 15 minutos de um edifício comercial da cidade de Fremont nos Estados Unidos da América e, posteriormente agregados em intervalos de 30 e 60 minutos de forma a realizar duas previsões de médio prazo com as resoluções temporais definidas. Os resultados deste estudo demonstraram que o método proposto superou os outros métodos nas duas resoluções temporais utilizadas. A principal diferença que fez este método apresentar melhores resultados, do ponto de vista dos autores, prende-se com o facto de que este não necessita de tanta afinação do modelo e possui uma arquitetura de aprendizagem mais profunda, assim como estratégias de aprendizagem mais otimizadas, o que permite uma melhor extração de características do consumo de energia do edifício, reduz o tempo de treino do modelo e melhora a precisão de previsão.

Uma pesquisa semelhante desenvolvida com edifícios residenciais do Reino Unido teve como objetivo realizar a previsão do consumo de energia anual com recurso a nove algoritmos de aprendizagem computacional a fim de os comparar e identificar o melhor algoritmo para o problema em estudo. Os algoritmos de aprendizagem computacional utilizados foram uma rede MLP, *Gradient Boosting*, *K-Nearest Neighbors*, uma rede neuronal profunda (*Deep Neural Network*), Florestas Aleatórias, Árvores de Decisão, *Stacking*, Máquinas de Suporte Vetorial e Regressão Linear. Os dados utilizados foram recolhidos através de entidades públicas de serviços e referem-se a cerca de 5000 edifícios. Neste estudo, as variáveis dependentes podem ser divididas em três grupos: as variáveis meteorológicas, as variáveis de construção e os dados de consumo energético. Quanto aos dados meteorológicos, estes contêm informação relativamente à temperatura, velocidade do vento e humidade relativa durante o ano de 2020. Quanto às variáveis de construção, foram

analisadas várias como, por exemplo, a altura dos andares a descrição dos telhados e das paredes e o número de quartos habitáveis durante o ano de estudo em referência. A variável dependente é a energia consumida por cada edifício. As variáveis de entrada, como os tipos de paredes e de janelas são consideradas variáveis importantes, pois têm efeito sobre o consumo de energia dos edifícios e a sua seleção apropriada poderá reduzir consideravelmente o consumo de energia. Estas variáveis, numa fase de projeto de novos edifícios, podem ser modificadas para alcançar edifícios mais sustentáveis e eficientes energeticamente. Neste estudo conclui-se que o melhor modelo para previsão de consumo de energia para um horizonte temporal de um ano, isto é, o modelo que apresentou as métricas de erro mais baixas e a métrica de explicação da variância dos dados mais alta, foi o algoritmo de rede neuronal profunda, seguindo-se a MLP e o *gradient boosting* em terceiro lugar [48]. O alto desempenho destes modelos apresenta uma motivação para a sua utilização na fase inicial de projetos de edifícios novos, de forma a prestarem auxílio na tomada de decisões de âmbito energético, material e estrutural.

Como demonstrado neste texto, as tecnologias de aprendizagem computacional têm sido aplicadas cada vez com mais regularidade na previsão de consumos de energia, uma vez que o rápido desenvolvimento de tecnologias modernas, incluindo sensores, comunicação em rede, transmissões sem fios e dispositivos inteligentes levou a uma acumulação enorme de dados, sem os quais estas tecnologias não poderiam ser aplicadas. As redes neuronais *feedforward* são ferramentas poderosas para a realização desta tarefa. Contudo, estas requerem uma escolha apropriada da sua estrutura e o ajuste preciso dos seus diversos parâmetros, sendo que o desempenho pode não ser garantido devido ao problema dos mínimos locais, e podem até ser superadas por modelos de regressões lineares múltiplas. Também outras técnicas de aprendizagem computacional, como o SVM e regressões baseadas em métodos gaussianos, possuem a capacidade de apresentar bons resultados, superando em alguns casos as redes neuronais *feedforward*, uma vez que não utilizam tantos parâmetros, logo não carecem da necessidade de ajuste tão rigoroso. Uma publicação que comparou os modelos mencionados anteriormente evidencia a grande dificuldade em eleger um melhor modelo de aprendizagem computacional, uma vez que na literatura entende-se que vários modelos fornecem previsões com uma precisão razoável quando alimentados com dados suficientes e parametrizados de forma adequada. Logo, torna-se imperativo analisar minuciosamente a natureza dos dados disponíveis e a aplicação que se pretende dar ao modelo. Por exemplo, redes neuronais *feedforward* fornecem previsões de curto prazo

rápidas e precisas quando alimentadas com os dados de temperatura e humidade, enquanto que outros métodos, como as regressões baseadas em métodos gaussianos, são mais benéficos para previsões de longo prazo e quando há incerteza nas variáveis de entrada [49]. Para além da modelação de um edifício para previsão de consumos, o agrupamento de edifícios através de *clustering* com base em vários parâmetros de entrada facilita e aprimora as metodologias de previsão de uma forma notável. A utilização de edifícios de referência com boas características energéticas leva a uma rotulagem energética mais precisa, o que permite simultaneamente estimar necessidades de construções futuras adequadas a um desenvolvimento sustentável, à medida que a preocupação global com a energia aumenta e muitos países se esforçam para regular as indústrias consumidoras de energia.

De todos os modelos baseados em dados apresentados não é possível identificar um melhor para todas as situações, uma vez que os resultados obtidos variam significativamente de acordo com muitas variáveis. Algumas destas estão relacionadas com a sua recolha e disponibilidade, como os dados de consumo de energia e a sua resolução temporal, assim como os mesmos aspetos para as variáveis exógenas que se relacionam intrinsecamente com o consumo energético de um edifício. Outra variável que afeta consideravelmente o resultado obtido é o horizonte temporal de previsão, sendo que esta poderá apontar para um modelo a utilizar. Para além disso, cada modelo encontra-se sujeito a várias configurações e parametrizações que terão que ser testadas de forma exaustiva de forma a encontrar a melhor configuração para os dados disponíveis e o problema em questão.

4. Metodologia, Modelos e Métricas

A eficácia da previsão dos consumos energéticos depende significativamente da abordagem aplicada, assim como dos modelos utilizados. Assim, de forma a realizar as previsões de consumos energéticos do edifício em estudo, foi desenvolvida uma metodologia a aplicar que engloba as fases, desde a recolha e tratamento de dados, até à obtenção das previsões e avaliação dos modelos através das métricas de erro mais apropriadas.

O presente capítulo pretende explicar a metodologia desenvolvida, apresentar detalhadamente os modelos selecionados, dividindo estes em dois grupos, nomeadamente os modelos de base estatística e modelos de aprendizagem computacional e, por fim, selecionar as métricas de erro que permitem comparar os resultados obtidos de uma forma imparcial.

4.1. Metodologia

No decorrer do trabalho foi desenvolvida uma metodologia a seguir na criação e teste dos modelos concebidos, a qual se apresenta na figura seguinte.

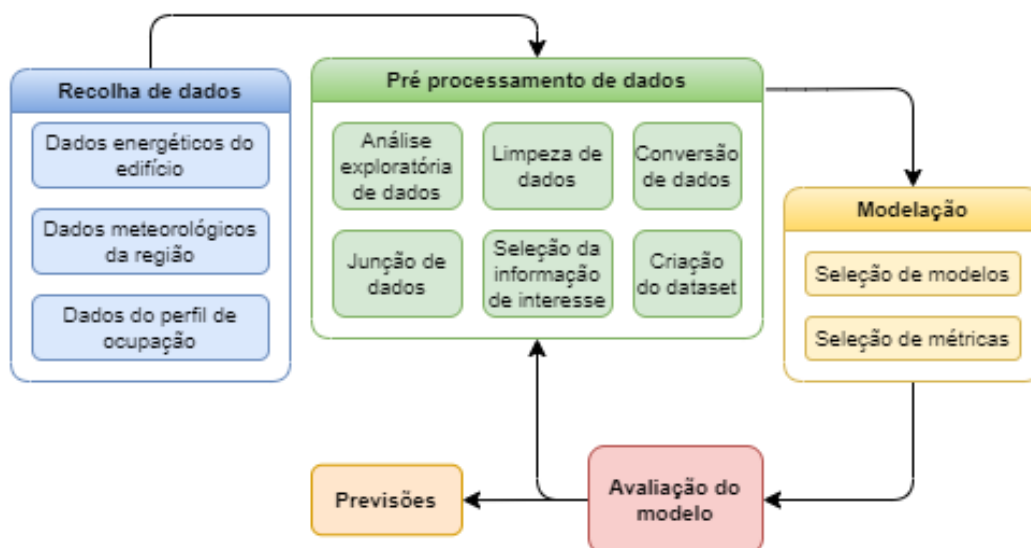


Figura 10 - Metodologia no tratamento e análise dos dados

Esta metodologia foi dividida em cinco atividades principais apresentadas seguidamente por ordem: Recolha de dados, Pré processamento de dados, Modelação, Avaliação do modelo e Previsões. Contudo, esta metodologia foi desenvolvida tendo em conta uma perspetiva

cíclica das atividades realizadas, uma vez que, a criação e avaliação dos modelos exige uma repetição de atividades com o objetivo de obter o melhor modelo. Como é visível na Figura 10, a recolha de dados refere-se à forma como os dados foram obtidos, e esta foi uma atividade realizada apenas uma vez, tendo sido posteriormente dada como concluída. Relativamente ao pré processamento de dados, esta atividade inclui vários procedimentos, nomeadamente a análise exploratória, a limpeza, a conversão e a junção dos dados, a seleção da informação de interesse e a criação do *dataset* final para aplicar ao modelo. No que diz respeito à modelação, existem apenas duas tarefas, tendo a seleção das métricas sido definida logo ao início e apresentada no decorrer deste capítulo, e a seleção dos modelos, onde se incluem as atividades de parametrização de cada modelo, isto é, o ajuste dos hiperparâmetros que constituem cada modelo aplicado.

4.2. Modelos de Base Estatística

Vários são os modelos de base estatística que se têm vindo a utilizar ao longo do tempo para modelar uma série temporal. Alguns destes modelos já foram apresentados neste trabalho e passam pela decomposição da série temporal nas suas diferentes componentes. Outros, mais complexos, surgem da junção de modelos mais simples, como é o caso do modelo ARIMA (*Autoregressive Integrated Moving Average*), que provem da junção do modelo AR (*Autoregressive*) com o modelo MA (*Moving Average*) e o fator *I* (estacionariedade obtida por diferenciação de ordem *I*). Nesta secção pretende-se dar a conhecer ao leitor os modelos SARIMA (*Seasonal Autoregressive Integrated Moving-Average*), o qual proveio do ARIMA, e SARIMAX (*Seasonal Autoregressive Integrated Moving Average with Exogenous Variables*), sendo este uma evolução do anterior.

4.2.1. SARIMA

Na análise de séries temporais, a previsão de valores futuros é uma tarefa realizada através da estimação de um modelo que represente da melhor forma possível os padrões da série. O modelo SARIMA deriva do modelo ARIMA, sendo que o primeiro tem em conta as componentes sazonais da série, logo, a sua utilização é mais indicada para séries temporais onde a componente sazonal esteja presente. O modelo ARIMA é composto pelas componentes p , d e q , sendo estas correspondentes às seguintes partes: autorregressiva, ordem de diferenciação e médias móveis. Já o modelo SARIMA, denominado na língua portuguesa como Modelo Sazonal Autorregressivo Integrado de Médias Móveis, é composto pelas componentes do modelo ARIMA e pelas mesmas representativas da parte sazonal e

representadas pelas letras P , D e Q , com a adição de uma outra, denominada “ m ” e que representa o período da sazonalidade. Desta forma, estes modelos são representados conforme a seguinte figura, onde são substituídas as letras representativas das diferentes componentes por números determinados através da análise dos gráficos de autocorrelação e de autocorrelação parcial e por ordens de diferenciação para tornar a série estacionária.

$$\text{SARIMA} \left(\underbrace{(p, d, q)}_{\text{non-seasonal}} \left(\underbrace{(P, D, Q)}_m \right) \right)$$

Figura 11 – Representação das diferentes ordens de um modelo SARIMA [50]

Quanto às diferentes componentes:

- Componente autorregressiva (p e P): Procura examinar a relação entre uma observação e um certo número de observações desfasadas (em *lags* diferentes).
- Componente de diferenciação (d e D): Representa a ordem de diferenciação. Como já foi explicado anteriormente, a aplicação destes modelos para análise de séries temporais exige que a série seja estacionária e, para isso, poderão ser aplicadas ordens de diferenciação de forma a remover as componentes de tendência e sazonalidade para tornar a série estacionária, isto é, tornar a média e a variância praticamente constantes ao longo do tempo.
- Componente de média móvel (q e Q): Considera a dependência entre uma observação e o erro residual de um modelo de médias móveis aplicado a observações desfasadas.
- Componente sazonal (m): Representa o período da sazonalidade de uma série temporal ou a duração do período sazonal que se observa na série.

Os modelos SARIMA possuem flexibilidade e são versáteis, o que os torna capazes de lidar com uma ampla gama de séries temporais com tendência e sazonalidade. Estes modelos são também facilmente interpretáveis, permitindo uma melhor compreensão da dinâmica subjacente da série temporal. Contudo, estes modelos assumem linearidade e podem ter um desempenho inferior quando confrontados com séries temporais complexas e não lineares.

4.2.2. SARIMAX

Um outro modelo semelhante ao apresentado anteriormente é o modelo SARIMAX. Este modelo estende os recursos do modelo SARIMA ao incorporar variáveis exógenas. Estas variáveis são geralmente fatores externos à série temporal que se encontram em análise e

que a podem influenciar. A inclusão de variáveis exógenas pode aumentar a precisão das previsões ao capturar informações adicionais que podem ser correlacionadas com a série temporal em análise. Todavia, a inclusão de variáveis exógenas aumenta consideravelmente a complexidade do modelo o que exige maiores requisitos computacionais.

4.3. Modelos de Aprendizagem Computacional

Como já foi abordado, as redes neuronais podem assumir vários tipos, sendo que de seguida são abordados dois: as MLP – *Multy Layer Perceptron* e as RNN – *Recurrent Neural Network*. As MLP são redes neuronais *feedforward* constituídas por várias camadas, cada uma delas constituída por neurónios ligados entre si de uma camada para a outra, como exemplifica a Figura 12. As redes neuronais são técnicas de aprendizagem computacional, sendo que a quantidade de camadas escondidas (*hidden layers*) é que determina se a rede neuronal é de aprendizagem profunda ou de aprendizagem computacional, sendo que a primeira é uma subárea da segunda.

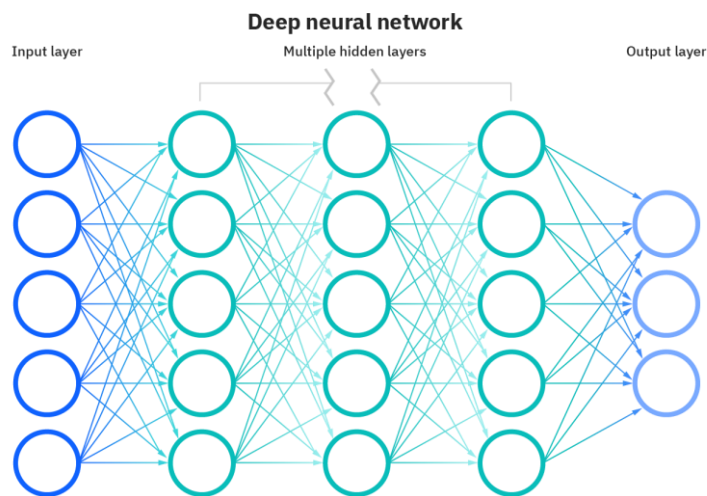


Figura 12 – Exemplo de uma rede neuronal artificial [51]

As MLP constituem-se uma arquitetura de rede neuronal que possui a capacidade de capturar padrões complexos e relações não lineares. Contudo, estas redes não possuem nenhum mecanismo para lidar com dados sequenciais, isto é, cada entrada de uma MLP é tratada de forma independente, não considerando nenhum relacionamento sequencial com entradas anteriores. Assim, estas redes são mais adequadas para tarefas em que a ordem dos elementos de entrada é irrelevante, como classificação de imagens. Tendo em conta esta característica, e atendendo ao objetivo deste trabalho, assim como ao facto que os conceitos relacionados

com as MLP foram já explicados no decorrer deste documento, a presente secção pretende explicar com mais detalhe como funcionam as RNN e as redes neuronais que advieram desta arquitetura de rede, nomeadamente as GRU – *Gated Recurrent Unit* e as LSTM - *Long Short-Term Memory*.

4.3.1. RNN

As RNN são redes neuronais que foram criadas com o intuito de lidar com dados sequenciais e séries temporais, estando estas representadas na Figura 13. Estas redes possuem a capacidade de capturar dependências relacionadas com o tempo, uma vez que, possuem um elemento de memória que as permite manter e processar informação ao longo do período temporal em análise.

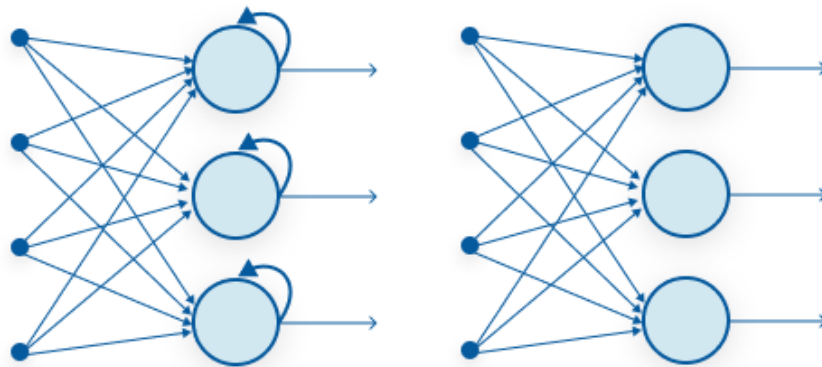


Figura 13 – Esquerda: Representação das entradas e saídas dos neurónios numa RNN. Direita: Representação das entradas e saídas dos neurónios numa MLP [52]

Dentro das RNN destacam-se as RNN simples, as LSTM e as GRU, sendo que estas últimas originaram a partir da primeira. Os neurónios destas redes possuem uma ligação recorrente que permite que certas informações fluam de uma etapa para a outra. Esta ligação cria um ciclo dentro da rede neuronal que permite que a informação do passado seja retida e utilizada no presente. Embora possuam diferentes arquiteturas, a forma como a informação flui dentro da rede é o fator de semelhança entre estas. Assim, a arquitetura dentro dos neurónios de cada camada de uma RNN é a componente chave da rede neuronal.

O neurónio de um RNN simples produz um estado oculto, sendo que este atua como uma unidade de memória. Desta forma, o neurónio de uma RNN recebe a entrada da etapa de tempo atual, combina-a com as informações da etapa de tempo anterior e produz uma saída e um novo estado oculto, tal como ilustrado na Figura 14. Este processo recursivo permite que a rede capture dependências ao longo do tempo.

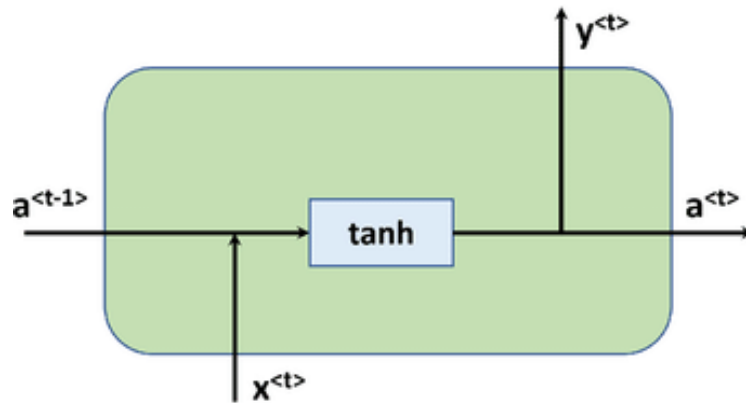


Figura 14 – Exemplo de um neurônio numa RNN [53]

Tendo em conta a capacidade das RNN de lidarem com dados sequencias e reterem informações passadas, estas são adequadas ao processamento de séries temporais [53]. Contudo, a complexidade deste tipo de redes faz com que o poder computacional necessário para as processar seja superior a redes mais simples, aumentando desta forma o tempo de treino destas redes.

4.3.2. LSTM

As LSTM são um tipo de rede neuronal recursiva que surgiu da evolução das RNN simples. Este tipo de rede neuronal possui um mecanismo de memória especializado, permitindo a captura de dependências de longo e curto prazo em dados sequenciais e de lidar com o problema de *vanishing gradient* que as RNN simples apresentam. Assim, as LSTM foram projetadas para superar os desafios de capturar dependências de longo prazo, possuindo a capacidade de armazenar, ler e atualizar informações seletivamente por períodos prolongados o que lhes dá a capacidade de reter informações vitais e descartar informações irrelevantes.

Este tipo de rede neuronal introduz um mecanismo de *gating* para atualizar e redefinir o estado oculto calculado nos neurónios. Esse mecanismo permite que as LSTM retenham informações importantes em sequências mais longas sem ser sobrecarregada por ruído ou detalhes irrelevantes. Desta forma, a ideia fundamental por detrás da LSTM é utilizar unidades de *gating* que regulam o fluxo de informações dentro da rede.

A configuração das LSTM encontra-se apresentada na Figura 15, sendo desenhada através de *gates* que regulam o fluxo de informação, permitindo e inibindo seletivamente a passagem de dados. As três *gates* principais numa LSTM são: a *input gate*, a *forget gate* e a *output gate*. A primeira *gate* controla até que ponto novas informações devem ser incorporadas à

memória, decidindo quais elementos da entrada atual são importantes para atualizar o estado da memória. A segunda *gate* determina quais as informações do intervalo de tempo anterior devem ser descartadas, permitindo que a rede foque nas informações passadas mais relevantes. Quanto à terceira *gate*, esta determina o novo estado oculto do neurónio para a próxima entrada de dados e a respetiva saída de dados.

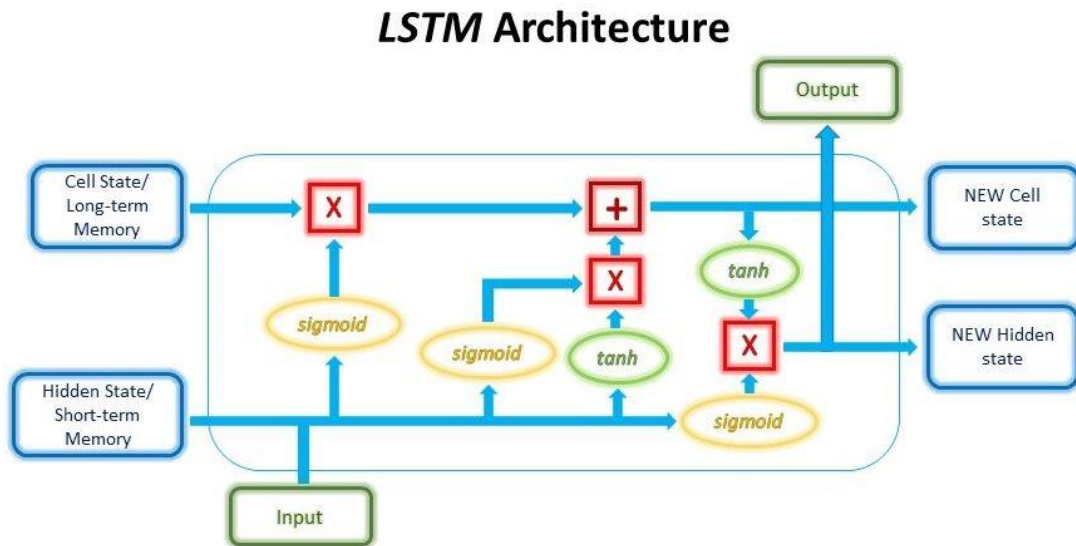


Figura 15 - Exemplo de um neurónio numa LSTM [54]

Estas redes neuronais, embora possuam uma maior complexidade computacional e consequente maior tempo de treino, possuem a capacidade de capturar as dependências de longo prazo de dados sequenciais, o que as torna ferramentas interessantes na análise de séries temporais, nomeadamente nas tarefas de previsão e deteção de anomalias. Contudo, estas redes são conhecidas por serem modelos de caixa preta, o que significa que carecem de interpretabilidade, o qual pode ser uma preocupação em áreas de estudo específicas [55].

4.3.3. GRU

As GRU são outro tipo de rede neuronal recursiva que surgiram da evolução das LSTM. Este tipo de rede neuronal é muito semelhante às LSTM, possuindo as mesmas vantagens no que toca à modelação das dependências de longo e curto prazo nos dados. Para além disso, este tipo de rede neuronal é também composto por um mecanismo de *gating* para atualizar e redefinir o estado oculto calculado nos neurónios.

Uma célula GRU consiste em dois componentes principais que geram a saída do neurónio, nomeadamente o *reset gate* e o *update gate*, como apresentado na Figura 16. O *reset gate* determina a quantidade de informação passada que deve ser esquecida, enquanto que o

update gate determina a quantidade de informação passada que deve ser transportada para o intervalo de tempo atual.

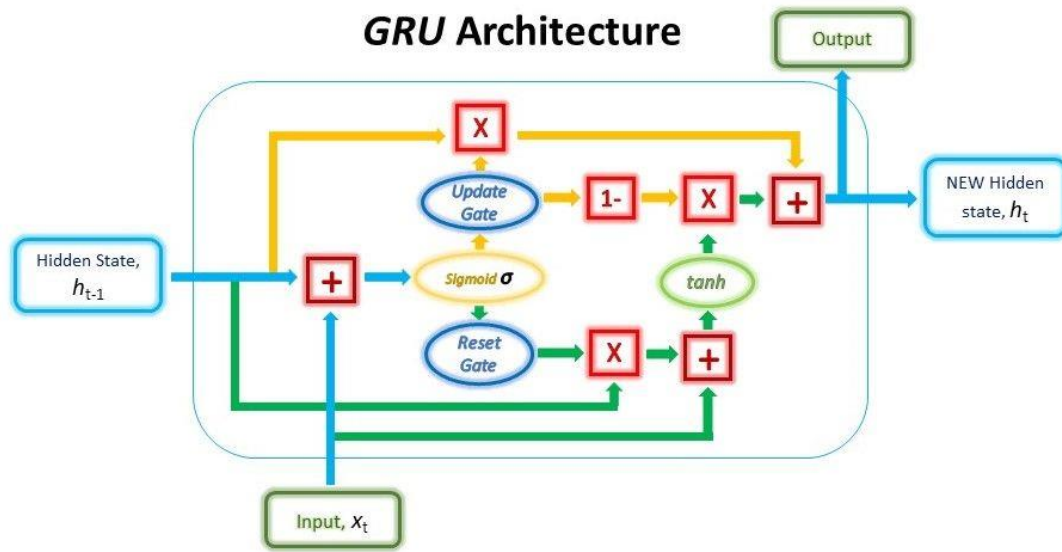


Figura 16 - Exemplo de um neurônio numa GRU [54]

A vantagem deste tipo de rede neuronal encontra-se relacionada com a estrutura simplificada da célula GRU, o que por sua vez, diminui a complexidade computacional e o tempo de treino da rede neuronal. Contudo, este tipo de rede também é sensível à inicialização dos pesos e ainda apresenta dificuldades a capturar dependências muito longas [56]. Pese embora os factos apresentados, este tipo de rede neuronal é apropriado para a análise de séries temporais no que toca a tarefas de previsão e deteção de anomalias.

4.4. Métricas de Erro

As métricas de erro são utilizadas para avaliar os modelos e quantificar a precisão das previsões. Neste trabalho foram seleccionadas cinco métricas de erro, nomeadamente MSE – *Mean Squared Error* (Erro Quadrático Médio), RMSE – *Root Mean Squared Error* (Raiz Quadrada do Erro Quadrático Médio), MAPE – *Mean Absolute Percentage Error* (Erro Percentual Absoluto Médio), MAE – *Mean Absolute Error* (Erro Absoluto Médio) e o R^2 – Coeficiente de determinação. De seguida, são apresentadas as equações matemáticas de cada métrica mencionada.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 ;$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} ;$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|;$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|;$$

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2};$$

Y_i – Valor real para a observação i

SSR – *Sum Squared Regression*

\hat{Y}_i – Valor previsto para a observação i

SST – *Total Sum of Squares*

n – Número de observações

\bar{Y}_i – Valor médio para a observação i

Estas métricas são frequentemente utilizadas na literatura para avaliar a qualidade das previsões obtidas pelos modelos matemáticos, sendo que cada uma destas fornece representações distintas dos valores obtidos. Nomeadamente, o MSE penaliza mais os erros grandes do que os erros pequenos, atendendo à operação de elevação ao quadrado. O RMSE é expresso nas unidades da variável que se está a prever e é preferido quando se quer interpretar a métrica de erro na escala original dos dados. Já o MAPE e o MAE são métricas que não são tão sensíveis a *outliers* [57], sendo importantes para entender o erro relativo e o erro absoluto das previsões com os valores reais. Quanto ao coeficiente de determinação, este é uma métrica que varia entre 0 e 1, e serve para averiguar o quão bem o modelo está ajustado aos dados originais, ou, por outras palavras, quanto da variância da variável dependente é explicada pelas variáveis independentes.

Esta página foi propositadamente deixada em branco

5. Descrição e Tratamento dos Dados

No presente capítulo são descritos todos os aspetos relacionados com a recolha e o tratamento dos dados, assim como as operações de limpeza realizadas. Para além disso, são mencionados outros aspetos relevantes relacionados com o edifício de onde os dados tiveram origem, assim como explicadas as variáveis do presente estudo.

5.1. Descrição dos Dados

Os dados de consumo energético aqui apresentados referem-se a um edifício de serviços de prestação de formação profissional nas áreas da metalurgia e metalomecânica localizado na Marinha Grande que pertence ao distrito de Leiria. O edifício em questão, como já mencionado anteriormente, é o CENFIM da Marinha Grande e este é constituído por salas de formação teóricas, teórico-práticas, laboratórios e oficinas. Excluindo os consumos energéticos gerais como a climatização e a iluminação, as salas de formação teóricas possuem apenas videoprojectores como equipamentos, as salas teórico-práticas possuem para além de videoprojector vários computadores e monitores para lecionar disciplinas como CAD/CAM, os laboratórios são de eletricidade, pneumática e robótica, possuindo equipamentos didáticos desse âmbito e as oficinas dividem-se em 5 grupos, nomeadamente mecânica convencional, mecânica CNC, fabricação aditiva, soldadura e serralharia. Grande parte do consumo energético do edifício provém da utilização das oficinas, uma vez que encontramos uma grande variedade de equipamentos que possuem uma potência elétrica considerável, como é o caso de centros de maquinação CNC, impressoras 3D de metal, fresadoras e tornos convencionais, engenhos de furar, máquinas de soldar e ainda outros como é o caso de compressores para produzir ar comprimido para as máquinas e sistemas de extração de gases.

5.1.1. Origem e Formato dos Dados

Os dados aqui trabalhados possuem três origens distintas. Os dados relacionados com os consumos energéticos foram retirados do site da E-REDES, sendo estes os comunicados para a faturação da empresa comercializadora de energia. Estes dados podem ser retirados na área do cliente do site da E-REDES em formato CSV, sendo apenas possível retirar os dados referentes a um mês por ficheiro. Desta forma, existem oitenta e quatro ficheiros, um para cada mês, começando no mês de outubro de 2015 até setembro de 2022. Estes ficheiros são

as contagens de potência elétrica do edifício, tendo uma resolução quarto horária para os valores de energia ativa, energia reativa indutiva e energia reativa capacitiva. A energia ativa é a energia que realiza trabalho num sistema/equipamento, isto é, converte a energia elétrica em outras formas de energia como luminosa, mecânica ou calorífica. Já a energia reativa pode ser indutiva ou capacitiva, sendo que é a parte de energia elétrica que não realiza trabalho. No entanto, a parcela de reativa indutiva existe por consequência do fluxo magnético necessário ao funcionamento de motores, transformadores ou outros equipamentos que utilizem bobinas elétricas e, a parcela de reativa capacitiva existe por consequência do funcionamento de condensadores utilizados em equipamentos elétricos.

Os dados relacionados com a ocupação do edifício foram pedidos ao Departamento de Sistemas de Informação do CENFIM e referem-se apenas ao edifício da Marinha Grande. Embora estes dados estejam alojados em bases de dados, foi possível extrair a informação através de uma consulta, tendo sido posteriormente exportados para um ficheiro Excel. Estes dados estão tabelados com várias colunas, sendo que as mais importantes são referentes à data, hora de início, hora de fim, sala e número de ocupantes.

Quanto aos dados meteorológicos, estes foram solicitados ao IPMA – Instituto Português do Mar e da Atmosfera para o período correspondente aos dados energéticos. Os dados enviados estão no formato Excel e são da temperatura, possuem uma resolução diária e são referentes ao concelho da Marinha Grande.

5.2. Tratamento dos Dados

Os dados meteorológicos e os dados de ocupação do edifício não possuem valores omissos, pelo que o seu tratamento foi relativamente simples. Os dados meteorológicos com a resolução diária mantiveram-se. Quanto aos dados referentes à ocupação, foi desenvolvida uma fórmula no ficheiro Excel para somar a ocupação por dia, tendo em conta os horários de início e de fim e as salas ocupadas. As salas ocupadas foram tidas em conta uma vez que, atendendo à denominação da sala, é possível separar estas por tipologia. Desta forma, para além de obter a ocupação total do edifício, desagregou-se esta ocupação em quatro tipologias de espaço, nomeadamente salas de formação teóricas, salas de formação teórico-práticas/informática, laboratórios e oficinas. Por fim, obteve-se um ficheiro com a ocupação por dia, o qual se juntou com os dados meteorológicos.

Quanto aos dados energéticos do edifício, começou-se por juntar todos os ficheiros CSV num só através de uma macro em VBA criada para esse efeito. Nesta fase, procedeu-se à análise de valores omissos ou valores de zero e valores duplicados, tendo-se concluído que, embora o *dataset* não possuísse valores omissos, existiam alguns períodos onde a potência elétrica instantânea era zero e tal não se deveria verificar mesmo em períodos sem ocupação, uma vez que existem sistemas que estão ligados de forma permanente e, embora tenham um consumo residual, este nunca pode ser zero. Para além dos valores zero, também se identificaram valores duplicados, tendo sido observados no total 103 valores zero e 32 valores duplicados, sendo parte destes correspondentes à mudança da hora. Para o tratamento destes dados foram utilizadas técnicas presentes na literatura, sendo que, quanto aos valores duplicados, foram realizadas as médias entre observações. Quanto aos valores zero referentes à mudança da hora, sendo estes também correspondentes a 32 observações, foram utilizados os valores médios para a mesma hora dos 3 dias anteriores e 3 dias posteriores. Para as restantes observações com valores zero, atendendo que estas se verificaram em períodos do dia onde deveria existir consumo e, tendo-se concluído que tal não se verificou devido a uma falha de comunicação do distribuidor de energia elétrica, utilizou-se o valor médio do dia e hora correspondente das 3 semanas anteriores e das 3 semanas posteriores. Assim, por exemplo, a observação do dia 21/05/2022 pelas 14:00 que possuía um valor 0 e correspondia a um sábado, foi substituída pela média do consumo dos 3 sábados anteriores e dos 3 sábados posteriores.

Após a resolução dos valores duplicados e omissos e, tendo em conta que o ficheiro obtido possuía uma resolução quarto horária para as potências elétricas instantâneas (kW), foi necessário transformar as potências elétricas em energia elétrica (kWh) e agregar a energia para obter uma frequência diária. Para isso, utilizaram-se fórmulas no ficheiro convertido anteriormente para Excel, para obter a energia elétrica por dia, onde se multiplicou a potência elétrica pela resolução horária correspondente e, posteriormente, somaram-se os consumos correspondentes a um só dia. Desta forma, os dados relacionados com a energia elétrica encontram-se na mesma frequência que os restantes, pelo que se pode proceder à sua junção num único *dataset* para análise.

Por fim, obteve-se um *dataset* com 2557 observações e cujas variáveis se apresentam de seguida na Tabela 2:

Tabela 2 - Variáveis do *Dataset*

Variável	Descrição	Tipo	Unidade
Data	Data em formato dd/mm/aaaa	Data	Data
EnergiaA	Energia ativa	Decimal	kWh
EnergiaR_I	Energia reativa indutiva	Decimal	kVArh
EnergiaR_C	Energia reativa capacitiva	Decimal	kVArh
OcupT	Ocupação total	Inteiro	unidade
Ocup_Teorica	Ocupação das salas teóricas	Inteiro	unidade
OcupComput	Ocupação das salas de computadores	Inteiro	unidade
OcupLab	Ocupação dos laboratórios	Inteiro	unidade
OcupOfi	Ocupação das oficinas	Inteiro	unidade
Tmin	Temperatura mínima	Decimal	°C
Tmax	Temperatura máxima	Decimal	°C

5.2.1. Dados de Energia

Os dados de energia ativa alvo de análise, como já foi referido anteriormente, possuem uma frequência diária e são relativos ao período de 01/10/2015 a 30/09/2022, apresentando-se na Figura 17 a série temporal resultante. Pela análise da série temporal, é possível verificar que os valores mínimos se encontram estáveis ao longo do período temporal, com a exceção do período de início de março de 2020 até meados de maio do mesmo ano. Este período correspondeu ao primeiro isolamento geral aquando da pandemia COVID-19, o que levou a um fecho total das instalações e, por consequência, uma redução significativa no consumo energético. Para além disso, na série temporal são visíveis vários picos e vales. Tendo em conta que os dados reportam ao consumo de energia ativa diária, os picos correspondem a dias da semana e os vales a fins-de-semana. Ainda é possível identificar o período de férias letivas, o qual corresponde ao mês de agosto com uma duração usual de três semanas. Contudo, esta pausa foi menor em 2020 e 2021, devido aos efeitos da pandemia.

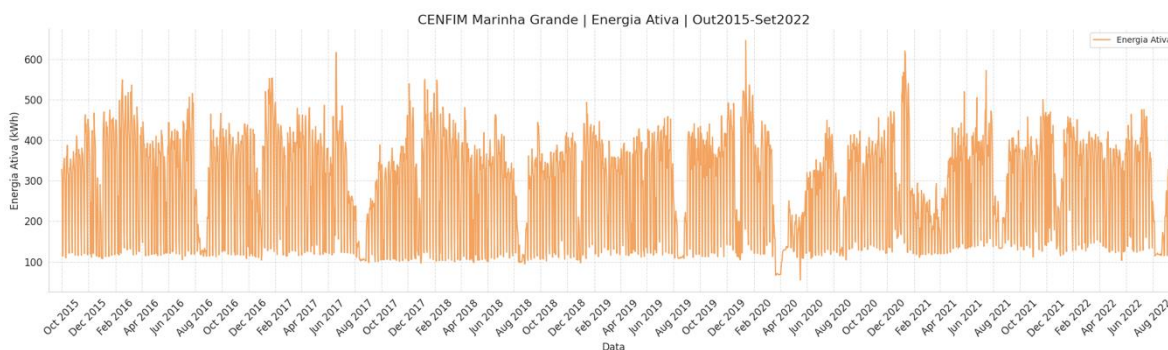


Figura 17 – Consumo de energia ativa durante o período de 01/10/2015 a 30/09/2022

Na Figura 18 é visível o consumo energético diário original, assim como o consumo energético agregado por semana e por mês.

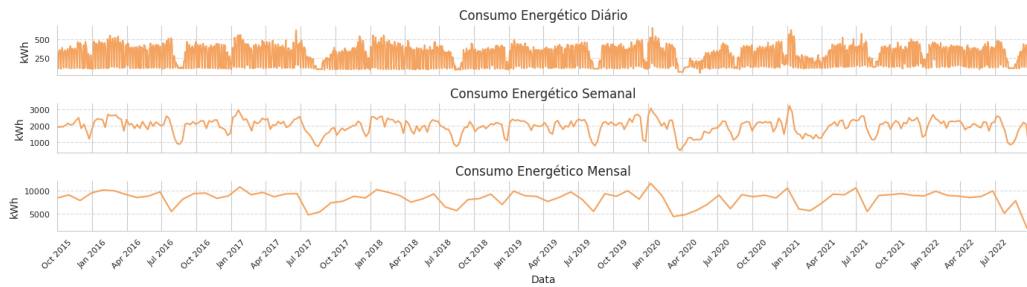


Figura 18 – Consumo de energia ativa diário e agregado por semana e por mês

Tendo em conta que a série temporal se refere a um período temporal relativamente grande, na Figura 19 é apresentado o consumo de energia elétrica por ano letivo para uma visão mais pormenorizada do comportamento do consumo energético em cada ano letivo.

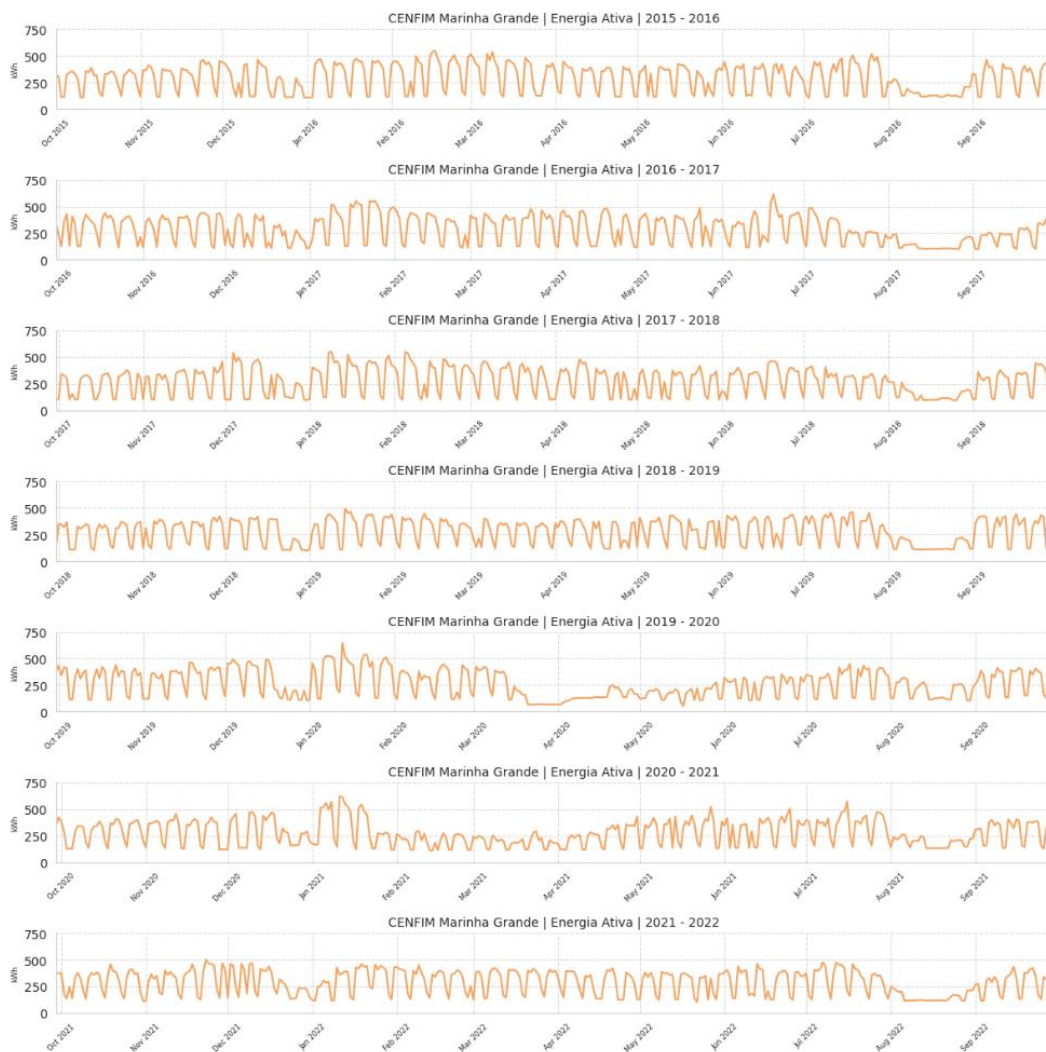


Figura 19 - Consumo de energia ativa por ano letivo

Da análise da Figura 19, é possível verificar que o consumo energético até ao ano de 2019 teve um comportamento semelhante ano após ano. Apenas após o ano de 2020, é que se torna visível o efeito da pandemia COVID-19 no consumo energético do edifício em análise. Mais especificamente, entre meados de março e junho de 2020 e entre meados de janeiro e maio de 2021 o consumo de energia elétrica sofreu alterações significativas relativamente aos anos transatos, sendo que no primeiro período a ocupação do edifício era completamente nula e no segundo período o edifício era ocupado apenas por colaboradores, o que levou a um consumo energético relativamente superior. Para visualizar o consumo energético por mês e por ano deverá ser consultado o Anexo I.

De forma a obter uma visão mais concreta dos consumos energéticos do edifício do CENFIM da Marinha Grande, foi criado o diagrama de extremos e quartis da energia ativa por ano e mês do período em análise, o qual se apresenta na Figura 20.

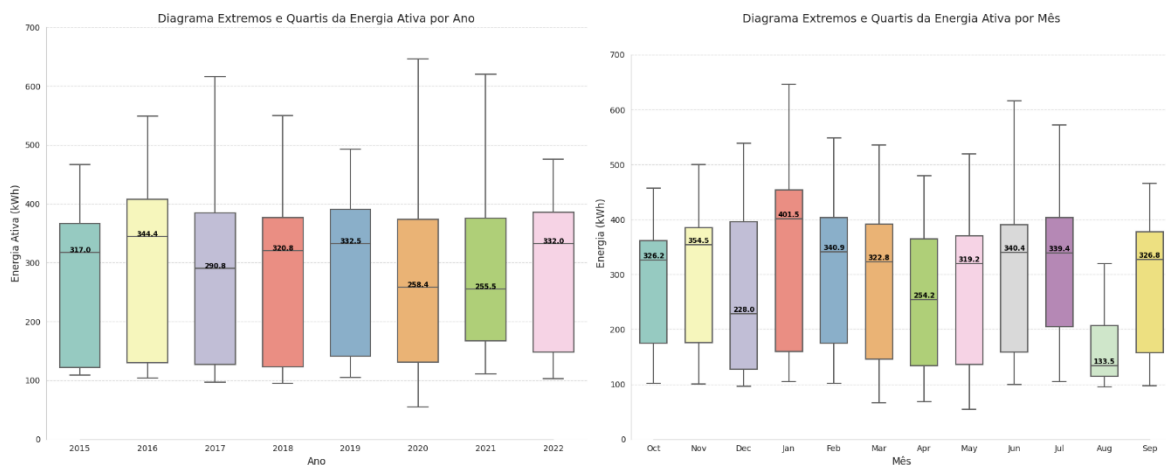


Figura 20 – Diagramas de extremos e quartis da energia ativa por ano e por mês

Através da análise da figura é possível verificar que, a mediana do consumo energético foi inferior nos dois anos onde existiu isolamento devido à pandemia. Para além disso, a pausa que resultou deste confinamento exigiu que nos restantes períodos de ocupação do edifício o consumo energético fosse mais intenso, o que originou valores máximos em 2020 e 2021 superiores aos restantes anos. No que toca ao consumo por mês, é visível de forma notória que o mês de agosto é o que apresenta o consumo mais baixo devido à pausa do período de férias letivas. Além disso, os meses de dezembro e abril também apresentam um consumo energético mediano inferior aos restantes meses, o que é explicado pelos períodos de pausa letiva relacionado com a época festiva do Natal e da Páscoa. Já janeiro é o mês que apresenta o consumo mediano mais alto, o que poderá ser explicado pela retoma da atividade de

formação e um alto nível de ocupação do edifício neste mês, onde a maior parte das turmas dos cursos de formação não se encontram em período de estágio.

5.2.2. Dados da Ocupação

A ocupação do edifício em estudo é apresentada de seguida nas Figuras 21 e 22. Indo ao encontro da análise efetuada anteriormente para o consumo de energia ativa, a Figura 21 apresenta a ocupação total ao longo da totalidade do período em estudo, enquanto que a Figura 22 representa a ocupação total por ano letivo.

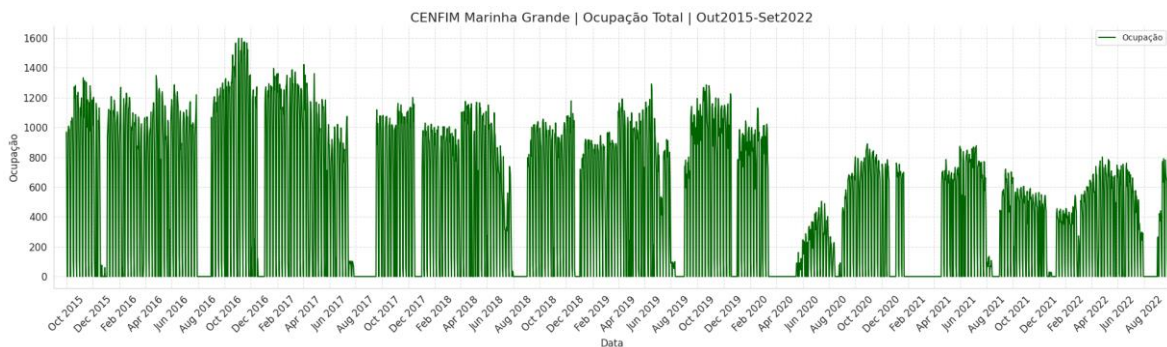


Figura 21 – Ocupação total do edifício durante o período de 01/10/2015 a 30/09/2022

Através da análise da Figura 21, podemos verificar que ao longo do tempo a ocupação tem vindo a diminuir de forma geral. Esta diminuição poderá estar relacionada com um número inferior de formandos que têm procurado a formação profissional. Nesta figura são ainda visíveis vários períodos de pausa que coincidem com períodos onde o consumo energético também diminui. Estes períodos são correspondentes às pausas letivas dos meses de agosto e dezembro. Embora a ocupação seja 0 durante esses períodos, é importante referir que a ocupação do edifício poderá não o ser. Isto é, os dados aqui apresentados referentes à ocupação são indicativos apenas do número de formandos que estiveram presentes no edifício, tendo em conta que a origem destes dados foi dos sistemas de gestão da formação do CENFIM. Assim, a ocupação dos colaboradores não se encontra aqui retratada, uma vez que esse registo é feito através das picagens de ponto, as quais não foram analisadas. É importante referir que esta parcela de ocupação inclui todos os colaboradores do CENFIM da Marinha Grande, onde se incluem todos os postos de trabalho, nomeadamente a receção, os serviços administrativos, os formadores, os serviços de direção e outros serviços de especialidade. A sua ocupação, embora residual no que toca ao consumo energético do edifício, impacta o total do consumo de energia elétrica.

Ainda analisando as referidas figuras, em ambas é visível o impacto do confinamento devido à pandemia. Mais especificamente, na Figura 22, podemos ver períodos em 2020 e 2021 onde a ocupação é 0, o que significa que durante esse período não existiu formação presencial nas instalações. Estes períodos coincidem com os já analisados da energia ativa, os quais produziram efeitos no consumo deste recurso.

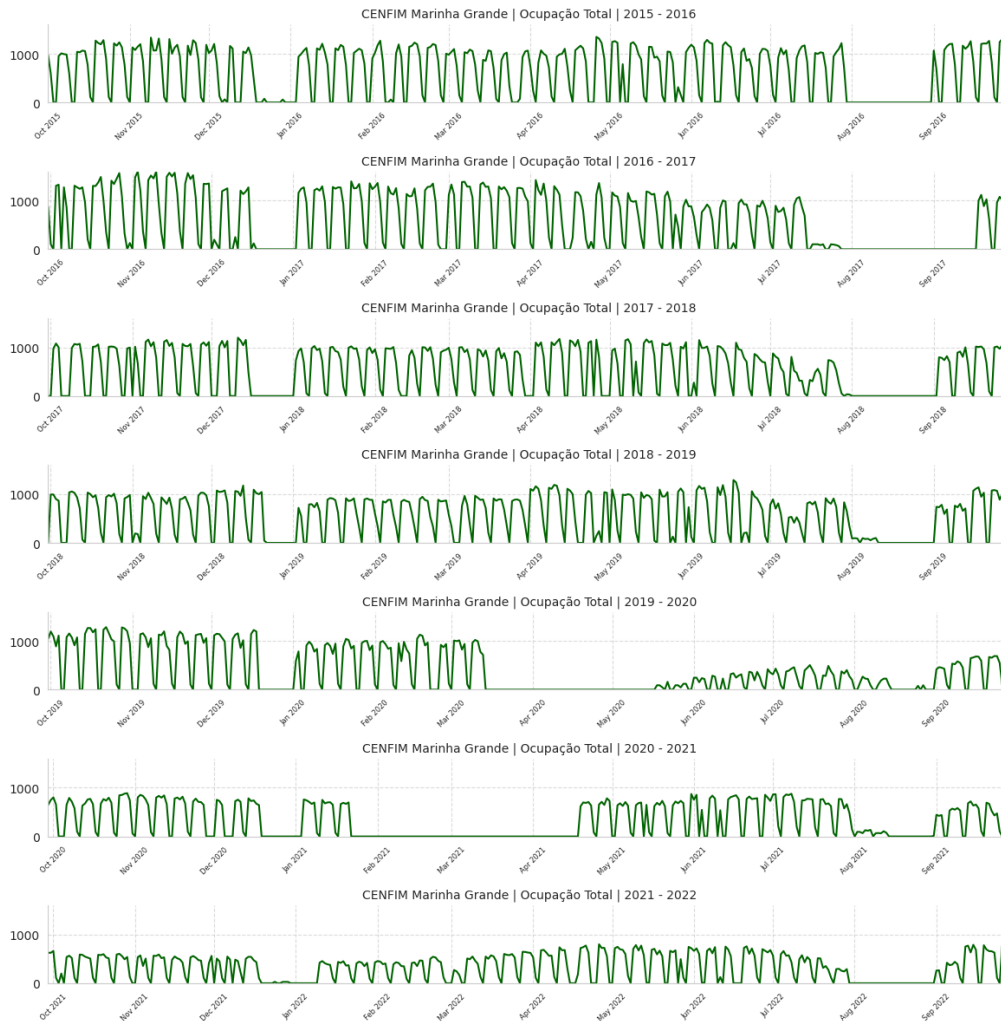


Figura 22 - Ocupação total do edifício por ano letivo

5.2.3. Dados da Temperatura

Relativamente aos dados de temperatura, estes foram cedidos pelo IPMA e correspondem aos valores da temperatura máxima e mínima diária para o mesmo período temporal das variáveis apresentadas anteriormente. Os dados são apresentados na Figura 23 e, como seria expectável, os valores de temperatura seguem o mesmo comportamento ao longo do tempo com um aumento da temperatura durante os meses de Verão e uma diminuição da temperatura nos meses de Inverno. Os dados de temperatura, em termos conceptuais,

configuram-se relevantes na análise do consumo energético de qualquer edifício, uma vez que, e tendo em conta o clima temperado de Portugal, as estações de Verão e Inverno exigem consumos energéticos superiores às estações da Primavera e Outono, sendo que tal dever-se-ia verificar nos consumos energéticos. Contudo, através da inspeção visual da série temporal da energia ativa, tal facto não se verificou, o que leva a crer que os consumos do edifício relacionados com aspetos de climatização não são relevantes. Este facto já era esperado, uma vez que a quantidade e potência elétrica dos equipamentos utilizados para lecionar cursos nas áreas da metalurgia e metalomecânica ultrapassam consideravelmente a potência elétrica instalada dos sistemas de climatização. Contudo, para averiguar esse facto foram calculados os índices de correlação entre as variáveis, os quais se apresentam de seguida na Figura 24.

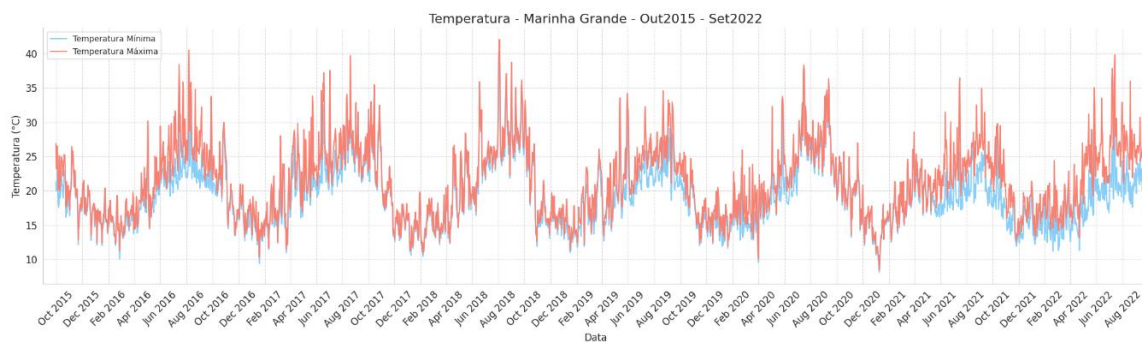


Figura 23 – Valores de temperatura máxima e mínima total durante o período de 01/10/2015 a 30/09/2022

Da análise da Figura 24, podemos verificar que a temperatura máxima e a temperatura mínima são as variáveis com o maior índice de correlação positivo, como tal seria de esperar. Contudo, nenhuma destas variáveis se encontra correlacionada com a energia ativa consumida. Quanto às restantes variáveis podemos observar que a ocupação total e a energia ativa possuem um índice de correlação positivo significativo.

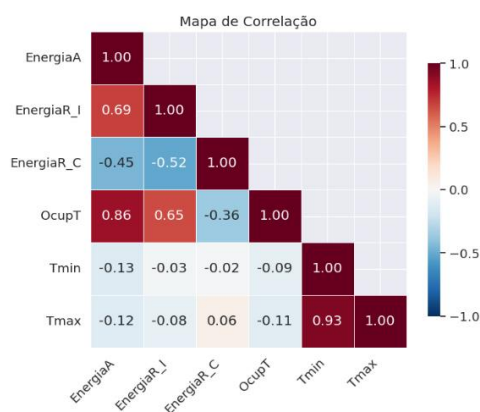


Figura 24 -Mapa de correlação das variáveis

Esta página foi propositadamente deixada em branco

6. Criação dos Modelos e Análise dos Resultados

Neste capítulo são apresentados os vários modelos utilizados, assim como os parâmetros para cada um deles. Os modelos utilizados são divididos em dois grupos, nomeadamente os modelos de base estatística e os modelos de aprendizagem computacional, os quais já foram explicados anteriormente. Deve referir-se também que todos os modelos foram aplicados com o objetivo de tirar o maior partido de cada um, tendo em conta as suas especificações e limitações.

6.1. Modelos de Base Estatística

Os primeiros modelos a serem testados foram os de base estatística, nomeadamente o SARIMA e as suas variações SARIMAX. Optou-se por escolher apenas o modelo SARIMA tendo em conta que a série temporal apresenta nitidamente uma componente sazonal. Elaborou-se a Figura 25 que representa a série temporal original do consumo de energia ativa e as médias móveis centradas com períodos de 7 e 30 dias.

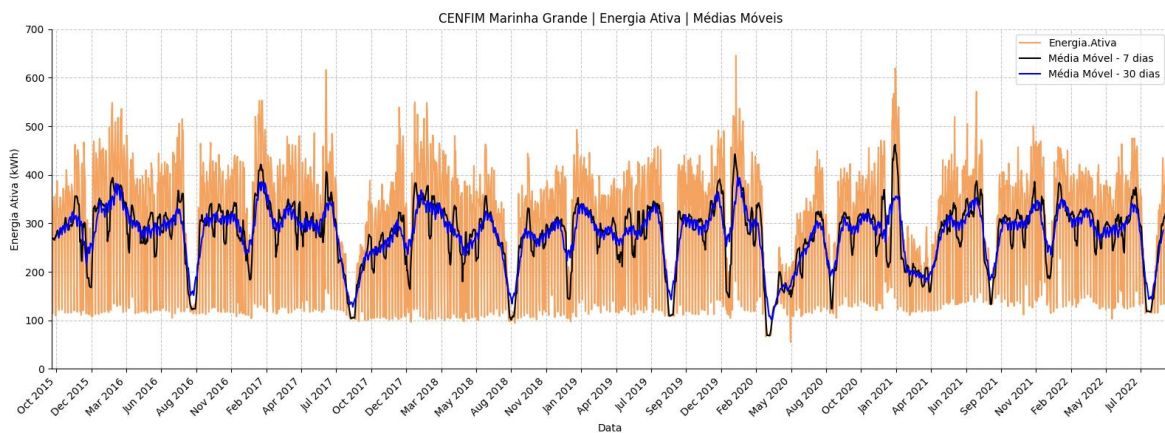


Figura 25 - Série temporal e médias móveis

Observando as médias móveis, a que aparenta representar a série temporal com mais detalhe é a de média móvel de 7 dias, tendo-se procedido à decomposição da série temporal com o modelo aditivo. A decomposição da série temporal é visível na Figura 26.

Através da análise da Figura 26, podemos observar que a tendência coincide com a média móvel de 7 dias e os períodos de tendência negativa mais acentuados coincidem com a pausa letiva do mês de agosto, excetuando-se o período da pandemia, onde também é visível uma diminuição. Através da Figura 26 não conseguimos observar com nitidez que esta série

temporal possui sazonalidade. Contudo, é possível afirmar que esta série possui uma componente sazonal semanal, sabendo que durante a semana existe um maior consumo energético comparativamente aos fins de semana e a qual é visível na análise exploratória efetuada no capítulo anterior.

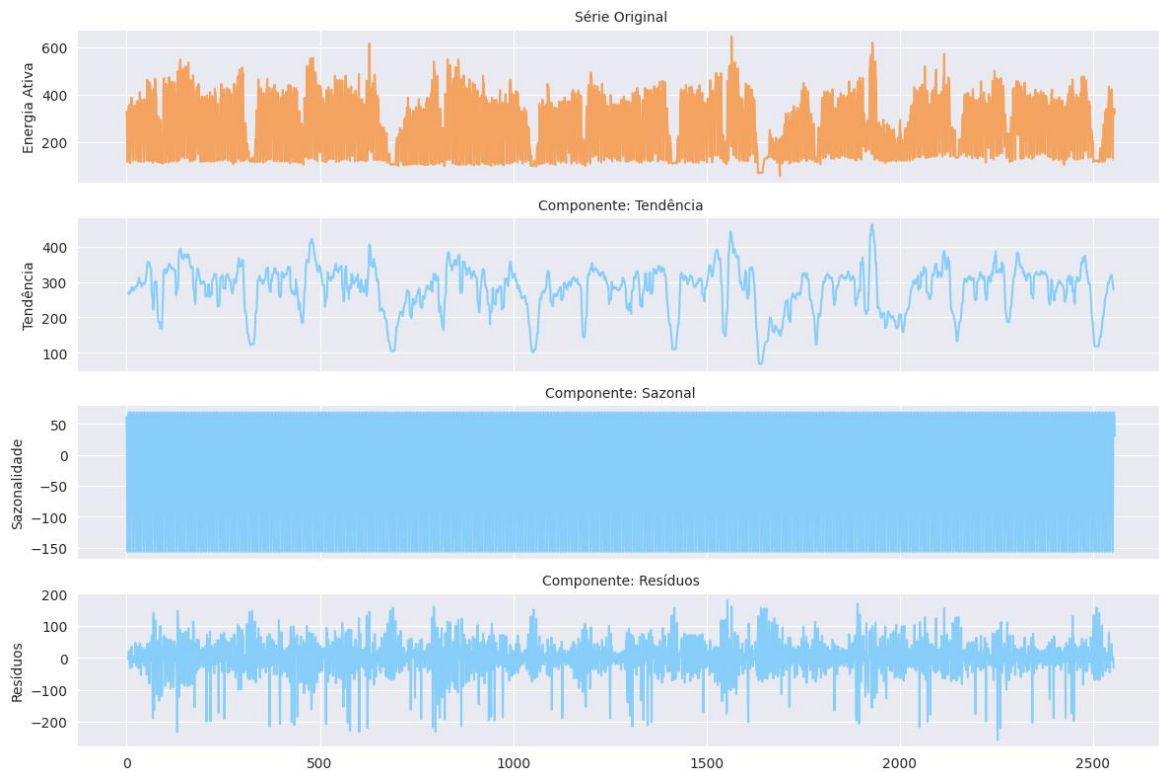


Figura 26 - Decomposição de série temporal através do modelo aditivo e com um período de 7

De seguida, procedeu-se à análise de estacionariedade da série temporal, tendo-se realizado um teste de hipóteses, especificamente o teste de *Dickey-Fuller* aumentado cujos resultados são visíveis na Tabela 3. Dos resultados obtidos podemos concluir que a série temporal é estacionária, pelo que não é necessário proceder a nenhum método para a tornar estacionária antes de estimar os modelos.

Tabela 3 - Teste de hipóteses para averiguar a estacionariedade

Teste de Hipóteses – Teste <i>Dickey-Fuller</i> aumentado (ADF)	
H₀ (Hipótese nula)	Os dados da série temporal são não estacionários
H₁ (Hipótese alternativa)	Os dados da série temporal são estacionários
α (alfa) – Nível de significância	0,05
<i>p-value</i>	9,373e-10
Decisão	$p\text{-value} \ll \alpha$ (alfa) \rightarrow Rejeita-se a hipótese nula, pelo que assumimos que a série temporal é estacionária

Posteriormente, procedeu-se à análise dos gráficos de autocorrelação e de autocorrelação parcial, os quais são visíveis na Figura 27. Como se pode observar, existem vários *lags* significativos em ambos os gráficos. Contudo, para obter o melhor modelo é utilizada a função *auto_arima* da biblioteca *pmdarima* com o *Python*. Esta função permite estimar inúmeros modelos utilizando várias combinações de ordens das componentes dos próprios modelos. Os parâmetros da ordem de diferenciação foram mantidos a 0, uma vez que a série já é estacionária. A sazonalidade foi definida como semanal e os parâmetros da parte autorregressiva e de médias móveis foram iniciados a 0 e testados até ao valor de 15, tendo em conta os *lags* significativos visíveis na Figura 27.

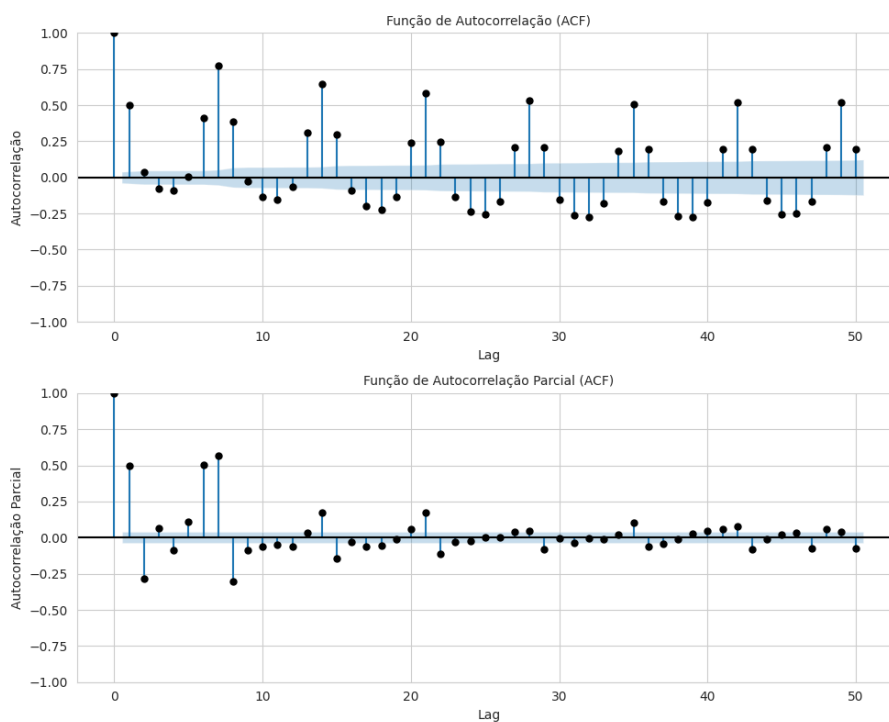


Figura 27 - Gráfico de autocorrelação e de autocorrelação parcial

6.1.1. Modelo SARIMA

O modelo SARIMA foi estimado com a utilização da variável da energia ativa por dia, sendo que o modelo que apresentou melhor desempenho, observado através da minimização do valor das métricas de AIC (*Akaike Information Criterion*) e BIC (*Bayesian Information Criterion*) é o que se apresenta na Tabela 4.

Tabela 4 - Critérios de Informação para o modelo SARIMA

Critérios de Informação		
Modelo	AIC	BIC
SARIMA (1,0,0) (1,0,2) [7]	28796,06	28831,14

Para estimar o modelo SARIMA e obter as ordens dos parâmetros, inicialmente foram utilizadas todas as observações disponíveis, estando disponível no Anexo II o resumo dos resultados dos resíduos do melhor modelo. Após a identificação dos parâmetros que configuram o modelo SARIMA, nomeadamente as ordens dos parâmetros autorregressivos e de médias móveis para ambas as componentes, o modelo foi aplicado para estimar o consumo do próximo dia.

De forma a conseguir comparar os resultados obtidos dos modelos de base estatística com os modelos de aprendizagem computacional, o modelo foi estimado 365 vezes tendo sido posteriormente feita uma previsão de um dia. Isto é, tendo em conta que os modelos de aprendizagem computacional requerem que os dados em análise sejam divididos em conjuntos de treino, validação e teste, e sendo normal o conjunto de treino ser o maior e os outros dois terem o mesmo comprimento e serem representativos de um ciclo completo do que se pretende analisar, para estes modelos, as 2557 observações, que totalizam 7 anos, foram divididas em cinco anos para o treino, um ano para a validação e um ano para o teste. Assim, tanto o conjunto de validação como o conjunto de teste têm 365 observações. Contudo, nesta fase onde se estimam os modelos de base estatística, por estes serem modelos autorregressivos e de natureza diferente das redes neuronais, não se implementa uma fase de treino onde os parâmetros são iterativamente atualizados, mas sim uma estimação do modelo consoante os parâmetros previamente selecionados e uma previsão onde se definiu um horizonte temporal.

Desta maneira, para uma comparação justa dos modelos e tirando o máximo proveito de cada um deles, decidiu-se criar um ciclo iterativo, onde o modelo é estimado sempre com os parâmetros obtidos previamente e com a janela de dados a aumentar em 1 dia a cada iteração, sendo que a previsão obtida é sempre para o dia seguinte. Com a criação deste ciclo é possível registar os valores das previsões num vetor, o qual é de seguida comparado com o vetor dos valores reais para proceder ao cálculo das métricas de erro. Para além disso, as previsões obtidas correspondem aos últimos 365 dias do conjunto de dados e ao conjunto de teste que será utilizado nos modelos de aprendizagem computacional.

Após a aplicação do modelo descrito anteriormente, foram calculadas as métricas de erro, as quais se apresentam na Tabela 5. Com a aplicação desta técnica, foi possível obter um erro percentual médio absoluto de 18,27% para as 365 previsões estimadas com o modelo SARIMA.

Tabela 5 – Métricas de Erro para o modelo SARIMA

Métricas de Erro					
Modelo	MSE	RMSE	MAE	MAPE	R ²
SARIMA (1,0,0) (1,0,2) [7]	3408,43	58,38	40,27	0,1827	0,7482

Para uma melhor visão do comportamento das previsões face aos valores reais são apresentados na figura abaixo ambos os valores em ambiente gráfico. Da análise da figura é possível verificar que as previsões correspondentes ao mês de agosto sofreram um grande erro, sendo que o modelo não conseguiu identificar uma pausa letiva apenas com a utilização da variável da energia ativa e, por isso, decidiu-se estudar, de seguida, um modelo SARIMAX com a introdução de outras variáveis exógenas no modelo.

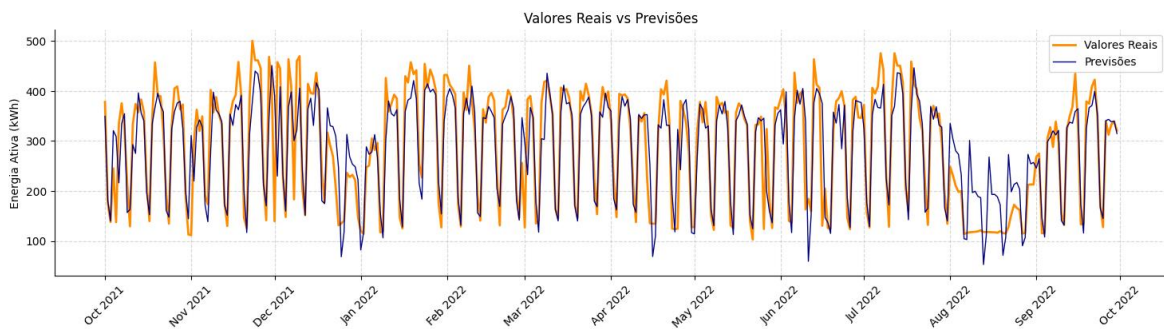


Figura 28 - Valores reais vs previsões para o modelo SARIMA

6.1.2. Modelo SARIMAX – com uma variável exógena

O modelo SARIMAX foi estimado com recurso a duas variáveis, nomeadamente a energia ativa como variável endógena e a ocupação total como variável exógena, com o objetivo de obter um modelo mais adequado aos dados em estudo e que realize previsões mais precisas. A variável exógena selecionada, como se previa e tendo em conta que se encontra fortemente correlacionada com a energia ativa, impacta significativamente a energia utilizada no edifício, assim, e como se pode comprovar na Tabela 6, o melhor modelo obtido apresenta tanto um valor de AIC como um valor de BIC inferiores ao modelo anterior, o que significa que este modelo representa melhor a série temporal em estudo.

Tabela 6 - Critérios de Informação para o modelo SARIMAX – com uma variável exógena

Critérios de Informação		
Modelo	AIC	BIC
SARIMAX (2,0,0) (7,0,0) [7]	26083,72	26153,88

Este modelo foi obtido da mesma forma que o anterior, isto é, com recurso à função *auto_arima* da biblioteca *pmdarima* e com a iniciação de todos os parâmetros a 0 até a um máximo de 15, excetuando-se o parâmetro de diferenciação, o qual foi mantido a 0, tendo em conta que a série temporal é estacionária. O Anexo III contém o resumo dos resultados dos resíduos do modelo obtido.

As previsões deste modelo foram obtidas da mesma forma que o modelo anterior, tendo-se criado um ciclo onde a previsão efetuada foi para o dia seguinte. Contudo, para o modelo estimar o consumo de energia ativa do dia seguinte com este modelo, é necessário fornecer ao modelo a série temporal da variável exógena com o dado correspondente da previsão, isto é, com informação da ocupação total do edifício do dia seguinte. Tanto as métricas de erro para as previsões como a comparação destas com os valores reais encontram-se apresentados de seguida e respetivamente na Tabela 7 e Figura 29.

Tabela 7 - Métricas de Erro para o modelo SARIMAX – com uma variável exógena

Métricas de Erro					
Modelo	MSE	RMSE	MAE	MAPE	R ²
SARIMAX (2,0,0) (7,0,0) [7]	1692,69	41,14	30,74	0,1281	0,8749

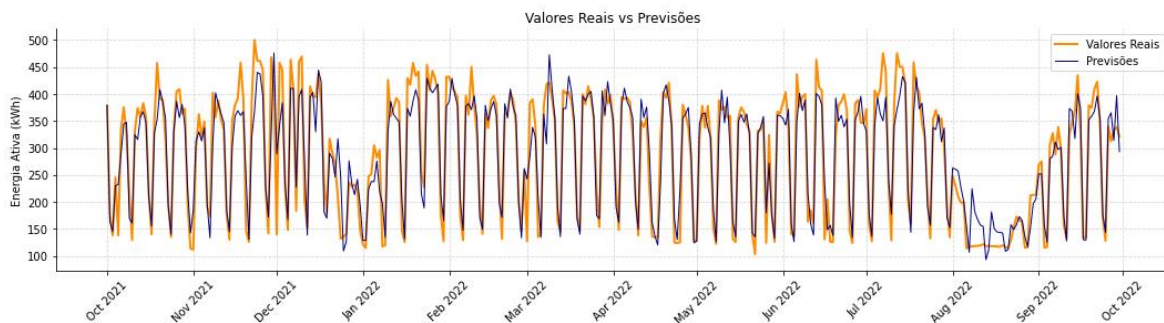


Figura 29 - Valores reais vs previsões para o modelo SARIMAX – com uma variável exógena

Através da análise de ambos os elementos, é possível verificar que os resultados obtidos melhoraram em comparação com o modelo anterior, tendo-se obtido um erro percentual médio absoluto de 12,81%. A previsão dos consumos energéticos ao longo do horizonte temporal captou de forma correta a evolução do consumo energético, sendo visíveis melhorias face aos resultados anteriores, especificamente as previsões dos pontos mínimos dos consumos energéticos. Quanto aos pontos máximos, o modelo continua com alguma dificuldade a estimá-los, atendendo que estes são os mais difíceis de estimar por vários

fatores, dentro dos quais alguma eventualidade da utilização fora do usual de máquinas e equipamentos durante os períodos de formação e mesmo até a versatilidade de utilização. Este modelo apresenta melhorias em relação ao anterior para o período de agosto, contudo, embora estime um consumo bastante inferior, este ainda não é representativo da pausa total da utilização do edifício.

6.1.3. Modelo SARIMAX – com quatro variáveis exógenas

Tendo em conta que o modelo SARIMAX com a adição de uma variável exógena apresentou melhores resultados que o modelo SARIMA, testou-se outro modelo SARIMAX agora com a adição de quatro variáveis exógenas, nomeadamente a ocupação total, as temperaturas mínimas e máximas e o valor de energia reativa indutiva. Apresenta-se novamente na Tabela 8 os critérios de informação para o melhor modelo obtido através da função *auto_arima*, onde é possível comprovar, tendo em conta os valores de AIC e BIC, que o modelo obtido é melhor representativo da série temporal de energia ativa. O Anexo IV contém o resumo dos resultados dos resíduos do modelo obtido.

Tabela 8 - Critérios de Informação para o modelo SARIMAX – com quatro variáveis exógenas

Critérios de Informação		
Modelo	AIC	BIC
SARIMAX (4,0,1) (1,0,7) [7]	24904,84	25015,92

As previsões geradas para este modelo foram obtidas da mesma forma que os dois modelos anteriores. No entanto, para o modelo SARIMAX com quatro variáveis estimar o consumo de energia ativa do dia seguinte foi necessário fornecer os valores das séries temporais das variáveis exógenas correspondentes à previsão, ou seja, ocupação total, as temperaturas máximas e mínimas e valor de energia reativa indutiva do dia seguinte. Os resultados obtidos são apresentados de seguida na Tabela 9 e na Figura 30.

Tabela 9 - Métricas de Erro para o modelo SARIMAX – com quatro variáveis exógenas

Métricas de Erro					
Modelo	MSE	RMSE	MAE	MAPE	R ²
SARIMAX (4,0,1) (1,0,7) [7]	968,28	31,12	22,89	0,0974	0,9284

Os resultados obtidos para o modelo SARIMAX com quatro variáveis exógenas apresentou os melhores resultados de todos os modelos de séries temporais testados com um erro percentual médio absoluto de 9,74% e um R² de 0,9284, o que significa que as variáveis

aplicadas para a criação deste modelo explicam 92,84% da variância da energia ativa para o horizonte temporal correspondente à previsão. Também através da inspeção visual dos valores previstos *versus* os valores reais, é possível verificar que o modelo conseguiu captar adequadamente a evolução do consumo energético ao longo do horizonte temporal de previsão. Apesar de o modelo aparentar previsões de valores máximos mais próximos dos reais, este ainda falha em vários pontos, os quais têm de ser considerados no caso de aplicação deste modelo para previsão do consumo. Também quanto ao mês de agosto, é possível verificar que nenhum modelo foi capaz de prever um consumo estável e baixo durante o período de férias onde a ocupação total do edifício é igual a 0, tendo as previsões mimicado o padrão normal do consumo durante as semanas do ano, todavia muito inferior ao normal durante o período letivo de formação.

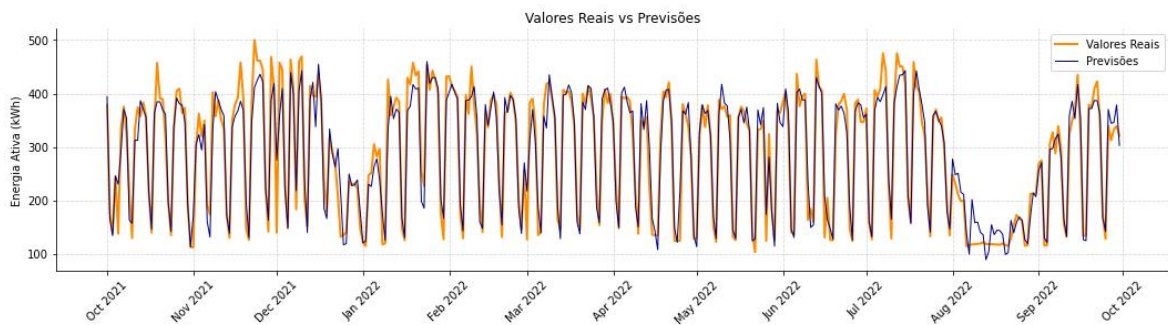


Figura 30 - Valores reais vs previsões para o modelo SARIMAX – com quatro variáveis exógenas

Embora estes resultados sejam consideravelmente melhores que os obtidos com os dois modelos anteriores, deverá evidenciar-se o facto de que foram obtidos com valores reais de variáveis exógenas que para aplicação do modelo poderão ser difíceis de estimar. Isto é, quanto à ocupação total, este valor é fácil de estimar com um erro relativamente baixo e poderá ser aplicado para previsão da energia ativa, assim como os valores de temperaturas. Contudo, a energia reativa indutiva varia consoante a quantidade de equipamentos específicos que se utiliza e poderá ser mais difícil de obter com precisão, o que poderá comprometer a qualidade da previsão obtida.

Outro aspeto importante que deve ser referido foi o tempo que demorou cada modelo a prever o consumo para o horizonte temporal de previsão. Tendo em conta que o modelo foi criado para prever a energia consumida do dia seguinte por forma a comparar os resultados obtidos com os modelos de aprendizagem computacional que se antevêm, a previsão para o horizonte temporal de um ano demorou significativamente algum tempo em termos de esforço computacional, sendo que este tempo aumentou sempre com a complexidade dos

modelos. Contudo, após a realização dos testes para averiguar o melhor modelo e uma vez utilizado numa aplicação, o modelo SARIMAX apenas terá que realizar a previsão para o dia seguinte, o que por sua vez se traduzirá num tempo reduzido de treino e de obtenção da previsão.

6.2. Modelos de Aprendizagem Computacional

Os modelos de aprendizagem computacional testados formam redes neuronais, as quais já foram apresentadas anteriormente neste documento. Dentro destas, e atendendo que o problema a resolver é um problema de regressão, foram aplicadas redes MLP – *Multylayer Perceptron*, RNN – *Recurrent Neural Network*, GRU – *Gated Recurrent Units* e LSTM – *Long-Short Term Memory*, sendo os resultados apresentados nas secções seguintes.

Inicialmente foram abordados quatro cenários distintos, todos para previsão do consumo de energia elétrica do dia seguinte com base no consumo do passado, isto é, em cada um destes cenários foi aplicada apenas uma variável, especificamente a variável que se pretende prever, sendo esta a energia ativa. Os cenários abordados foram os seguintes:

- **C1:** Utilização do consumo energético dos 7 dias anteriores para prever o consumo energético do 8.º dia;
- **C2:** Utilização do consumo energético dos 14 dias anteriores para prever o consumo energético do 15.º dia;
- **C3:** Utilização do consumo energético dos dias das últimas 4 semanas para prever o consumo energético do dia correspondente da 5.ª semana, por exemplo, prever o consumo da próxima terça-feira com base nos valores das quatro terças feiras anteriores;
- **C4:** Utilização do consumo energético dos dias das últimas 8 semanas para prever o consumo energético do dia correspondente da 9.ª semana;

O treino das redes neuronais envolveu a divisão do conjunto de dados em conjunto de treino, conjunto de validação e conjunto de teste, onde se teve como objetivo o dimensionamento para 365 observações dos conjuntos de validação e teste, como observado na Figura 31.

Assim, para estes conjuntos, os dados obtidos são representativos de um horizonte temporal longo o suficiente para possuir informação relativa aos eventos que ocorrem durante um ano letivo, dando ao mesmo tempo aos conjuntos de treino observações suficientes para os

parâmetros das redes neuronais se ajustarem. Os conjuntos de treino possuem sensivelmente cinco anos de informação, estando esta no formato de energia ativa consumida por dia, a qual possui pequenas oscilações dependendo do cenário em estudo.

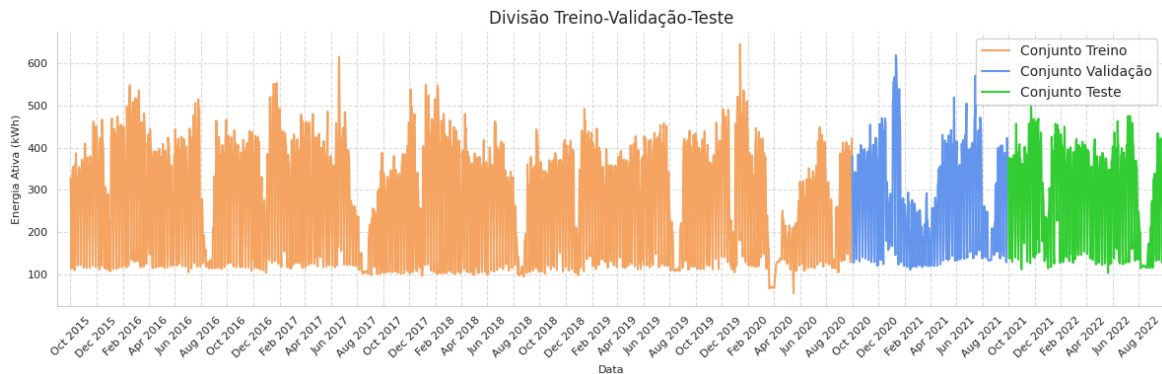


Figura 31 – Divisão conjunto treino, validação e teste

Outro aspeto importante que deve ser referido nesta fase é a normalização dos dados. No treino de redes neuronais é recomendado que os dados sejam normalizados à entrada, contudo, é comum na literatura para previsão de consumos energéticos com redes neuronais que os dados de saída não sejam normalizados. Assim, procedeu-se à normalização dos valores correspondentes à variável independente do conjunto de treino com a função *MinMaxScaler*. A fórmula de cálculo da função apresenta-se de seguida, sendo que X , X_{max} , X_{min} e X_{scaled} representam respetivamente o valor real, o valor máximo e o valor mínimo do conjunto de dados e o valor normalizado.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Desta forma os dados de entrada variam entre o valor de 0 e 1, sendo que os dados de previsão de energia ativa, correspondentes à variável dependente, mantiveram-se com os valores reais. Para além disso, o conjunto de dados de validação e de teste foram normalizados com o valor máximo e mínimo do conjunto de dados de treino. Assim, o modelo tem a capacidade de generalizar corretamente para novos dados de entrada, uma vez que os valores são normalizados com a mesma escala com que foram os dados utilizados para treinar o modelo.

As redes neuronais foram testadas com recurso a várias bibliotecas de *Python* com destaque para as bibliotecas *Pandas* e *NumPY* para manipulação e análise de dados, *Matplotlib* e *Seaborn* para visualização de dados e *Scikit-learn*, *TensorFlow* e *Keras* para aplicação dos

modelos de aprendizagem computacional. Para além das bibliotecas mencionadas, atendendo à grande quantidade de modelos que se pretende testar, foi também utilizada a biblioteca *Optuna*, aplicada para automatizar o processo de busca pelos melhores hiperparâmetros. Utilizando esta ferramenta é possível testar uma rede neuronal com vários hiperparâmetros e configurações de rede de forma automatizada objetivando a minimização de uma métrica de erro que se selecione.

Todos os treinos realizados manualmente para uma análise inicial, assim como os treinos que envolveram a utilização da biblioteca *Optuna*, foram realizados com o conjunto de treino e o conjunto de validação. Só após a seleção dos melhores modelos para cada tipo de rede neuronal é que se procedeu ao teste da rede com o conjunto de teste. Esta divisão entre treino, validação e teste garante que o modelo seja treinado de forma adequada, ajustado corretamente e avaliado de forma imparcial, permitindo que o conjunto de teste obtenha uma estimativa imparcial do desempenho do modelo.

Tendo em conta o problema que se pretende resolver, todas as redes neuronais foram criadas com apenas um neurónio como camada de saída sem nenhuma função de ativação, sendo o resultado deste o produto escalar das entradas proveniente da camada oculta anterior. Assim, encontram-se resumidos todos os parâmetros testados com o *Optuna* na Tabela 10. Deve referir-se ainda que, de acordo com a forma como os dados entraram na rede, isto é, por exemplo, para o primeiro cenário onde se quer prever o consumo do 8.º dia, os dados de entrada correspondem a um vetor com 7 observações, as quais correspondem ao consumo energético dos 7 dias anteriores. Assim, para o caso das MLP, introduziu-se uma camada *Flatten*, a qual tem em conta a estrutura dos dados de entrada da rede. Esta camada tem como objetivo transformar os dados numa sequência unidimensional para que sejam processados pelas camadas seguintes da MLP.

Tabela 10 - Arquiteturas e hiperparâmetros testados com o Optuna

Arquitetura				Hiperparâmetros		Treino
Quantidade de camadas ocultas	Quantidade de neurónios em cada camada oculta	Tipo de neurónio	Camada de saída	Funções de ativação dos neurónios das camadas ocultas	Optimizers	Batch size
1, 2 e 3	16, 32, 64, 128, 256, 512	Dense, SimpleRNN, GRU e LSTM	Dense(1)	tanh, relu e sigmoid	adam e rmsprop	16, 32 e 64

Durante os treinos das redes neuronais foram utilizados *callbacks*, que são objetos que permitem realizar ações específicas em momentos determinados durante o treino de uma rede neuronal e são úteis para acelerar os processos de treino, objetivando a obtenção de um modelo parcimonioso. Estes estão presentes na biblioteca *Keras* e os dois *callbacks* utilizados foram o *EarlyStopping* e o *ModelCheckpoint*. O primeiro permite estipular um valor de *patience*, por exemplo de 20, e faz com que o treino corra todas as épocas que foram determinadas até os resultados obtidos não melhorarem nas próximas 20 épocas. Isto é, pode-se definir um total de 1000 épocas para treinar o modelo, contudo, se os modelos que estiverem a ser obtidos não apresentarem melhorias 20 épocas consecutivas, o treino é terminado. Assim, se um modelo que estiver a ser treinado termine na época 300, e tendo em conta que o *patience* foi determinado para 20, conclui-se que o melhor modelo alcançado pelo treino dessa rede corresponde ao modelo obtido na época 280. Já o segundo *callback*, o *ModelCheckpoint*, permite que o treino do modelo monitorize uma métrica de desempenho e crie um ficheiro para guardar o melhor modelo de um treino, garantindo desta forma que o modelo guardado foi o melhor do treino executado e que pode ser utilizado posteriormente.

A utilização da biblioteca *Optuna* permite visualizar o impacto de cada hiperparâmetro no resultado obtido, sendo este um aspeto relevante da utilização desta ferramenta para a análise de redes neuronais. Para o leitor interessado, encontram-se a título de exemplo no Anexo V os resultados gráficos do *Optuna* para uma rede neuronal MLP com uma arquitetura de três camadas ocultas, a qual será apresentada mais à frente.

Tendo-se testado com o *Optuna* todos os tipos de arquitetura de rede com os conjuntos de dados de treino e validação apenas com a variável da energia ativa, assim como a aplicação de diferentes hiperparâmetros para cada uma das arquiteturas testadas, para os quatro cenários anteriormente apresentados, são posteriormente selecionados os melhores modelos para treino com os conjuntos de treino e de teste, os quais são apresentados nos seguintes capítulos para cada tipo de rede neuronal. Nos Anexos VI e VII apresentam-se todos os resultados obtidos com a aplicação desta ferramenta, assim como os resultados das métricas de erro para cada processo de treino.

6.2.1. Rede Neuronal – MLP

Na aplicação da MLP para previsão do consumo energético do dia seguinte foram escolhidos os melhores modelos para cada cenário obtidos através do *Optuna*, os quais se apresentam na Tabela 11.

Através da análise da tabela seguinte, é possível verificar que o modelo com o melhor desempenho foi obtido com o cenário C2, sendo este a utilização dos 14 dias anteriores para previsão do consumo energético do décimo quinto. Este modelo é ligeiramente melhor que o seu semelhante C1, contudo, ambos são bastantes melhores comparativamente aos modelos onde se utilizaram os dados passados semanais. O modelo MLP obtido com o conjunto de treino e de validação é constituído com duas camadas ocultas, com 512 e 256 neurónios na primeira e na segunda camadas, respetivamente, ambas com a função de ativação *relu*.

Tabela 11 - Resultados *Optuna* - MLP

Cenário	Unidades	Função ativação	Optimizer	MSE	RMSE	MAE	MAPE	R ²
C1	Flatten() + Dense(64) + Dense(64) + Dense(512)	relu + tanh + relu	adam	3828,06	61,87	39,76	0,1741	0,7360
C2	Flatten() + Dense(512) + Dense(256)	relu + relu	adam	3689,33	60,74	38,86	0,1626	0,7456
C3	Flatten() + Dense(512) + Dense(16) + Dense(512)	relu + sigmoid + relu	adam	10003,82	100,02	79,08	0,3504	0,3100
C4	Flatten() + Dense(64) + Dense(512)	relu + relu	adam	9984,89	99,92	74,54	0,3338	0,3113

Tendo-se obtido o melhor modelo com os conjuntos de treino e validação, posteriormente procedeu-se ao treino do modelo com o conjunto de treino utilizando o conjunto de teste para avaliar o desempenho. Para isso, foram realizados 30 treinos onde a média das métricas de erro e os valores do melhor modelo se apresentam de seguida na Tabela 12.

Tabela 12 - Resultados MLP

Modelo	MSE	RMSE	MAE	MAPE	R ²
MLP11	2699,95	51,96	34,01	0,1495	0,8005
Média dos 30	2917,11	54,00	35,32	0,1570	0,7845

A média do erro percentual médio absoluto das 30 redes neuronais MLP é de 15,70%, tendo a melhor rede apresentado um MAPE de 14,95%. O progresso do treino desta rede neuronal e as previsões realizadas para o período correspondente ao conjunto de teste são visíveis respetivamente nas Figura 32 e 33.

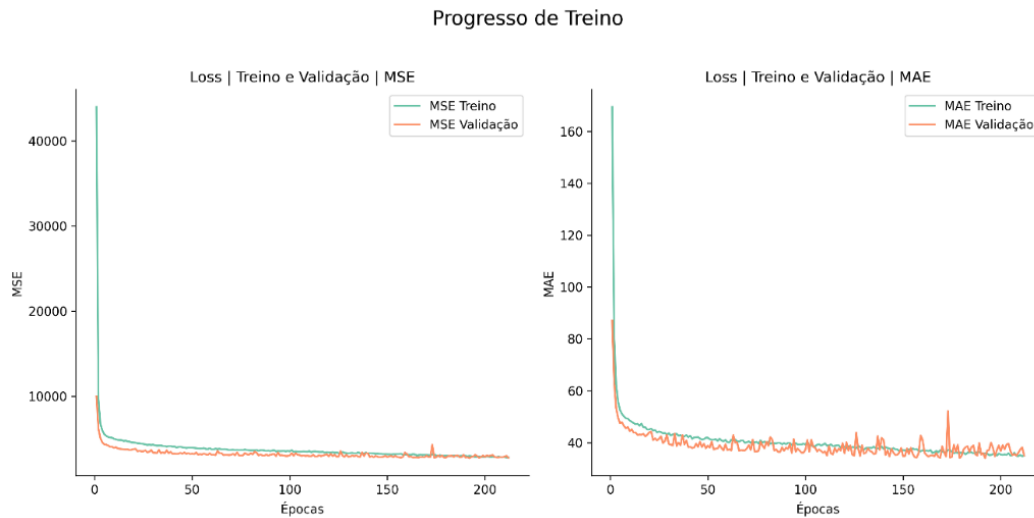


Figura 32 - Progresso de treino da MLP

Através da análise da Figura 31 podemos observar que a rede progrediu com o objetivo de minimizar tanto o MSE como o MAE, tendo sido alcançada uma época onde ambos os valores parecem ter estabilizado. Deve-se lembrar que para o treino foi utilizado o *EarlyStopping* com um *patience* de 20, o que quer dizer que o processo de treino é interrompido se se verificar que a métrica de avaliação não melhora durante 20 épocas consecutivas. Quanto à Figura 32, é possível verificar que o modelo conseguiu captar a tendência do consumo energético diário, isto é, as previsões mostram um consumo superior durante os dias da semana, sendo estes os picos na figura, assim como um consumo inferior durante os fins de semana, correspondentes aos vales na figura. Para além disso, as previsões obtidas captaram corretamente o consumo energético registado durante a época festiva do Natal, assim como a pausa letiva durante o mês de agosto, embora esta última tenha sido prevista de uma forma que não representa precisamente a realidade. Podemos ainda observar que, as previsões efetuadas possuem uma precisão superior a prever os mínimos dos consumos energéticos diários do que os máximos.

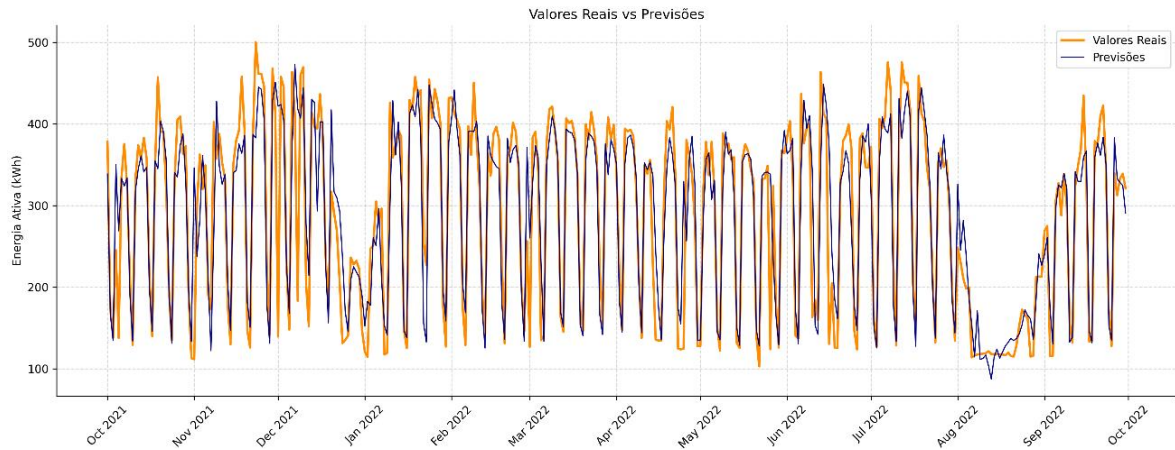


Figura 33 - Previsões vs valores reais da MLP

6.2.2. Rede Neuronal – RNN

Quanto às redes neuronais RNN, para os 4 cenários testados, a que apresentou melhores resultados com o *Optuna* foi a rede com a arquitetura apresentada na seguinte Tabela 13 para o primeiro cenário.

Tabela 13 - Resultados Optuna - RNN

Cenário	Unidades	Função ativação	Optimizer	MSE	RMSE	MAE	MAPE	R2
C1	SimpleRNN(128) + Dense(128)	tanh + relu	adam	4136,97	64,32	41,13	0,1729	0,7147
C2	SimpleRNN(64) + Dense(256)	sigmoid + relu	rmsprop	4245,80	65,16	42,56	0,1785	0,7072
C3	SimpleRNN(256) + SimpleRNN(512) + Dense(256)	tanh + tanh + tanh	rmsprop	10290,35	101,44	75,69	0,3438	0,2902
C4	SimpleRNN(128) + Dense(128)	tanh + relu	adam	10533,08	102,63	75,39	0,3610	0,2735

As redes RNN, de uma forma geral, apresentaram um desempenho pior que as redes MLP, tendo-se procedido ao treino da melhor rede obtida através de um ciclo de 30 iterações, sendo os resultados apresentados na Tabela 14.

Tabela 14 - Resultados RNN

Modelo	MSE	RMSE	MAE	MAPE	R2
RNN17	3420,96	58,49	36,83	0,1629	0,7472
Média dos 30	3476,76	58,96	37,53	0,1722	0,7431

A média destas redes apresentou um MAPE de 17,22%, sendo que a melhor apresenta um MAPE de 16,29%, ambos os valores inferiores às redes MLP. Nas Figuras 34 e 35 são apresentados, respetivamente, o progresso de treino desta rede e as previsões obtidas com o modelo. Podemos verificar que o treino deste modelo estabilizou de uma forma mais lenta que as previsões são ligeiramente piores que no modelo anterior, com o período correspondente ao período do Natal a última semana de julho/primeira de agosto com previsões bastante distantes da realidade.

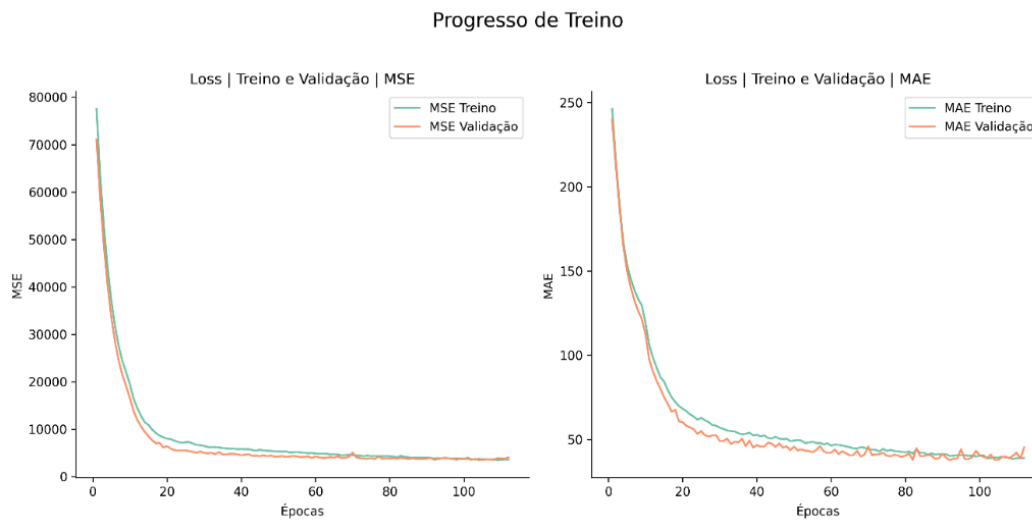


Figura 34 - Progresso de treino da RNN

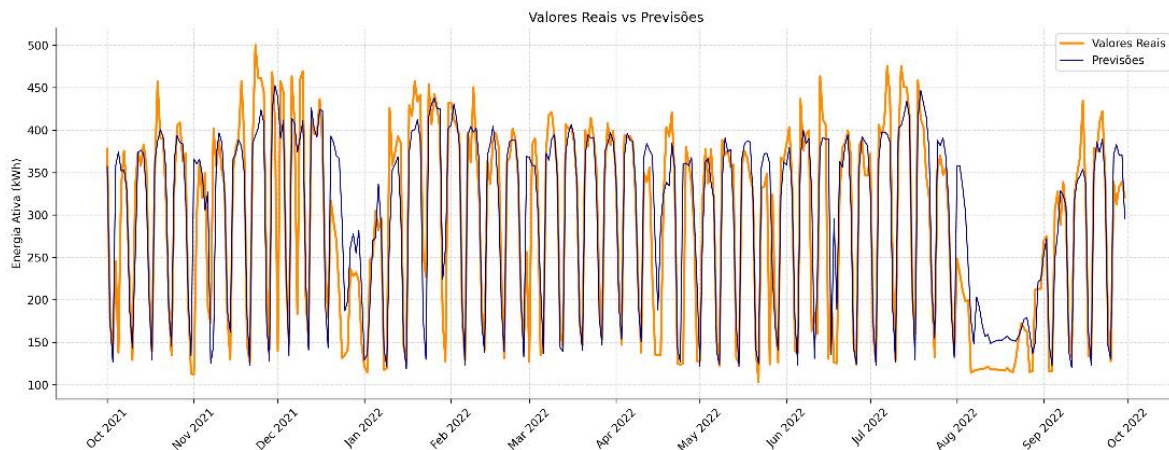


Figura 35 - Previsões vs valores reais da RNN

6.2.3. Rede Neuronal – LSTM

Para a rede neuronal LSTM, através do *Optuna* obteve-se a melhor rede com uma arquitetura bastante simples com apenas uma camada oculta com 256 neurónios para o cenário 1, estando os resultados apresentados na Tabela 15.

Tabela 15 - Resultados Optuna - LSTM

Cenário	Unidades	Função ativação	Optimizer	MSE	RMSE	MAE	MAPE	R ²
C1	LSTM(256)	tanh	rmsprop	4102,90	64,05	41,3732	0,1799	0,7170
C2	LSTM(256) + LSTM(128) + Dense(64)	tanh + sigmoid + relu	adam	4375,88	66,15	41,10	0,1802	0,6982
C3	LSTM(16)	sigmoid	adam	10492,24	102,43	79,25	0,3398	0,2763
C4	LSTM(64) + LSTM(128) + Dense(256)	sigmoid + relu + relu	rmsprop	10624,98	103,08	76,00	0,3352	0,2672

Esta rede, à semelhança das outras, foi treinada 30 vezes com o conjunto de treino e o conjunto de teste, tendo-se obtido um MAPE de 17,40% para a média de todos os modelos e de 16,25% para o melhor modelo obtido, de acordo com a Tabela 16.

Tabela 16 - Resultados LSTM

Modelo	MSE	RMSE	MAE	MAPE	R2
LSTM11	3373,31	58,08	36,65	0,1625	0,7508
Média dos 30	3490,63	59,07	37,84	0,1740	0,7421

Já o progresso de treino, visível na Figura 36, assim como as previsões na Figura 37, não apresentam diferenças significativas comparativamente aos modelos anteriores.

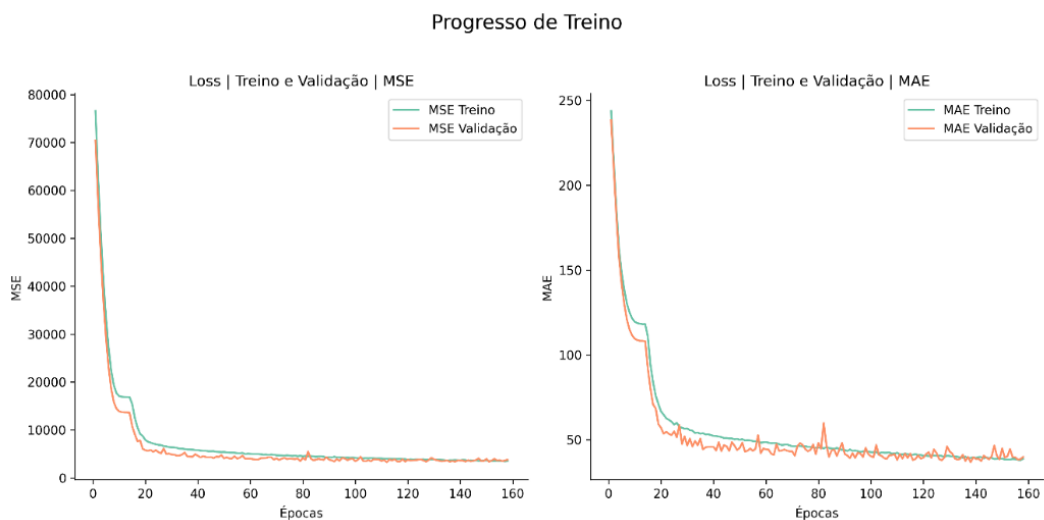


Figura 36 - Progresso de treino da LSTM

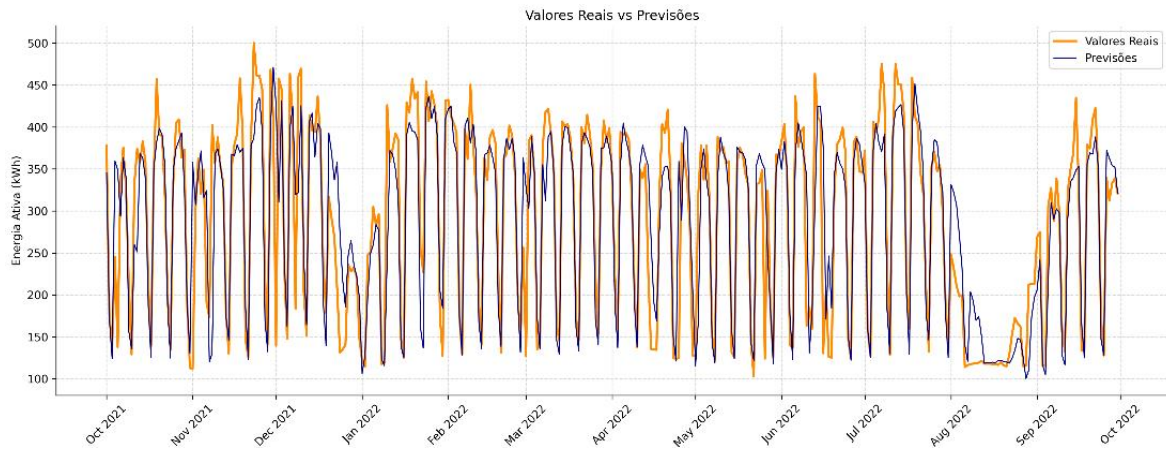


Figura 37 - Previsões vs valores reais da LSTM

6.2.4. Rede Neuronal – GRU

Relativamente às redes neuronais GRU, podemos verificar na Tabela 17 que a que apresentou melhor resultado foi a rede com duas camadas ocultas com neurónios GRU e com, respetivamente, 32 e 512 unidades para o cenário 2.

Tabela 17 - Resultados Optuna - GRU

Cenário	Unidades	Função ativação	Optimizer	MSE	RMSE	MAE	MAPE	R ²
C1	GRU(64) + Dense(128)	tanh + relu	Adam	3852,60	62,07	39,4424	0,1734	0,7343
C2	GRU(32) + GRU(512)	tanh + tanh	rmsprop	4413,67	66,44	41,42	0,1705	0,6956
C3	GRU(128) + GRU(32) + Dense(512)	tanh + sigmoid + relu	Adam	10445,63	102,20	76,26	0,3469	0,2795
C4	GRU(64) + GRU(128) + Dense(256)	tanh + sigmoid + relu	rmsprop	10541,06	102,67	72,86	0,3317	0,2730

Esta arquitetura foi selecionada e o modelo foi treinado novamente 30 vezes, com o conjunto de treino e de teste, tendo os resultados de MAPE sido de 15,45% para a médias dos 30 modelos e de 14,58% para o melhor modelo obtido.

Tabela 18 - Resultados GRU

Modelo	MSE	RMSE	MAE	MAPE	R ²
GRU13	2905,54	53,90	33,80	0,1458	0,7853
Média dos 30	3107,49	55,72	35,12	0,1545	0,7704

O progresso do treino desta rede neuronal e as previsões efetuadas encontram-se visíveis nas Figuras 38 e 39.

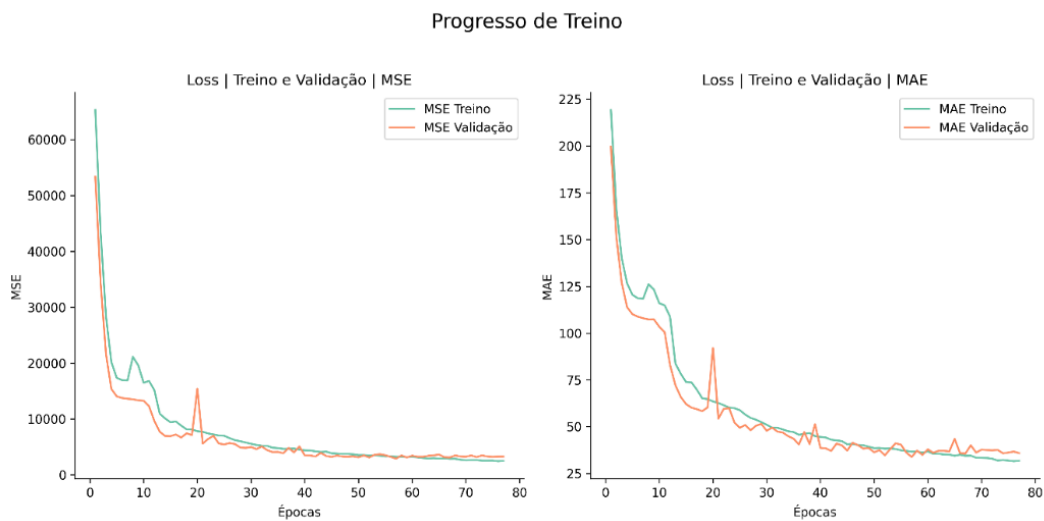


Figura 38 - Progresso de treino da GRU

É possível verificar que o treino com a rede neuronal GRU foi o mais curto em termos de épocas. Contudo, este estabilizou de uma forma mais errática que os anteriores, podendo este aspeto estar relacionado com os dados ou com os hiperparâmetros utilizados neste modelo. Quanto às previsões, é possível verificar que tanto a época de Natal como a pausa de agosto foram previstas com alguma precisão, embora a rede ainda falhe bastante em algumas situações, como por exemplo a última semana de julho, assim como alguns valores máximos e mínimos que se registam no decorrer da série temporal.

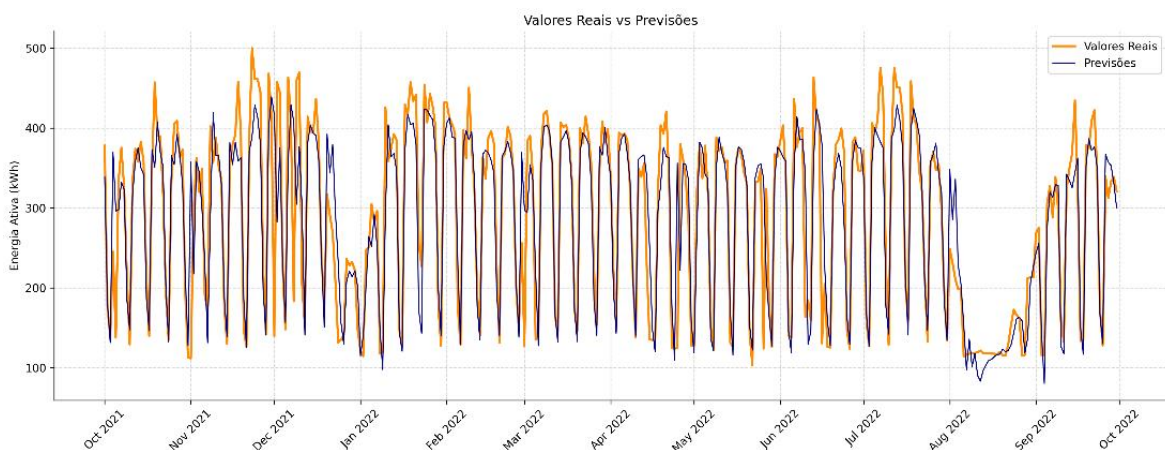


Figura 39 - Previsões vs valores reais da GRU

6.3. Análise de Resultados

A aplicação dos modelos mencionados para realizar previsões de consumos energéticos do dia seguinte, gerou resultados de MAPE que variam sensivelmente entre os 9% e os 18%, os quais, conjuntamente com os resultados das outras métricas seleccionadas estão visíveis na Tabela 19. O pior modelo obtido foi o SARIMA e o melhor modelo foi o SARIMAX com a utilização da variável endógena (energia ativa) e quatro variáveis exógenas, nomeadamente a energia reativa, ocupação total, e as temperaturas máxima e mínima.

Tabela 19 – Resumo dos resultados dos modelos utilizados para efetuar previsões

Cenário	Variáveis	Modelo	MSE	RMSE	MAE	MAPE	R ²
Previsão do dia à frente com base em todos os dados do passado	EnergiaA	SARIMA (1,0,0) (1,0,2) [7]	3408,43	58,38	40,27	0,1827	0,7482
	EnergiaA + OcupT	SARIMAX (2,0,0) (7,0,0) [7]	1692,69	41,14	30,74	0,1281	0,8749
	EnergiaA + OcupT + EnergiaR_I + Tmax + Tmin	SARIMAX (4,0,1) (1,0,7) [7]	968,28	31,12	22,89	0,0974	0,9284
C2	EnergiaA	MLP Flatten() + Dense(512) + Dense(256)	2699,95	51,96	34,01	0,1495	0,8005
C1	EnergiaA	SimpleRNN(128) + Dense(128)	3420,96	58,49	36,83	0,1629	0,7472
C2	EnergiaA	GRU(32) + GRU(512)	2905,54	53,90	33,80	0,1458	0,7853
C1	EnergiaA	LSTM(256)	3373,31	58,08	36,65	0,1625	0,7508

Porém, a utilização do SARIMAX com a adição de apenas uma variável exógena, a ocupação total diária, apresentou um modelo razoável com o MAPE de 12,81%, sendo que este exige um poder computacional bastante inferior ao modelo anterior o que poderá contribuir para a sua seleção.

Quanto às redes neuronais testadas, estas apresentaram um MAPE entre 14,58% e 16,29%, sendo a melhor a GRU e a pior a RNN, o que se traduz num desempenho inferior quando comparado com os métodos de base estatística para análise de séries temporais. Outro aspeto que deverá ser referido relativamente às redes neuronais e o seu desempenho, é o resultado obtido das duas melhores redes, nomeadamente a GRU e a MLP. Embora o MAPE seja inferior na GRU, podemos verificar que a MLP apresentou um MSE e um RMSE inferior, assim como um R² superior. Este facto, e embora os resultados obtidos sejam bastante

semelhantes para ambos os modelos, leva a afirmar que a MLP explica uma maior proporção da variabilidade dos dados.

Quando analisados os resultados das redes neuronais, foi possível verificar que os cenários que utilizaram dias passados consecutivos apresentaram sempre melhores resultados comparativamente aos modelos que utilizaram os dias passados semanais, tendo estes últimos apresentado um MAPE de sensivelmente 33%. Este facto poderá dever-se à imprevisibilidade de utilização dos equipamentos do edifício em análise.

Quando comparados os modelos SARIMA e SARIMAX, podemos verificar que o SARIMA apresenta um erro bastante grande, pelo que a sua utilização não fornece resultados com uma precisão considerável. Por outro lado, os modelos SARIMAX, apresentam resultados precisos, contudo, ainda longe do ideal, com a agravante de que o modelo SARIMAX com quatro variáveis exógenas exige a estimação de valores para as variáveis exógenas, algumas difíceis de estimar. Logo, o modelo mais parcimonioso e passível de se utilizar em contexto de previsão seria o modelo SARIMAX com a variável exógena da ocupação diária total do edifício, uma vez que este parâmetro é possível de ser estimado com alguma precisão, tendo em conta o planeamento das ações de formação e das turmas inscritas.

6.4. Caso de Retroanálise

O processo de previsão de consumos energéticos pode ser aplicado num caso de retroanálise, isto é, a aplicação de técnicas para previsão de consumos energético que envolve a modelação de dados históricos para estimar o consumo de energia de um determinado período passado. Este processo pode ser particularmente útil quando se deseja verificar a eficácia de um modelo de previsão em dados passados e entender como o modelo se teria comportado em períodos passados.

Uma vez que as redes neuronais, de forma geral, apresentam um desempenho superior quando alimentadas com um número maior de dados, e uma vez que os valores horários da energia ativa estão disponíveis decidiu-se utilizar este conjunto de dados para testar algumas redes neuronais. Assim, a utilização de modelos com os valores horários de consumos energéticos poderá resultar em previsões com um erro inferior. Contudo, tendo em conta que a rede foi criada com os valores horários, esta apenas tem a capacidade de prever o consumo efetivo da próxima hora, e não do próximo dia como os modelos anteriormente apresentados.

Desta forma, caso o modelo obtido possua um erro bastante pequeno, isto é, uma precisão de previsões bastante alta, este poderia ser utilizado para comparar com os valores reais com o objetivo de identificar desvios consideráveis. Estes desvios poderão representar eventuais comportamentos anómalos com o intuito de averiguar se existiu algum problema durante um certo período e, se sim, tentar identificar a sua causa.

Para esta análise utilizou-se a mesma metodologia aplicada ao longo do trabalho já descrito, com a experimentação de várias arquiteturas de rede e hiperparâmetros com recurso ao *Optuna* e com os conjuntos de treino e de validação. O Anexo VIII contém todos os resultados obtidos para todas as redes testadas para o leitor interessado.

A rede neuronal utilizada para o treino com os conjuntos de treino e de teste foi a RNN, tendo esta sido treinada 30 vezes com os resultados apresentados na Tabela 20. Para todos os conjuntos de previsão obtidos com as 30 redes neuronais treinadas, os valores das previsões foram agregados para os dias correspondentes, tendo os resultados destas somas sido comparados com os valores reais dos consumos energéticos diários. Os resultados das métricas de erro da operação mencionada encontram-se também apresentados na Tabela 20.

Podemos observar que a rede neuronal obteve erros inferiores com os valores horários quando comparados com os valores diários, com um MAPE na ordem dos 12%. Quando analisadas as métricas de erro para os valores de previsão horárias agregados em períodos diários, podemos observar que estas diminuíram substancialmente com a média do MAPE para os 30 modelos de 2,82% e um MAPE melhor do modelo RNN2 de 1,82%.

Tabela 20 - Resultados RNN para valores horários

<i>Modelo</i>	MSE	RMSE	MAE	MAPE	R2
<i>RNN12</i>	7,3379	2,7089	1,5316	0,1194	0,9025
<i>Média de 30</i>	7,5774	2,7522	1,6124	0,1309	0,8993
<i>Agregados os consumos a cada 24 horas</i>					
<i>Modelo</i>	MSE	RMSE	MAE	MAPE	R2
<i>RNN2</i>	43,4169	2,7964	4,8593	0,0182	0,9968
<i>Média de 30</i>	100,1189	2,7522	7,6279	0,0282	0,9926

De seguida apresenta-se a Figura 40 das previsões *versus* valores reais dos valores agregados por dia. Podemos observar que as previsões obtidas com este método se apresentam perfeitamente alinhadas com os valores reais, tendo um pior desempenho apenas em alguns

valores máximos da série temporal. Contudo, relativamente a este método é importante referir que estas previsões não podem ser utilizadas para previsão do consumo energético do dia seguinte, uma vez que a rede neuronal não foi criada dessa forma, isto é, uma vez que a rede neuronal utiliza os 7 valores horários passados de energia ativa para prever o consumo energético da hora seguinte e, por isso, esta quando introduzida numa aplicação requererá sempre os valores horários do consumo energético das 7 horas anteriores para prever a próxima. Assim, é expectável que, para se conseguir obter uma estimativa do consumo num determinado dia, seja necessário o consumo energético horário até à penúltima hora do dia que se pretende analisar, permitindo desta forma agregar posteriormente o consumo horário para um dia completo.

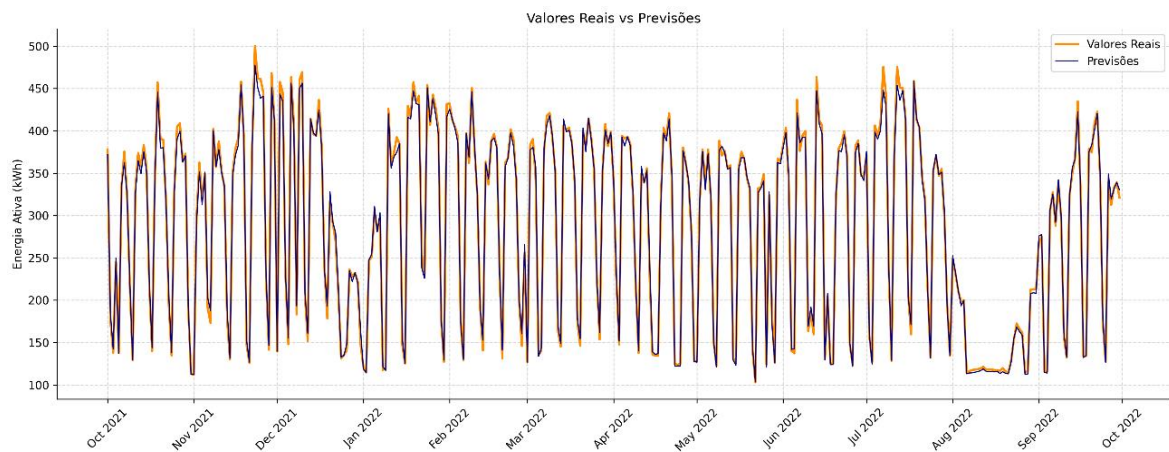


Figura 40 – Previsões vs valores reais da RNN com os valores horários agregados por dia

Embora a utilização desta rede neuronal não seja possível para o exercício de previsão do consumo energético diário, é de salientar que para uma análise do consumo energético do passado apresenta um erro bastante reduzido. É possível verificar que em nenhum momento do horizonte temporal analisado se identificam períodos onde a previsão e o consumo real de energia são significativamente diferentes, o que significa que não existiu nenhum evento onde o consumo se tenha identificado como anómalo. A utilização deste modelo apenas é possível com um modelo com um erro muito pequeno, com o objetivo de identificar eventos que possam ter um impacto significativo no consumo energético.

Esta página foi propositadamente deixada em branco

7. Conclusão

O presente projeto de mestrado investigou a previsão de consumos energéticos através da utilização de diferentes técnicas de modelação, com foco especial em redes neuronais, como a MLP, RNN, GRU e LSTM, além de outros modelos de análise de séries temporais como o SARIMA e o SARIMAX com diferentes configurações de variáveis exógenas.

Ao longo desta pesquisa, foi possível observar que a utilização de modelos para a previsão de consumos energéticos apresentou resultados interessantes, demonstrando uma capacidade significativa de capturar padrões complexos e não lineares nos dados de consumo energético. Além disso, a incorporação de variáveis exógenas permitiu aumentar a eficácia das previsões, levando em conta fatores externos que influenciam o consumo energético. A utilização de uma única variável exógena já demonstrou benefícios significativos, enquanto que a inclusão de múltiplas variáveis exógenas revelou-se ainda mais eficaz. Porém, a obtenção das variáveis exógenas de forma a permitir a realização de uma previsão é um facto que deverá ser considerado na aplicação de um modelo, uma vez que, a estimação da variável exógena se não for efetuada de forma fiável poderá prejudicar a precisão das previsões.

Os resultados deste trabalho destacam a importância de explorar diferentes técnicas de previsão de séries temporais para abordar problemas complexos, sendo praticamente impossível definir um modelo ótimo para diferentes casos de estudo. A seleção do modelo e a inclusão de variáveis exógenas relevantes podem aprimorar significativamente a qualidade das previsões, fornecendo informações valiosas para o planeamento e a gestão eficientes de recursos energéticos. Contudo, a qualidade das previsões depende fortemente do caso de estudo em análise. Para o edifício em estudo, é possível admitir que o planeamento das sessões de formação afeta significativamente o consumo expectável de energia elétrica no edifício. Uma vez que as formações iniciam mediante a inscrição de formandos, os consumos podem ser afetados de tal forma que o modelo não consegue prever apenas com a informação do consumo passado. Por exemplo, tendo em conta o início de uma sessão de formação em soldadura de 250 horas, e dependendo do horário estipulado, o qual por sua vez poderá ser todos os dias da semana ou mais desfasado com várias sessões por semana, a utilização dos equipamentos associados a esta formação poderá gerar padrões de consumo de energia elétrica que os modelos terão dificuldade em captar. Para além disso, o modelo também necessita de modelar todo o consumo considerado de padrão de utilização normal do

edifício, podendo este também ser alvo de alterações consoante os formandos estejam em período de estágio nas empresas ou não. Estes factos contribuem significativamente para os resultados obtidos durante a realização deste trabalho.

De todos os modelos obtidos, embora o SARIMAX com quatro variáveis exógenas tenha sido o que apresentou melhor resultado, consideramos que o modelo mais parcimonioso e passível de ser utilizado numa aplicação seja o SARIMAX com uma variável exógena, nomeadamente a variável do número de ocupantes total diário do edifício. Este modelo apresentou um MAPE sensivelmente de 12%, o que permitiria uma previsão aproximada do consumo energético do dia seguinte com a utilização do histórico do consumo energético até ao dia anterior e a ocupação prevista do dia seguinte, sendo esta uma variável exógena suscetível de ser prevista no dia anterior com alguma precisão. Porém, deve-se referir que a utilização do modelo SARIMAX implica a recriação do modelo dia após dia, uma vez que, os dados reais dos consumos energéticos podem ser obtidos com o intuito de aperfeiçoar a previsão obtida. No entanto, este aspeto configura-se um processo leve que pode ser facilmente automatizado.

Devem-se salientar também os resultados obtidos com a desfragmentação do *dataset* nos valores horários de consumo energético com a rede neuronal RNN. Embora o modelo obtido apenas permita a realização de previsões de consumos energéticos para a hora seguinte, este quando aplicado a uma análise retrospectiva do consumo registado poderá revelar eventos onde o consumo energético se encontrou bastante distante da previsão.

Por fim, é essencial salientar que, o campo da previsão de consumos energéticos continua a evoluir, com um gradual crescente de publicações sobre este tema e com bastantes variações de casos de estudo e formas de analisar o problema. Espera-se que este projeto tenha apresentado uma visão global do trabalho que se tem desenvolvido nesta área e que contribua para o desenvolvimento contínuo do conhecimento. No entanto, a fim de melhorar os resultados obtidos neste projeto apresentam-se de seguida algumas propostas para trabalho futuro que poderão melhorar o trabalho até aqui realizado.

7.1. Trabalho Futuro

Tendo os resultados obtidos revelado que as redes neuronais, de uma forma geral, tiveram um desempenho inferior aos modelos estatísticos SARIMAX, o trabalho futuro deverá incidir sobre as redes neuronais com o objetivo de melhorar os resultados obtidos. Das

tarefas que podem configurar-se como trabalho futuro, destaca-se o teste de um maior número de arquiteturas de redes neuronais, a utilização de um maior número de técnicas de ciência de dados como o *FB Prophet*, *NeuralProphet* ou uma rede neuronal *Transformer* e a utilização de técnicas de regularização de redes neuronais, como *L1*, *L2* e *Dropout*. Para além destas, e tendo sido utilizada uma configuração de rede de *many to one*, poderá testar-se uma configuração diferente, como a *many to many*, com o objetivo de averiguar o impacto que esta configuração originaria. Este modelo poderia utilizar os dados de consumo energético para prever mais do que um dia a fim de averiguar o impacto das suas previsões. Por fim, mas não menos importante, procurar treinar as redes neuronais com um conjunto maior de dados resulta, na maior parte das vezes, em resultados melhores, assim como a utilização de variáveis exógenas para os diferentes tipos e configurações de redes neuronais, como a utilização dos dias da semana como variável exógena para treinar o modelo através de técnicas de *One-hot encoding*, ou mesmo a criação de variáveis que especifiquem que equipamentos estão em utilização num determinado período.

Bibliografia

- [1] Observatório da Energia, DGEG – Direção Geral de Energia e Geologia, Direção de Serviços de Planeamento Energético e Estatística e ADENE – Agência para a Energia, Direção de Formação, Informação e Educação (2022) "Energia em Números".
- [2] Davis, H. T., & Nelson, W. F. C. (1935). "The Analysis of Time Series". In H. T. Davis & W. F. C. Nelson, "Elements of Statistics with Applications to Economic Data" (pp. 119–145). Principia Press.
- [3] De Gooijer, J. G. e Hyndman, R. J. (2006). "25 Years of Time Series Forecasting". "International Journal of Forecasting", 22(3), pp. 443–473.
- [4] Time series analysis: Definition, types, techniques, and when it's used (s.d.) "Tableau".
- [5] 6 real-world examples of tableau customers using time series analysis (s.d.) "Tableau".
- [6] Schlittgen, R. (2008). "Robert H. Shumway and David S. Stoffer: Time Series Analysis and Its Applications with R Examples, 2nd edn". "AStA Advances in Statistical Analysis", 92(2), 233–234.
- [7] Dickey, D. A. e Fuller, W. A. (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root". "Journal of the American Statistical Association", 74, 427-431.
- [8] Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1992). "Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root". "Journal of Econometrics", 54 (1-3), 159-178.
- [9] Morettin, P. A. e Tolo, C. M. C. (2006). "Análise de Séries Temporais". 2.^a Edição. São Paulo. Editora Edgard Blücher.
- [10] Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). "STL: A Seasonal-Trend Decomposition Procedure Based on Loess". "J. Off. Stat", 6(1), 3-73.
- [11] Peixeiro, M. (2022). "The Complete Guide to Time Series Analysis and Forecasting", Medium. "Towards Data Science".

- [12] Brownlee, J. (2020). "A Gentle Introduction to Autocorrelation and Partial Autocorrelation", "MachineLearningMastery.com".
- [13] Pandian, S. (2023). "Time Series Analysis and Forecasting: Data-driven Insights", "Analytics Vidhya".
- [14] Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). "DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks". "International Journal of Forecasting", 36(3), 1181–1191.
- [15] Gupta, S. C. (2022). "Machine Learning vs. Traditional Software Development (ML4Devs ..."
- [16] Mahesh, Batta. (2019). "Machine Learning Algorithms - A Review". 10.21275/ART20203995.
- [17] ABY, A. V. S. (2022). "Machine Learning Mind Map: Data Science and Machine Learning", Kaggle.
- [18] Silva, A. R. da (2022). "Uma Visão Geral sobre Machine Learning - Classificação", "Statplace a Estatística ao Alcance de Todos. Cursos e Consultoria.", Statplace.
- [19] LeCun, Y., Bengio, Y. & Hinton, G. (2015). "Deep Learning". "Nature", 521, 436–444.
- [20] Peixoto, F. (2020). "A Simple Overview of Multilayer Perceptron (MLP)", "Analytics Vidhya".
- [21] Schmidhuber, J. (2015). "Deep Learning in Neural Networks: An Overview". "Neural Networks", 61, 85-117.
- [22] Collis, J. (2017). "Glossary of Deep Learning: Bias", "Medium".
- [23] Ceccon, D. (2020). "Funções de Ativação: Definição, Características, E quando usar cada uma", "IA Expert Academy".
- [24] Zhou, V. (2022). "Machine Learning for Beginners: An Introduction to Neural Networks", "Victor Zhou".
- [25] A Look at IEA's New Global Energy Forecast (2022). "Global Energy Institute".

- [26] Iruela, J. R. (2021). "A TensorFlow Approach to Data Analysis for Time Series Forecasting in the Energy-Efficiency Realm", "Energies", 14(13), p. 4038.
- [27] Sun, Y., Haghghat, F. and Fung, B. C. M. (2020). "A Review of the-State-of-the-Art in Data-Driven Approaches for Building Energy Prediction", "Energy and Buildings", 221, p. 110022.
- [28] Runge, J. and Zmeureanu, R. (2021). "A Review of Deep Learning Techniques for Forecasting Energy Use in Buildings", "Energies", 14(3), p. 608.
- [29] Deb, C. et al. (2017). "A Review on Time Series Forecasting Techniques for Building Energy Consumption", "Renewable and Sustainable Energy Reviews", 74, pp. 902–924.
- [30] Amasyali, K. & El-Gohary, N. (2022). "Hybrid Approach for Energy Consumption Prediction: Coupling Data-Driven and Physical Approaches". "Energy and Buildings", 259, 111758.
- [31] Dong, B., Li, Z., Rahman, S. M. M., & Vega, R. (2016). "A Hybrid Model Approach for Forecasting Future Residential Electricity Consumption". "Experts@Syracuse".
- [32] Vivas, E., Allende-Cid, H., & Salas, R. (2020). "A Systematic Review of Statistical and Machine Learning Methods for Electrical Power Forecasting with Reported MAPE Score". "Entropy", 22(12), 1412.
- [33] Ghalekhondabi, I., Ardjmand, E., Weckman, G. R., & Young, W. A. (2016). "An Overview of Energy Demand Forecasting Methods Published in 2005–2015". "Energy Systems", 8(2), 411–447.
- [34] Divina, F. et al. (2019). "A Comparative Study of Time Series Forecasting Methods for Short Term Electric Energy Consumption Prediction in Smart Buildings", "Energies", 12(10), p. 1934.
- [35] Runge, J. & Zmeureanu, R. (2019). "Forecasting Energy Use in Buildings Using Artificial Neural Networks: A Review". "MDPI".
- [36] Mosavi, A. & Bahmani, A. (2019). "Energy Consumption Prediction Using Machine Learning; A Review".

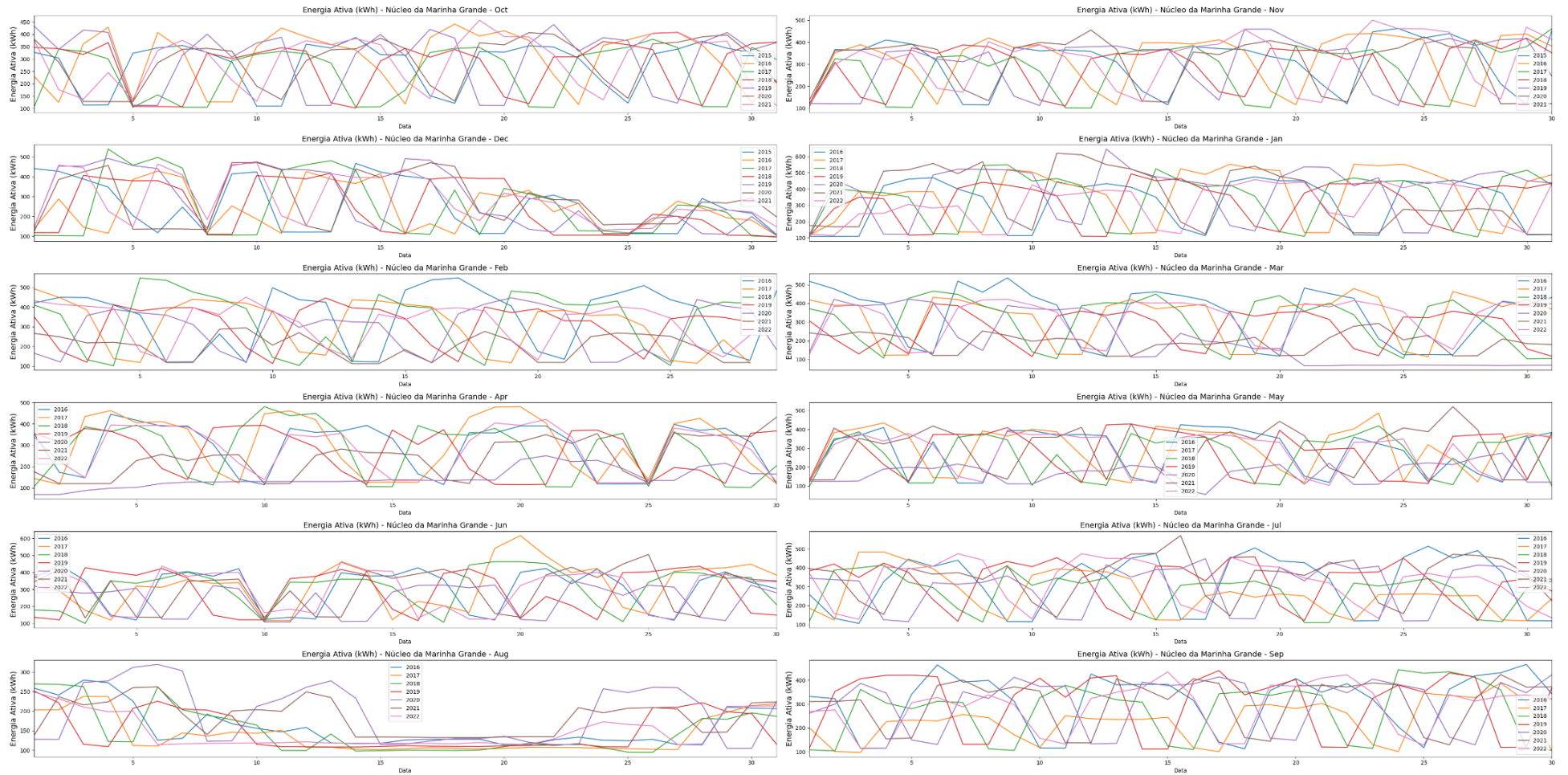
- [37] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). "Time Series Analysis: Forecasting and Control". John Wiley & Sons.
- [38] Newsham, G. R. & Birt, B. J. (2010). "Building-level occupancy data to improve ARIMA-based electricity use forecasts", Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building [Preprint].
- [39] Alharbi, F. R. & Csala, D. (2022). "A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time Series Approach". "Inventions", 7(4), 94.
- [40] Vagropoulos, S. I. et al. (2016). "Comparison of SARIMAX, Sarima, Modified Sarima and ANN-Based Models for Short-Term PV Generation Forecasting", 2016 IEEE International Energy Conference (ENERGYCON) [Preprint].
- [41] Friedrich, L., & Afshari, A. (2015). "Short-Term Forecasting of the Abu Dhabi Electricity Load Using Multiple Weather Variables". "Energy Procedia", 75, 3014–3026.
- [42] Nainwal, R., & Sharma, A. (2022). "Comparison of Multi Linear Regression and Artificial Neural Network to Predict the Energy Consumption of Residential Buildings". "IOP Conference Series: Earth and Environmental Science", 1084(1), 012005.
- [43] Pino-Mejías, R., Pérez-Fargallo, A., Rubio-Bellido, C., & Pulido-Arcas, J. A. (2017). "Comparison of Linear Regression and Artificial Neural Networks Models to Predict Heating and Cooling Energy Demand, Energy Consumption and CO2 Emissions". "Energy", 118, 24–36.
- [44] Luo, X. J. et al. (2020). "Comparative Study of Machine Learning-Based Multi-Objective Prediction Framework for Multiple Building Energy Loads", "Sustainable Cities and Society", 61, p. 102283.
- [45] Liu, Z. et al. (2019). "Accuracy Analyses and Model Comparison of Machine Learning Adopted in Building Energy Consumption Prediction", "Energy Exploration & Exploitation", 37(4), pp. 1426–1451.
- [46] Abdelaziz, A., Santos, V. e Dias, M. S. (2021). "Machine Learning Techniques in the Energy Consumption of Buildings: A Systematic Literature Review Using Text Mining and Bibliometric Analysis", "Energies", 14(22), p. 7810.

- [47] Li, C. et al. (2017). "Building Energy Consumption Prediction: An Extreme Deep Learning Approach", "Energies", 10(10), p. 1525.
- [48] Olu-Ajayi, R. et al. (2022). "Building Energy Consumption Prediction for Residential Buildings Using Deep Learning and Other Machine Learning Techniques", "Journal of Building Engineering", 45, p. 103406.
- [49] Seyedzadeh, S. et al. (2018). "Machine Learning for Estimation of Building Energy Consumption and Performance: A Review", "Visualization in Engineering", 6(1).
- [50] Admin (2023). "Seasonal Autoregressive Integrated Moving Average (SARIMA)", Helical IT Solutions Pvt Ltd.
- [51] AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference? (s.d.) "IBM".
- [52] Pai, A. (2023). "CNN vs. RNN vs. ANN - Analyzing 3 Types of Neural Networks in Deep Learning", "Analytics Vidhya".
- [53] Chen, G. (2021). "Recurrent Neural Networks (RNNs) Learn the Constitutive Law of Viscoelasticity", "Computational Mechanics", 67(3), pp. 1009–1019.
- [54] Loye, G. (2023). "Long Short-Term Memory: From Zero to Hero with PyTorch", "FloydHub Blog".
- [55] Hippert, H. S., Pedreira, C. E. e Souza, R. C. (2001). "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation", "IEEE Transactions on Power Systems", 16(1), pp. 44–55.
- [56] Wu, W. et al. (2019). "Using Gated Recurrent Unit Network to Forecast Short-Term Load Considering Impact of Electricity Price", "Energy Procedia", 158, pp. 3369–3374.
- [57] Zhang, G. P. (2003). "Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model", "Neurocomputing", 50, pp. 159–175.

Anexos

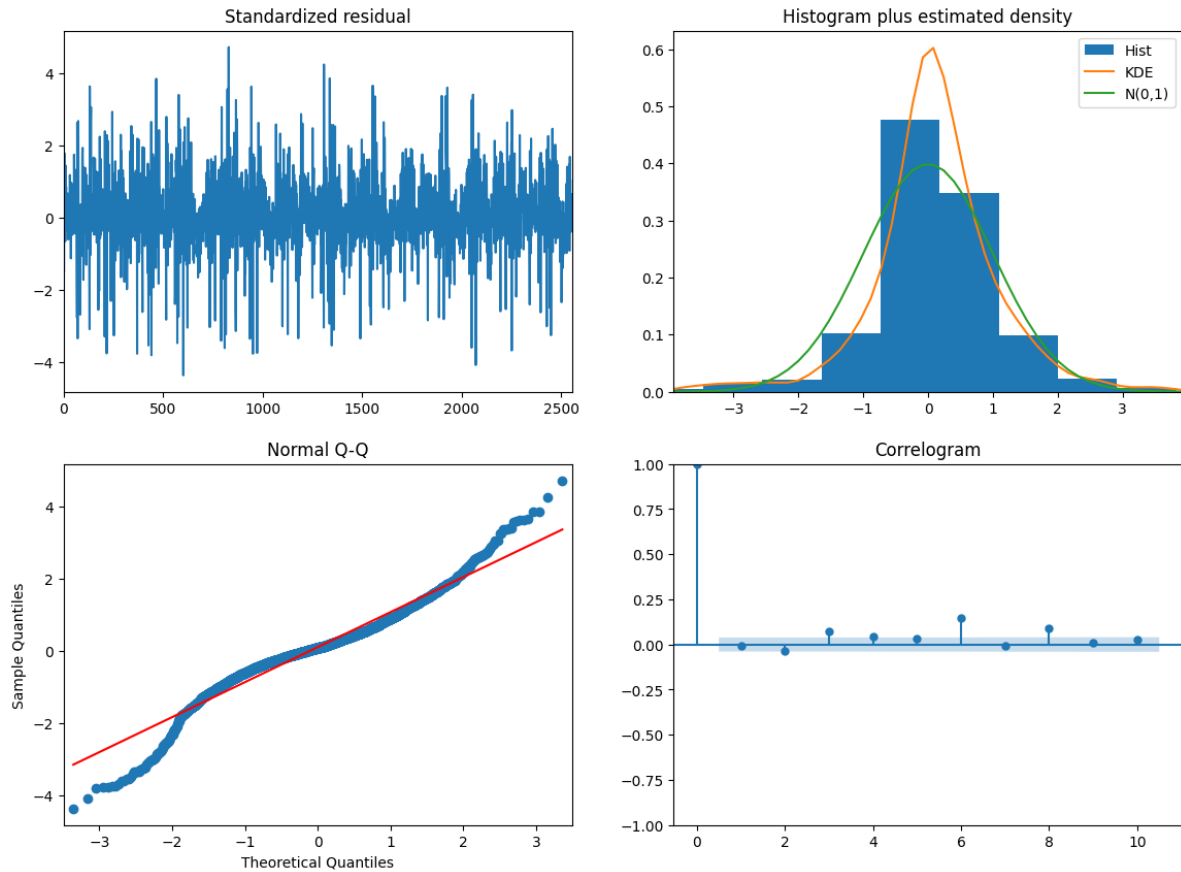
ANEXO I

Consumo energético diário por mês e por ano



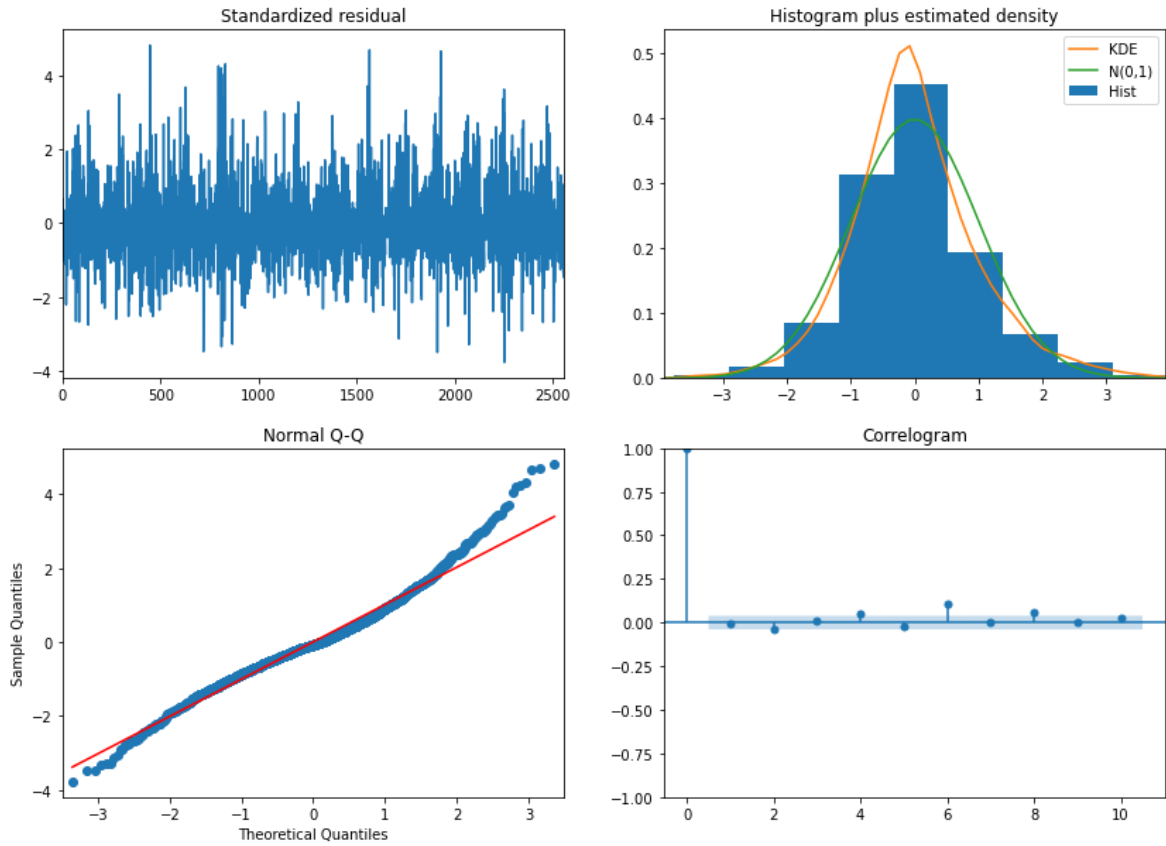
ANEXO II

Resíduos estandardizados, histograma dos resíduos, gráfico Q-Q distribuição normal dos resíduos e correlograma dos resíduos do modelo SARIMA



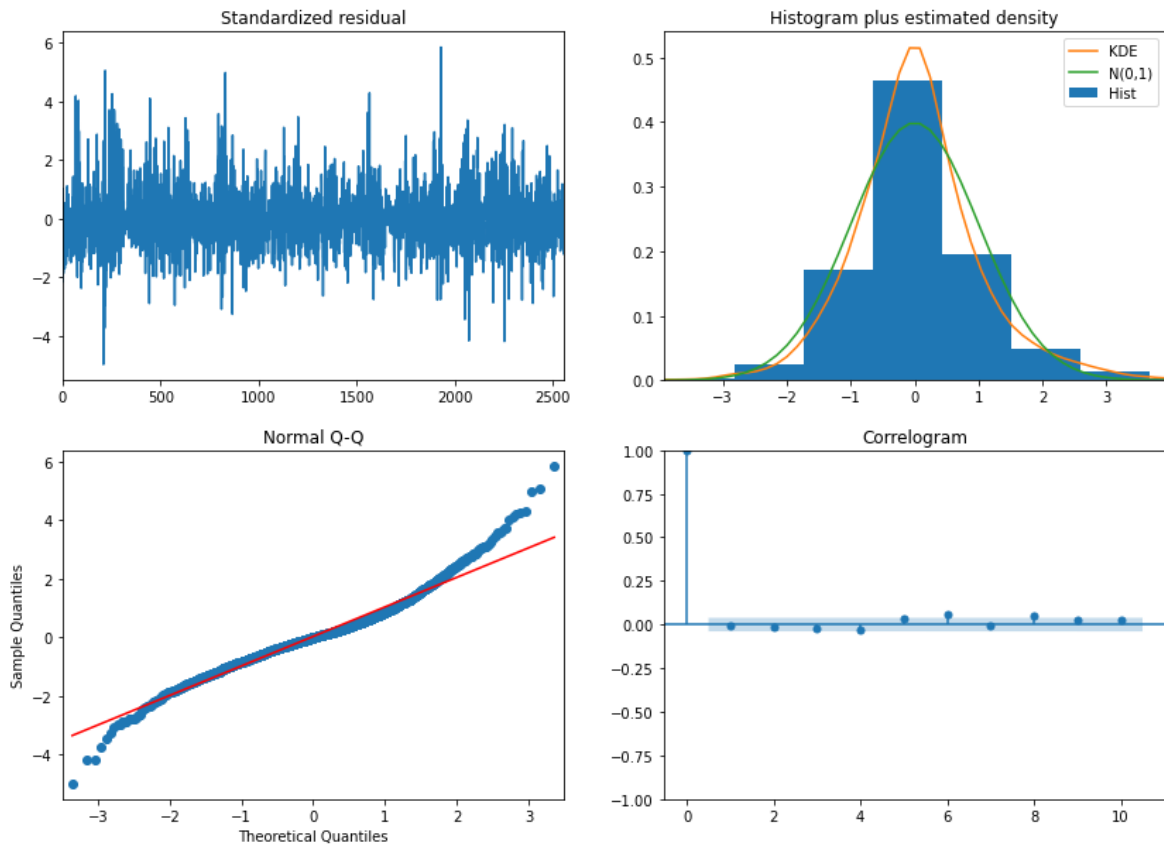
ANEXO III

Resíduos estandardizados, histograma dos resíduos, gráfico Q-Q distribuição normal dos resíduos e correlograma dos resíduos do modelo SARIMAX com 1 variável exógena



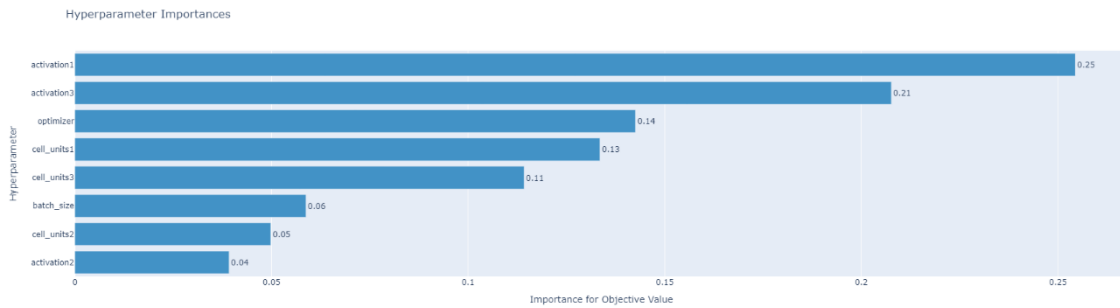
ANEXO IV

Resíduos estandardizados, histograma dos resíduos, gráfico Q-Q distribuição normal dos resíduos e correlograma dos resíduos do modelo SARIMAX com 4 variáveis exógenas

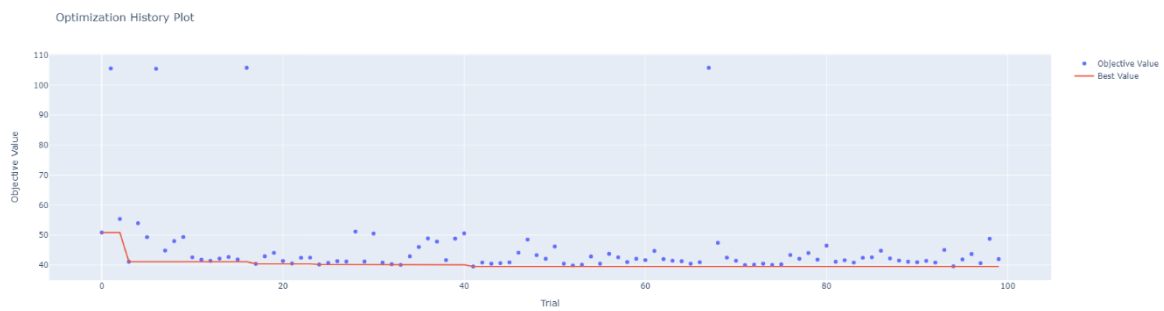


ANEXO V

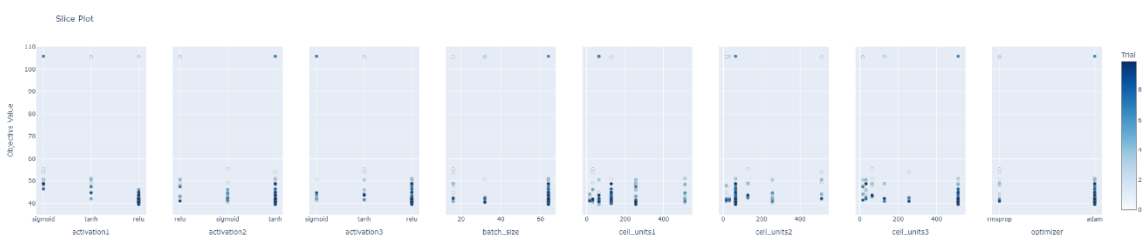
Importância de hiperparâmetros - Percentagem



Importância de hiperparâmetros - Histórico



Importância de hiperparâmetros – Gráfico de cada hiperparâmetro



ANEXO VI

Resultados obtidos com o OPTUNA para o C1 e C2

OPTUNA														
Dataset	Modelo	Unidades	Função ativação	Optimizer	Loss	Épocas - Total	Batch	Épocas - Terminus	Patience	Métrica				
Variável										MSE	RMSE	MAE	MAPE	R2
Utilização dos valores dos 7 dias anteriores para previsão do 8º														
1 variável: EnergiaA	Flatten()+Dense()+Dense(1)	256	relu	adam	mse	1000	32	806	20	4036,08	63,53	42,4341	0,1880	0,7216
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense(1)	16 + 128	relu+relu	rmsprop	mse	1000	32	484	20	3907,39	62,51	40,9767	0,1808	0,7305
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense()+Dense(1)	64+64+512	relu+tanh+relu	adam	mse	1000	64	233	20	3828,06	61,87	39,7614	0,1741	0,7360
1 variável: EnergiaA	SimpleRNN+Dense(1)	128	tanh	rmsprop	mse	1000	16	138	20	5463,47	73,92	48,6583	0,2146	0,6232
1 variável: EnergiaA	SimpleRNN+Dense()+Dense(1)	128+128	tanh+relu	adam	mse	1000	16	239	20	4136,97	64,32	41,1363	0,1729	0,7147
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense(1)	64+256	relu+tanh	adam	mse	1000	16	138	20	4166,24	64,55	40,3114	0,1736	0,7127
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense()+Dense(1)	32+16+512	relu+tanh+relu	rmsprop	mse	1000	32	135	20	3918,85	62,60	41,8421	0,1762	0,7297
1 variável: EnergiaA	GRU+Dense(1)	256	tanh	rmsprop	mse	1000	16	183	20	4306,94	65,63	41,7477	0,1841	0,7030
1 variável: EnergiaA	GRU+Dense()+Dense(1)	64+128	tanh+relu	adam	mse	1000	16	187	20	3852,60	62,07	39,4424	0,1734	0,7343
1 variável: EnergiaA	GRU+GRU+Dense(1)	256+512	relu+relu	adam	mse	1000	16	171	20	4023,87	63,43	40,5943	0,1760	0,7225
1 variável: EnergiaA	GRU+GRU+Dense()+Dense(1)	32+16+32	relu+tanh+relu	adam	mse	1000	16	164	20	4120,91	64,19	41,9542	0,1808	0,7158
1 variável: EnergiaA	LSTM+Dense(1)	256	tanh	rmsprop	mse	1000	16	146	20	4102,90	64,05	41,3732	0,1799	0,7170
1 variável: EnergiaA	LSTM+Dense()+Dense(1)	32+256	sigmoid+relu	rmsprop	mse	1000	16	480	20	4060,22	63,72	41,9440	0,1823	0,7200
1 variável: EnergiaA	LSTM+LSTM+Dense(1)	256+64	relu+relu	adam	mse	1000	32	331	20	4391,62	66,27	43,3090	0,1916	0,6971
1 variável: EnergiaA	LSTM+LSTM+Dense()+Dense(1)	64+128+128	sigmoid+tanh+relu	adam	mse	1000	16	235	20	3962,57	62,95	41,2989	0,1807	0,7267
Utilização dos valores dos 14 dias anteriores para previsão do 15º														
1 variável: EnergiaA	Flatten()+Dense()+Dense(1)	512	sigmoid	rmsprop	mse	1000	16	423	20	4317,22	65,71	44,67	0,1896	0,7023
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense(1)	512+256	relu+relu	adam	mse	1000	16	122	20	3689,33	60,74	38,86	0,1626	0,7456
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense()+Dense(1)	64+128+512	relu+relu+tanh	rmsprop	mse	1000	32	104	20	3839,52	61,96	39,18	0,1709	0,7352
1 variável: EnergiaA	SimpleRNN+Dense(1)	64	tanh	rmsprop	mse	1000	16	193	20	5442,62	73,77	46,93	0,1945	0,6246
1 variável: EnergiaA	SimpleRNN+Dense()+Dense(1)	64+256	sigmoid+relu	rmsprop	mse	1000	32	194	20	4245,80	65,16	42,56	0,1785	0,7072
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense(1)	16+32	relu+relu	adam	mse	1000	64	115	20	6473,47	80,46	53,37	0,2295	0,5536
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense()+Dense(1)	512+128+128	relu+tanh+relu	rmsprop	mse	1000	16	70	20	NAN	NAN	NAN	NAN	NAN
1 variável: EnergiaA	GRU+Dense(1)	16	relu	adam	mse	1000	32	79	20	8619,28	92,84	66,49	0,3020	0,4056
1 variável: EnergiaA	GRU+Dense()+Dense(1)	512+256	tanh+sigmoid	rmsprop	mse	1000	32	123	20	4458,87	66,77	43,04	0,1820	0,6925
1 variável: EnergiaA	GRU+GRU+Dense(1)	32+512	tanh+tanh	rmsprop	mse	1000	32	116	20	4413,67	66,44	41,42	0,1705	0,6956
1 variável: EnergiaA	GRU+GRU+Dense()+Dense(1)	64+512+64	relu+sigmoid+relu	rmsprop	mse	1000	16	103	20	4293,16	65,52	42,75	0,1884	0,7039
1 variável: EnergiaA	LSTM+Dense(1)	64	tanh	rmsprop	mse	1000	16	147	20	5161,14	71,84	46,85	0,1993	0,6441
1 variável: EnergiaA	LSTM+Dense()+Dense(1)	16+512	sigmoid+relu	adam	mse	1000	16	102	20	4398,11	66,32	44,18	0,1945	0,6967
1 variável: EnergiaA	LSTM+LSTM+Dense(1)	256+128	tanh+relu	adam	mse	1000	32	101	20	5317,30	72,92	46,83	0,2034	0,6333
1 variável: EnergiaA	LSTM+LSTM+Dense()+Dense(1)	256+128+64	tanh+sigmoid+relu	adam	mse	1000	32	121	20	4375,88	66,15	41,10	0,1802	0,6982

ANEXO VII

Resultados obtidos com o OPTUNA para o C3 e C4

OPTUNA														
Dataset	Modelo	Unidades	Função ativação	Optimizer	Loss	Épocas - Total	Batch	Épocas - Terminus	Patience	Métrica				
Variável										MSE	RMSE	MAE	MAPE	R2
Utilização dos valores das 4 semanas anteriores para previsão do dia correspondente da 5ª semana														
1 variável: EnergiaA	Flatten()+Dense()+Dense(1)	512	relu	adam	mse	1000	64	355	20	9933,17	99,67	78,41	0,3585	0,3149
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense(1)	64+64	relu+relu	adam	mse	1000	16	161	20	9895,03	99,47	77,64	0,3564	0,3175
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense()+Dense(1)	512+16+512	relu+sigmoid+relu	adam	mse	1000	32	54	20	10003,82	100,02	79,08	0,3504	0,3100
1 variável: EnergiaA	SimpleRNN+Dense(1)	512	sigmoid	rmsprop	mse	1000	16	222	20	10466,41	102,31	83,31	0,3918	0,2781
1 variável: EnergiaA	SimpleRNN+Dense()+Dense(1)	256+64	sigmoid+relu	adam	mse	1000	16	127	20	10038,41	100,19	80,44	0,3623	0,3076
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense(1)	512+64	sigmoid+relu	adam	mse	1000	64	195	20	11222,33	105,94	78,31	0,3641	0,2260
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense()+Dense(1)	256+512+256	tanh+tanh+tanh	rmsprop	mse	1000	32	118	20	10290,35	101,44	75,69	0,3438	0,2902
1 variável: EnergiaA	GRU+Dense(1)	512	relu	adam	mse	1000	16	167	20	10899,53	104,40	80,82	0,3675	0,2482
1 variável: EnergiaA	GRU+Dense()+Dense(1)	32+32	relu+relu	adam	mse	1000	16	172	20	10582,55	102,87	77,94	0,3504	0,2701
1 variável: EnergiaA	GRU+GRU+Dense(1)	16+32	tanh+relu	adam	mse	1000	16	288	20	10707,02	103,47	78,44	0,3597	0,2615
1 variável: EnergiaA	GRU+GRU+Dense()+Dense(1)	128+32+512	tanh+sigmoid+relu	adam	mse	1000	32	110	20	10445,63	102,20	76,26	0,3469	0,2795
1 variável: EnergiaA	LSTM+Dense(1)	16	sigmoid	adam	mse	1000	16	432	20	10492,24	102,43	79,25	0,3398	0,2763
1 variável: EnergiaA	LSTM+Dense()+Dense(1)	32+512	relu+relu	rmsprop	mse	1000	16	112	20	10509,91	102,52	78,78	0,3590	0,2751
1 variável: EnergiaA	LSTM+LSTM+Dense(1)	256+256	sigmoid+relu	adam	mse	1000	64	180	20	10700,72	103,44	77,60	0,3580	0,2619
1 variável: EnergiaA	LSTM+LSTM+Dense()+Dense(1)	16+128+32	tanh+tanh+relu	rmsprop	mse	1000	16	146	20	10907,18	104,44	79,53	0,3729	0,2477
Utilização dos valores das 8 semanas anteriores para previsão do dia correspondente da 9ª semana														
1 variável: EnergiaA	Flatten()+Dense()+Dense(1)	512	sigmoid	rmsprop	mse	1000	16	229	20	10385,85	101,91	80,95	0,3850	0,2837
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense(1)	64+512	relu+relu	adam	mse	1000	16	172	20	9984,89	99,92	74,54	0,3338	0,3113
1 variável: EnergiaA	Flatten()+Dense()+Dense()+Dense()+Dense(1)	128+128+256	relu+sigmoid+relu	rmsprop	mse	1000	32	58	20	9785,22	98,92	76,18	0,3600	0,3251
1 variável: EnergiaA	SimpleRNN+Dense(1)	16	tanh	adam	mse	1000	16	102	20	NAN	NAN	NAN	NAN	NAN
1 variável: EnergiaA	SimpleRNN+Dense()+Dense(1)	128+128	tanh+relu	adam	mse	1000	16	91	20	10533,08	102,63	75,39	0,3610	0,2735
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense(1)	32+16	tanh+sigmoid	adam	mse	1000	64	89	20	NAN	NAN	NAN	NAN	NAN
1 variável: EnergiaA	SimpleRNN+SimpleRNN+Dense()+Dense(1)	64+64+512	relu+tanh+relu	rmsprop	mse	1000	64	65	20	10558,04	102,75	76,30	0,3729	0,2718
1 variável: EnergiaA	GRU+Dense(1)	64	sigmoid	adam	mse	1000	16	354	20	10145,63	100,73	79,73	0,3531	0,3002
1 variável: EnergiaA	GRU+Dense()+Dense(1)	16+512	tanh+relu	rmsprop	mse	1000	32	198	20	10435,44	102,15	77,57	0,3422	0,2802
1 variável: EnergiaA	GRU+GRU+Dense(1)	128+512	tanh+tanh	rmsprop	mse	1000	64	104	20	9908,31	99,54	74,06	0,3467	0,3166
1 variável: EnergiaA	GRU+GRU+Dense()+Dense(1)	128+512+256	tanh+sigmoid+relu	rmsprop	mse	1000	16	80	20	10541,06	102,67	72,86	0,3317	0,2730
1 variável: EnergiaA	LSTM+Dense(1)	16	tanh	adam	mse	1000	16	396	20	10730,26	103,59	85,85	0,3944	0,2599
1 variável: EnergiaA	LSTM+Dense()+Dense(1)	128+128	tanh+relu	rmsprop	mse	1000	16	116	20	9947,28	99,74	76,31	0,3360	0,3139
1 variável: EnergiaA	LSTM+LSTM+Dense(1)	256+64	sigmoid+relu	adam	mse	1000	16	179	20	10192,13	100,96	77,90	0,3450	0,2970
1 variável: EnergiaA	LSTM+LSTM+Dense()+Dense(1)	64+128+256	sigmoid+relu+relu	rmsprop	mse	1000	16	130	20	10624,98	103,08	76,00	0,3352	0,2672

ANEXO VIII

Resultados obtidos com o OPTUNA para a Retro análise

OPTUNA														
Dataset	Modelo	Unidades	Função ativação	Optimizer	Loss	Épocas - Total	Batch	Épocas - Terminus	Patience	Métrica				
Variável										MSE	RMSE	MAE	MAPE	R2
Utilização dos valores das 7 horas anteriores para previsão da 8ª														
1 variável: EnergiaA	MLP + Dense(1)	128	relu	rmsprop	mse	1000	64	95	10	9,32	3,05	1,74	0,1471	0,8762
1 variável: EnergiaA	MLP + MLP + Dense(1)	512+512	relu+relu	adam	mse	1000	16	39	10	6,92	2,63	1,48	0,1238	0,9080
1 variável: EnergiaA	MLP + MLP + MLP + Dense(1)	64+256+256	relu+relu+sigmoid	adam	mse	1000	16	63	10	6,95	2,64	1,48	0,1217	0,9077
1 variável: EnergiaA	SimpleRNN + Dense(1)	256	relu	adam	mse	1000	32	37	10	6,88	2,62	1,53	0,1298	0,9086
1 variável: EnergiaA	SimpleRNN + Dense() + Dense(1)	128+128	sigmoid+sigmoid	adam	mse	1000	16	80	10	6,68	2,59	1,45	0,1161	0,9112
1 variável: EnergiaA	SimpleRNN + SimpleRNN + Dense(1)	64+256	relu+relu	adam	mse	1000	16	34	10	6,48	2,54	1,43	0,1156	0,9139
1 variável: EnergiaA	SimpleRNN + SimpleRNN + Dense() + Dense(1)	16+128+64	tanh+relu+relu	adam	mse	1000	16	43	10	6,57	2,56	1,45	0,1192	0,9126
1 variável: EnergiaA	GRU + Dense(1)	64	sigmoid	adam	mse	1000	32	43	10	8,01	2,83	1,71	0,1469	0,8935
1 variável: EnergiaA	GRU + Dense() + Dense(1)	256+128	relu+tanh	adam	mse	1000	16	40	10	6,74	2,60	1,46	0,1177	0,9105
1 variável: EnergiaA	GRU + GRU + Dense(1)	32+64	tanh+sigmoid	rmsprop	mse	1000	64	37	10	7,10	2,66	1,51	0,1224	0,9057
1 variável: EnergiaA	GRU + GRU + Dense() + Dense(1)	16+256+32	relu+sigmoid+relu	adam	mse	1000	16	44	10	7,11	2,67	1,49	0,1162	0,9055
1 variável: EnergiaA	LSTM + Dense(1)	512	relu	rmsprop	mse	1000	16	19	10	7,44	2,73	1,72	0,1609	0,9011
1 variável: EnergiaA	LSTM + Dense() + Dense(1)	128+32	sigmoid+tanh	adam	mse	1000	16	42	10	7,00	2,65	0,12	1,4979	0,9070
1 variável: EnergiaA	LSTM + LSTM + Dense(1)	512+32	sigmoid+relu	rmsprop	mse	1000	16	35	10	7,29	2,70	0,12	1,5137	0,9032
1 variável: EnergiaA	LSTM + LSTM + Dense() + Dense(1)	256+32+64	relu+tanh+relu	rmsprop	mse	1000	64	53	10	6,46	2,54	0,12	1,4449	0,9141