



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ – UFOPA
INSTITUTO DE ENGENHARIA E GEOCIÊNCIAS – IEG
BACHARELADO INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA**

YASMIN BRAGA TEXEIRA

**VALIDAÇÃO DE SENSOR DE CORRENTE EM UM SENSOR BOX IoT PARA
MONITORAMENTO DE ENRGIA ELÉTRICA**

SANTARÉM

2019

YASMIN BRAGA TEXEIRA

**VALIDAÇÃO DE SENSOR DE CORRENTE EM UM SENSOR BOX IoT PARA
MONITORAMENTO DE ENERGIA ELÉTRICA**

Trabalho de Conclusão de Curso apresentado ao Bacharelado Interdisciplinar em Ciência e Tecnologia para obtenção do grau de Bacharel em Ciência e Tecnologia na Universidade Federal do Oeste do Pará, Instituto de Engenharia e Geociências.

Orientador: Anderson Alvarenga De Moura Meneses

SANTARÉM

2019



SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DO OESTE DO PARÁ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO TECNOLÓGICA - PROPPIT
DIRETORIA DE PESQUISA
PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO CIENTÍFICA

RELATÓRIO TÉCNICO-CIENTÍFICO PIBIC/PIBITI

1. IDENTIFICAÇÃO

Bolsista: Yasmin Braga Teixeira

E-mail: yasminbragateixeira@gmail.com

Telefone: (93) 991964619

Título do Plano de Trabalho: Classificação de Padrões de Consumo em Edificações com Redes Neurais tendo em vista Eficiência Energética.

Título do Projeto ao qual está vinculado o plano de trabalho: Inteligência Computacional Aplicada a Recursos Energéticos: Modelagem e Análise de Padrões de Consumo e de Geração de Energia Solar Fotovoltaica.

Orientador: Anderson Alvarenga de Moura Meneses

E-mail do orientador: anderson.meneses@ufopa.edu.br

Telefone: (93) 991269294

Instituto: (ao qual o professor é vinculado) IEG

Bolsa: () PIBIC/UFOPA () PIBIC/FAPESPA () PIBIC/CNPq
() PIBITI/UFOPA () PIBITI/CNPq () PIBIC-AF/CNPq () PIBIC-AF/UFOPA
() PIBIC-AF/UFOPA – Indígena () PIBIC-AF/UFOPA - Quilombola

Vigência da bolsa: Agosto/2019 a Jul/2020

2. INTRODUÇÃO

Reduzir o consumo de energia e controlar o desempenho de energia elétrica em edifícios é um grande desafio nos dias atuais, visto que são eles que configuram as maiores taxas nesse contexto (DONG, XING e ZHANG, 2017). A eficiência energética dos prédios está relacionada com a diminuição na utilização de energia elétrica sem alterar o conforto dos ocupantes das edificações (NIKOLAOU, KOLOKOTSA e STAVRAKAKIS, 2011). Para a redução de energia, porém é necessário conhecer as configurações de cada prédio bem como os padrões energéticos de cada tipo de edificação. Esse reconhecimento é feito com bases de dados obtidas através do monitoramento do consumo (MORÁN, FUERTES, *et al.*, 2013). Assim, ao identificar os padrões de desempenho

energético e fazer a avaliação quanto a eficiência energética dos prédios é possível visualizar aqueles que necessitam de ações para redução de energia elétrica (FAYAZ e KIM, 2018).

A geração de grande quantidade de dados com o monitoramento de parâmetros de energia dos prédios abre caminho para a aplicação de Ciência de Dados (*Data Science*). (KOSELEVA e ROPAITE, 2017) afirmam que mais importante que a obtenção de dados em si, é crucial extrair significado desses dados, o que define Ciência de Dados segundo CIELEN, MEYSMAN e ALI (2016).

Ao extrair conhecimento dos dados é possível utilizá-los para previsões com Aprendizado de Máquina (*Machine Learning*) (KUBAT, 2017) que é um ramo da Inteligência Artificial caracterizado por algoritmos que aprendem com os dados.

Dentre os diversos algoritmos de aprendizado de máquina utilizados, este trabalho aborda a classificação utilizando o algoritmo K-Vizinho Mais Próximo (*K-Nearest Neighbor – KNN*) (ERTEL, 2017) e Redes Neurais Artificiais (RNAs) (AGGARWAL, 2018). O algoritmo KNN calcula a distância Euclidiana dos pontos e os classifica a partir de suas proximidades. Já as RNAs simulam neurônios artificiais com conjuntos de entradas e saídas em que cada conexão (neurônio) tem um peso associado (ELSALAMONY e ELSAYAD, 2013).

No contexto de RNAs existem as redes Perceptron Multi-Camadas (*Multi-Layer Perceptron; MLP*) com camadas ocultas adicionadas. A adição de camadas ocultas às RNAs gerando uma quantidade massiva de parâmetros a serem ajustados com a utilização de uma grande quantidade de dados é o que define Aprendizado Profundo (*Deep Learning*) (SKANSI, 2018) com RNAs. Porém, como especificado em PATTERSON e GIBSON (2017), são RNAs com mais parâmetros que aumentam seu poder com arquiteturas mais robustas.

Neste contexto, devido à importância econômica e social da eficiência energética e ao grande volume de dados gerados, faz-se cada vez mais necessário o uso de Deep Learning para análise de problemas e tomadas de decisão na área de Energia. Assim, o objetivo do trabalho é aplicar conhecimentos de análise de dados e Deep Learning na classificação de métricas energéticas de edifícios. Para isto, o *dataset* utilizado chama-se *2016 Building Energy Benchmarking* (<https://data.seattle.gov/dataset/2016-Building-Energy-Benchmarking/2bpz-gwpy>) e faz parte de um programa público da cidade de Seattle para redução de energia e de emissão de gases do efeito estufa com o acompanhamento e reporte anual do desempenho de energia. O *dataset* possui 46 colunas e 3376 linhas em que cada linha corresponde a um prédio. Após o pré-processamento do *dataset* percebeu-se um desbalanceamento na quantidade de dados da variável escolhida para o alvo, sendo necessário direcionar a pesquisa para estratégias que solucionassem esse problema. Sendo assim foi pesquisado dois métodos de equilibrar os valores dos dados, sendo eles *Oversampling* e *Undersampling*.

No método de *Oversampling* (LAST, DOUZAS e BACAO, 2017) os valores da classe minoritária são multiplicados no tamanho da classe majoritária igualando seus tamanhos. No método de *Undersampling* (LIU, WU e ZHOU, 2009) é selecionado uma amostra da classe majoritária na quantidade de amostras da classe minoritária do dataset.

A pesquisa busca seguir como em ROBINSON, DILKINA, *et al.* (2017) que utilizou *Machine Learning* para prever o Consumo Anual de prédios comerciais utilizando características de prédios de Nova York com dados disponibilizados pela Administração de Informações Energéticas da cidade. Além dele, busca-se basear em SAYADI, PATEL, *et al.* (2017) que abordou a previsão de configurações com Machine Learning que maximizem a eficiência energética dos prédios com variação de carga em seus diversos núcleos. E por fim, MURPHEY, GUO e FELDKAMP (2004) que utilizam RNAs para solucionar problemas de classificação em *datasets* desbalanceados comparando vários algoritmos e técnicas para treinamento e manipulação do *dataset* e obtenção de seu equilíbrio.

3. OBJETIVOS

- Objetivo geral:

O objetivo geral é classificar edifícios em residencial ou não residencial através da aplicação de RNAs do tipo MLP na análise de energia das construções prediais, fazendo uma comparação com o algoritmo KNN.

- Objetivos específicos:

- a) Revisão bibliográfica dos métodos de Undersampling e Oversampling e métricas para dados desbalanceados.
- b) Pré-processamento dos dados do *dataset*;
- c) Análise Exploratória de Dados;
- d) Treinamento e Teste da Rede Neural MLP e execução do KNN;
- e) Análise Estatística;

4. METODOLOGIA

A metodologia da pesquisa segue os procedimentos da Figura 1:

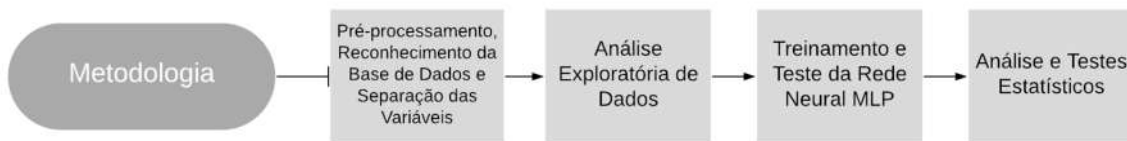


Figura 1 - Processos da Metodologia.

- Pré-processamento: Reconhecer a base de dados bem como identificar o problema a ser resolvido, assim, separar o *dataset* em uma matriz de variáveis preditoras e um vetor com a variável alvo a ser classificada e manipulação do dataset para resolver o problema de desbalanceamento com os métodos *oversampling* e *undersampling*. Para tais procedimentos de manipulação de dados a biblioteca a ser utilizada é a Pandas (<https://pypi.org/project/pandas/>).
- Análise Exploratória de Dados (SCHUTT e O'NEIL, 2014): Preparar os dados com modos de limpeza, transformação e combinação de dados com base na visualização com as bibliotecas *Matplotlib* e *Seaborn* (VANDERPLAS, 2017) e posteriormente dividir esses dados em conjunto de teste e treinamento.
- Treinamento e Teste da Rede MLP: Aplicar o modelo de RNAs, *KNN* e Aprendizado profundo nos dados processados ao utilizar bibliotecas do Python para aprendizado de máquina como *Scikit Learn* (*Sklearn* - <https://scikit-learn.org/stable/>) e biblioteca *Keras* (<https://keras.io/>).
- Análise Estatística: Aplicar métricas estatísticas como Acurácia, Precisão Recall e F1 Score.

Pré-processamento de dados

O primeiro procedimento realizado foi um estudo das variáveis do dataset, onde foram definidas as variáveis que não seriam relevantes para a avaliação e por isso, retiradas do dataset. Sendo assim, o dataset ficou com um tamanho de 3225 linhas por 18 colunas. A Tabela 1 mostra todas as colunas do dataset, onde *SecondLargestPropertyUseType*, *SecondLargestPropertyUseTypeGFA*, *ThirdLargestPropertyUseType*, *ThirdLargestPropertyUseTypeGFA*, *YearsENERGYSTARCertified*, *ENERGYSTARScore* são as colunas que possuem mais de 90% dos valores em *null values* e por isso foram apagadas e *OSEBuildingID*, *DataYear*, *PrimaryPropertyType*, *PropertyName*, *Address*, *City*, *State*, *ZipCode*, *TaxParcelIdentificationNumber*, *CouncilDistrictCode*, *Neighborhood*, *Latitude*, *Longitude*, *YearBuilt*, *ListOfAllPropertyUseTypes*, *LargestPropertyUseType*, *LargestPropertyUseTypeGF* são colunas que possuem informações irrelevantes para o treinamento, como ID do prédio, endereço, código postal, entre outros.

Tabela 1: Colunas do dataset.

Colunas do dataset	Colunas do dataset
OSEBuildingID	SecondLargestPropertyUseType
DataYear	SecondLargestPropertyUseTypeGFA
BuildingType	ThirdLargestPropertyUseType
PrimaryPropertyType	ThirdLargestPropertyUseTypeGFA
PropertyName	YearsENERGYSTARCertified
Address	ENERGYSTARScore
City	SiteEUI(kBtu/sf)
State	SiteEUIWN(kBtu/sf)
ZipCode	SourceEUI(kBtu/sf)
TaxParcelIdentificationNumber	SourceEUIWN(kBtu/sf)
CouncilDistrictCode	SiteEnergyUse(kBtu)
Neighborhood	SiteEnergyUseWN(kBtu)
Latitude	SteamUse(kBtu)
Longitude	Electricity(kWh)
YearBuilt	Electricity(kBtu)
NumberofBuildings	NaturalGas(therms)
NumberofFloors	NaturalGas(kBtu)
PropertyGFATotal	DefaultData
PropertyGFAParking	Comments
PropertyGFABuilding(s)	ComplianceStatus
ListOfAllPropertyUseTypes	Outlier
LargestPropertyUseType	TotalGHGEmissions
LargestPropertyUseTypeGF	GHGEmissionsIntensity

Definiu-se como variável alvo a coluna *Compliance Status* que indica se um prédio atingiu ou não o *benchmark* de energia do ano em questão. Essa estratégia foi definida pensando em aplicações locais em que pode ser feito tal estudo e definir as métricas para prédios da universidade.

O dataset continha ainda muitos dados corrompidos ou faltantes, sendo assim, foram analisadas separadamente as linhas que ainda continham *null values* com o objetivo de melhor preenchê-los ou deletá-los.

Desbalanceamento dos Dados

O *dataset* é considerado desbalanceado, ou seja, possui um desequilíbrio nas variáveis alvo, sendo 3195 com status *compliant* e 30 com status *non-compliant*. Com base nisso, definiu-se duas estratégias para análise de equilíbrio do dataset: *Oversampling* e *Undersampling*.

Na pesquisa foi utilizado o método de *undersampling*, onde é selecionado uma amostra de *compliant* na quantidade de amostras de *non-compliant*. Como o número de *non-compliant* é cerca de 30 amostras, o dataset fica com um tamanho de 60 linhas para dividir entre conjunto de teste e conjunto de treinamento.

Sem essa estratégia e com a característica de desbalanceamento do *dataset* porém, não seria possível avaliar o desempenho de classificação da rede com as métricas padrão, sendo elas acurácia ou precisão pois elas avaliam através da quantidade de acertos da classe positiva do *dataset*. Como no *dataset* em questão a classe positiva é majoritária seu resultado anularia o resultado da classe minoritária e assim, as métricas teriam valores altos, porém o que não pode ser considerado um bom desempenho.

Para isso, foi utilizado as métricas curva ROC (*Receiver Operating Characteristic Curve*) que relaciona a sensibilidade (taxa de verdadeiros positivos) com a especificidade (taxa de falsos positivos) e PRC (*Precision Recall Curve*) que relaciona as métricas precisão e recall, nos testes em que não houve balanceamento no *dataset* para fins de comparação de metodologias.

5. RESULTADOS OBTIDOS

Análise Exploratória de dados

A Figura 2 apresenta os tipos de prédios classificados pela base de dado. Sendo em sua maioria prédios não residenciais que configuram o maior consumo de energia da cidade de Seattle.

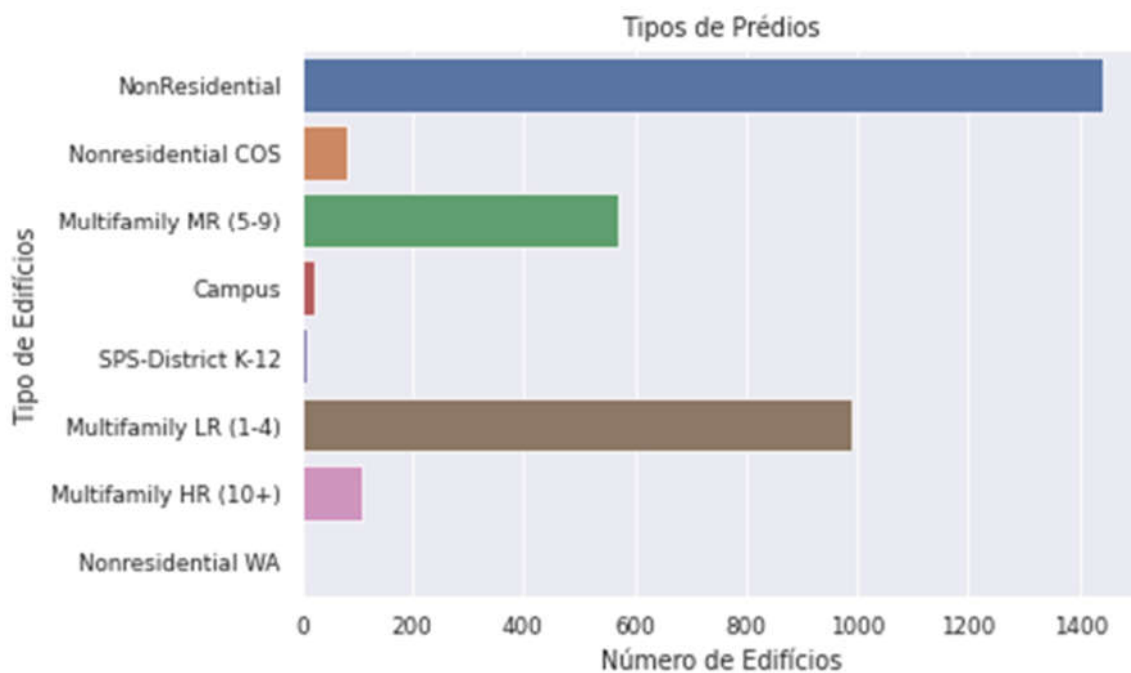


Figura 2: Classificação de Tipos de Prédios pela Base de Dados. Fonte: autor.

A Figura 3 relaciona o consumo de energia anual do tipo principal ocupado pela propriedade, o que se pode observar uma maior utilização de energia em Hospitais e Universidades.

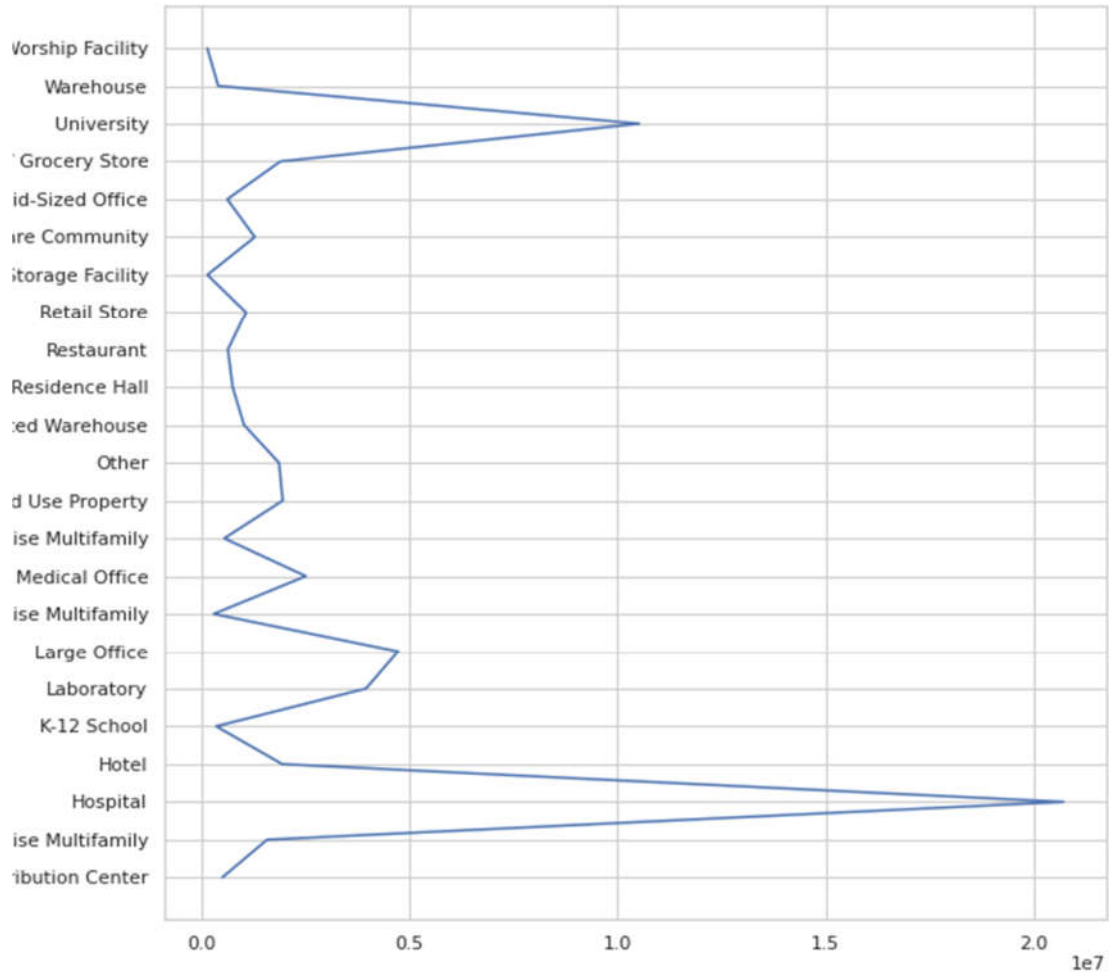


Figura 3: Relação entre Tipo Principal de Utilização das Propriedades com o Consumo de Energia Anual. Fonte: autor.

O gráfico da Figura 4 relaciona a interação entre o consumo anual de eletricidade de cada prédio, valores da coluna *SiteEnergyUse(kBtu)* e os valores de emissões totais de gases de efeito estufa representados na coluna *TotalGHGEmissions*. O gráfico mostra um aumento nas emissões de gases com o aumento do consumo de energia dos prédios.

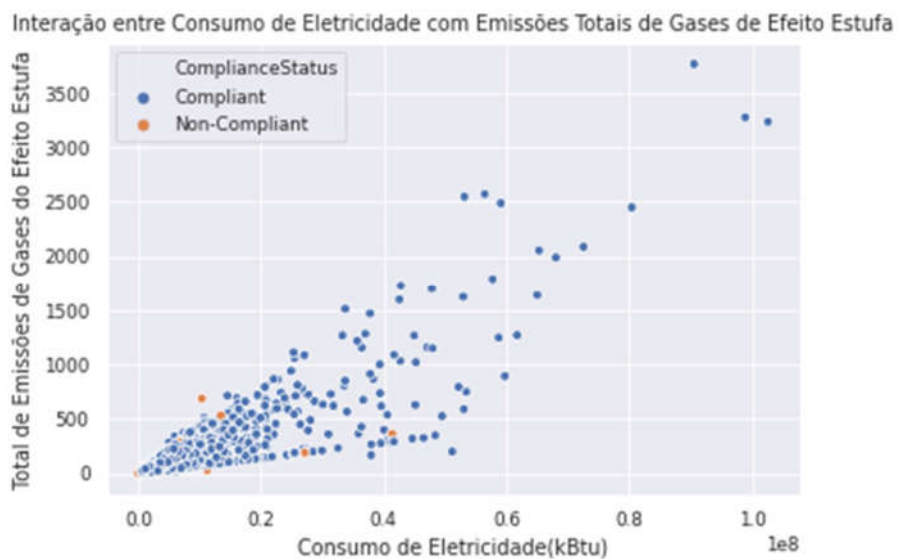


Figura 4: Relação entre Consumo de Eletricidade com Emissões Totais de Gases de Efeito Estufa de cada Propriedade. Fonte: autor.

A Figura 5 representa o gráfico que mostra o nível de desbalanceamento do *dataset* com 3195 valores *Compliant* e apenas 30 *non-compliant*.

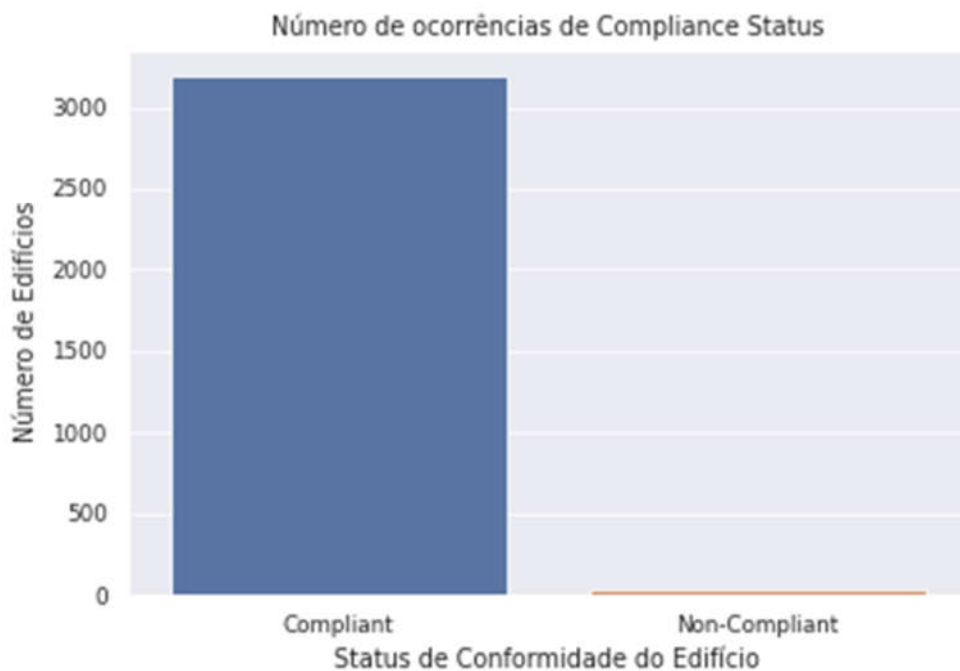


Figura 5: Contagem dos Valores da Coluna Compliance Status escolhida como variável alvo demonstrando nível de desbalanceamento entre variáveis. Fonte: autor.

Treinamento e Teste

- **Rede Neural MLP**

Para treinamento da rede neural MLP foi utilizada a biblioteca de aprendizado de máquina *Scikit-Learn* do Python. O modelo de Rede Neural utilizado foi o *MLPClassifier* que otimiza a função de perda por gradiente descendente estocástico. Foi utilizado uma camada oculta com 250 neurônios, uma com 150, uma com 200 neurônios e o número de iterações igual a 1000. Os resultados constam na Tabela 2 as curvas ROC e PRC respectivamente na Figura 6 e na Figura 7.

Tabela 2: Valores de Cross Val Score do teste no modelo de Redes Neurais com balanceamento do dataset.

Cross Val Score (cv=5)					Média
0.75	1	0.833	0.666	0.75	0.8

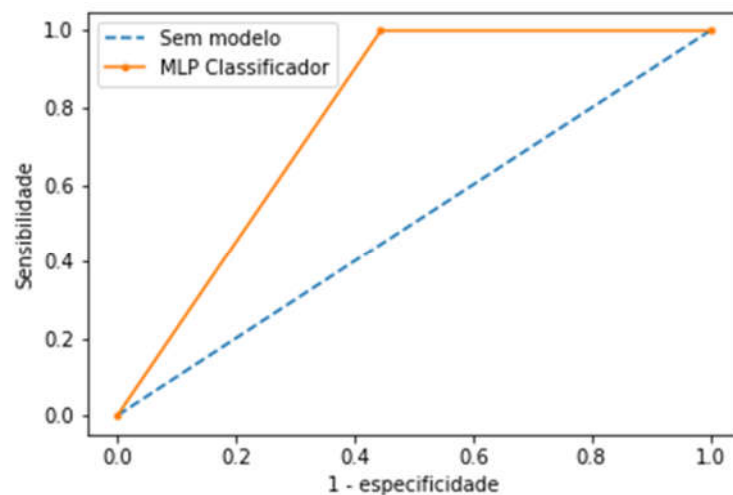


Figura 6: Curva ROC teste Redes Neurais com dataset balanceado. Fonte: Autor.

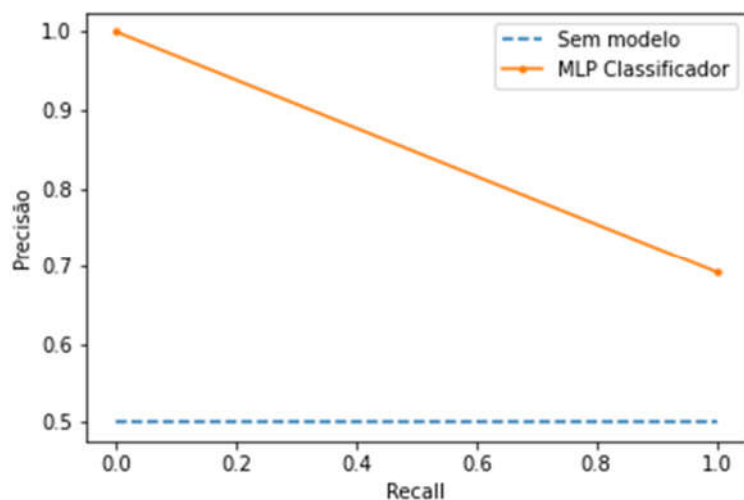


Figura 7: Curva PRC do teste com Redes Neurais e dataset com balanceamento. Fonte: Autor.

Sem a aplicação de *undersampling* os resultados seguem a Tabela 3 com os gráficos ROC na Figura 8 e PRC na Figura 9.

Tabela 3: Valores de Cross Val Score do teste no modelo de Redes Neurais sem balanceamento do dataset.

Cross Val Score (cv=5)					Média
0.979	0.989	0.990	0.989	0.990	0.987

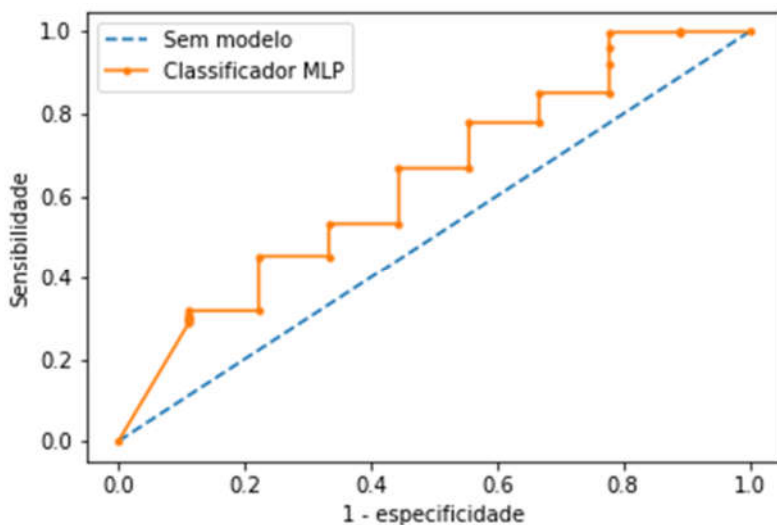


Figura 8: Curva ROC do teste com Redes Neurais e dataset sem balanceamento. Fonte: Autor.

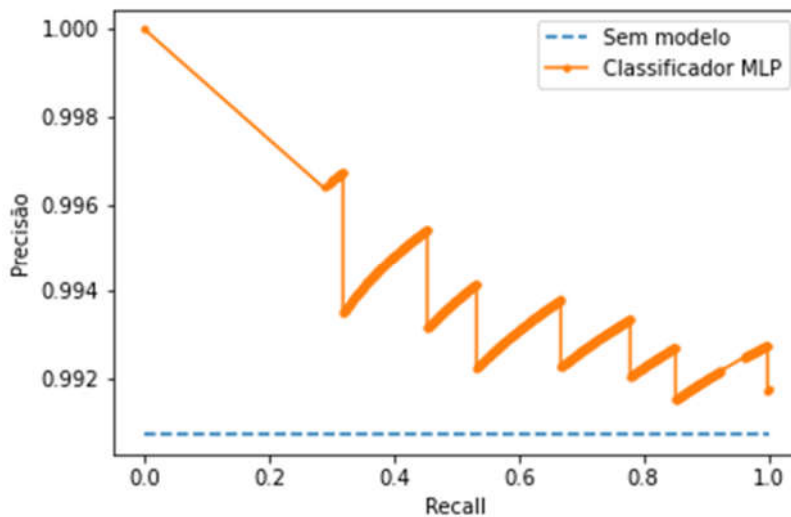


Figura 9: Curva PRC do teste com Redes Neurais e dataset sem balanceamento. Fonte: Autor.

- **KNN**

Para o treinamento com o KNN foi utilizado também a biblioteca *scikit-learn* com o *KNeighborsClassifier*. O número de vizinhos utilizados foi 5. Os resultados constam na Tabela 4 as curvas ROC e PRC respectivamente na Figura 10 e na Figura 11.

Tabela 4: Valores de Cross Val Score para o teste no modelo de KNN com balanceamento do dataset.

Cross Val Score (cv = 5)					Média
0.833	0.833	0.916	0.833	0.833	0.85

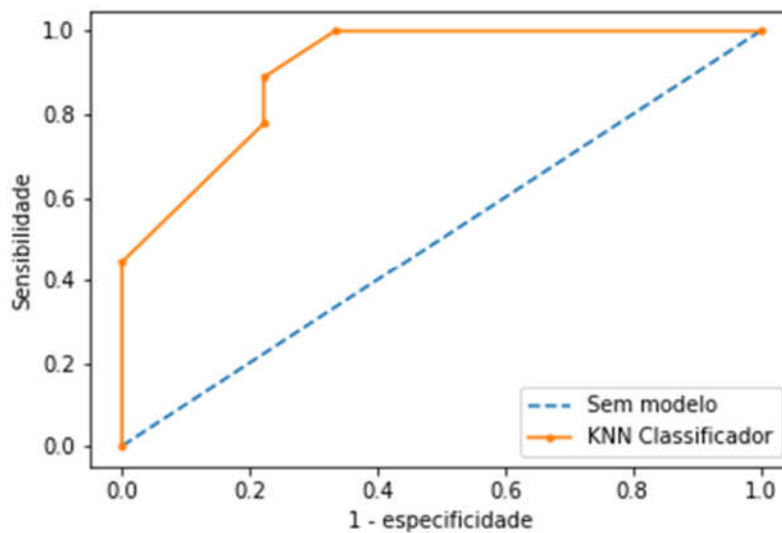


Figura 10: Curva ROC para o teste com o KNN com balanceamento. Fonte: Autor.

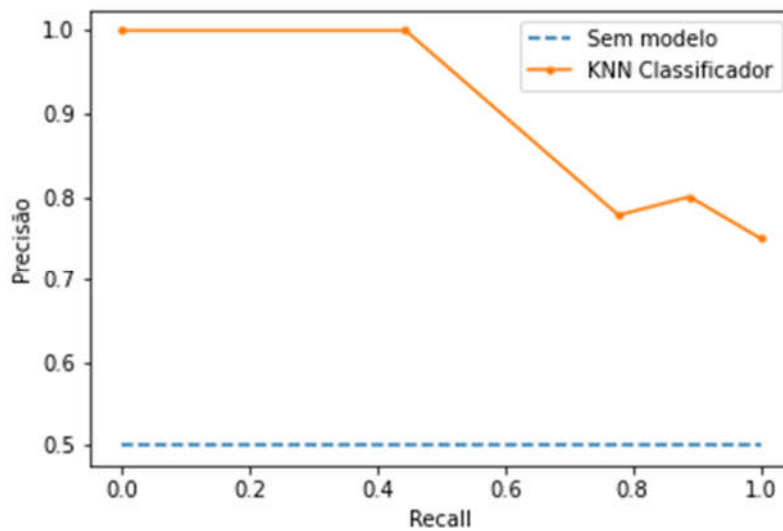


Figura 11: Curva PR para o teste com o KNN com balanceamento. Fonte: Autor.

Sem a aplicação de undersampling os resultados seguem a Tabela 5 com os gráficos ROC na Figura 12 e PRC na Figura 13.

Tabela 5: Valores de Cross Val Score do teste no modelo KNN sem balanceamento do dataset.

Cross Val Score (cv = 5)					Média
0.990	0.992	0.990	0.987	0.99	0.990

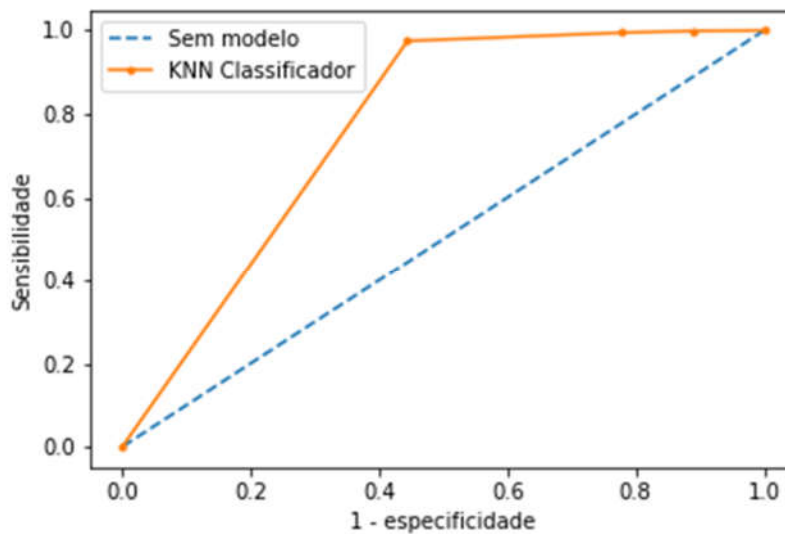


Figura 12: Curva ROC para o teste com o KNN sem balanceamento. Fonte: Autor.

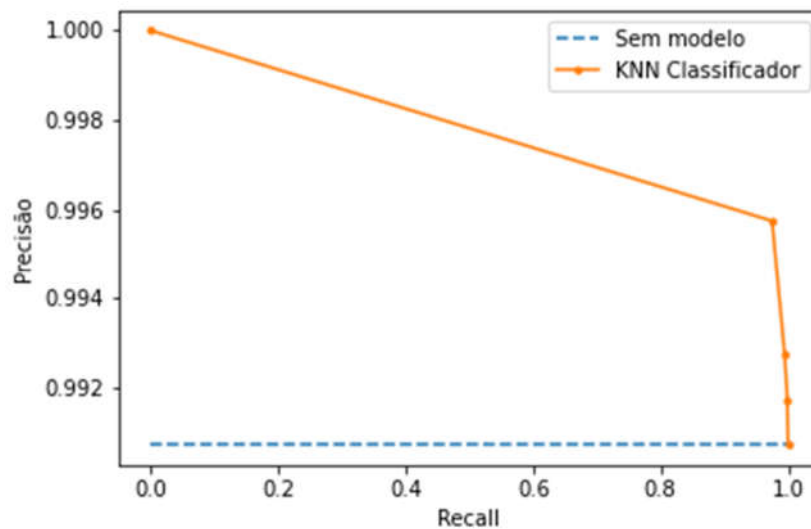


Figura 13: Curva PR para o teste com o KNN sem balanceamento. Fonte: Autor.

- **Deep Learning**

Para o treinamento com aprendizado profundo foi utilizada a biblioteca Keras com o pacote contido nela do scikit Learn com o modelo KerasClassifier. Foi definida uma função que carrega o modelo *Sequential* com adição de camadas com o Dense. Foram adicionadas três camadas: duas com 100 neurônios e função de ativação ReLU e uma com 2 neurônios e função de ativação *Softmax* que configura a camada de saída. Os resultados constam na Tabela 6 as curvas ROC e PRC respectivamente na Figura 14 e na Figura 15.

Tabela 6: Valores de Cross Val Score para o teste no modelo de Aprendizado Profundo.

Cross Val Score (cv = 5)					Média
0.75	0.666	0.583	0.916	0.5	0.68

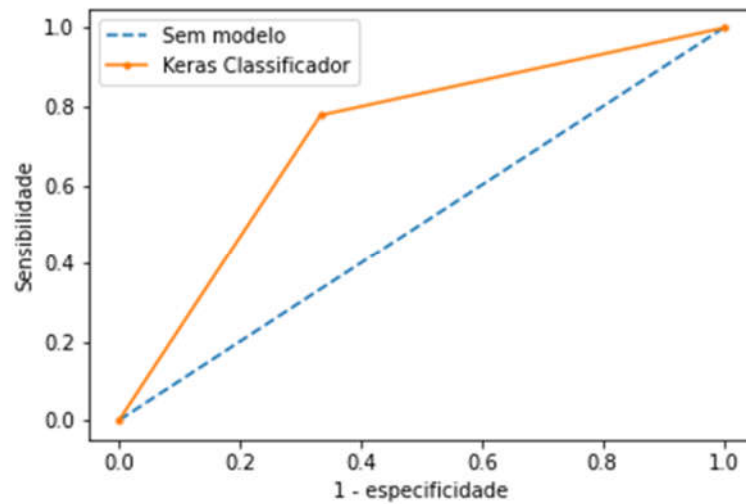


Figura 14: Curva ROC para o teste com Deep Learning com balanceamento. Fonte: Autor.

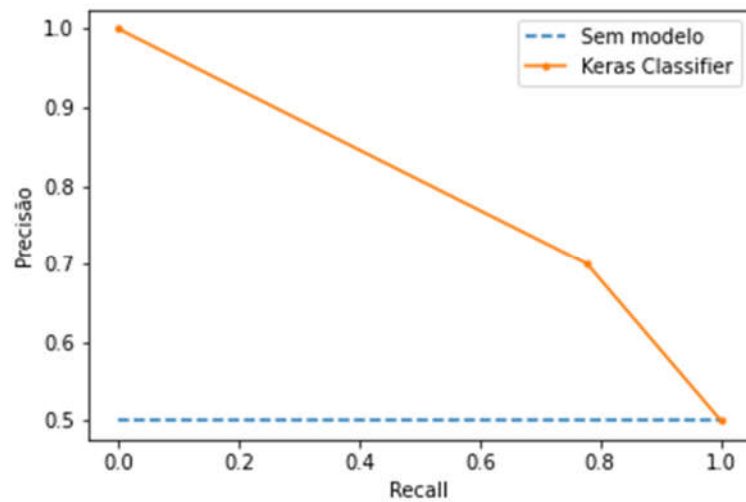


Figura 15: Curva PR para o teste com Deep Learning com balanceamento. Fonte: Autor.

Sem a aplicação de undersampling os resultados seguem a Tabela 7 com os gráficos ROC na Figura 16 e PRC na Figura 17.

Tabela 7: Valores de Cross Val Score para o teste no modelo de Aprendizado Profundo sem balanceamento do dataset.

Cross Val Score (cv = 5)					Média
0.993	0.993	0.992	0.990	0.98	0.991

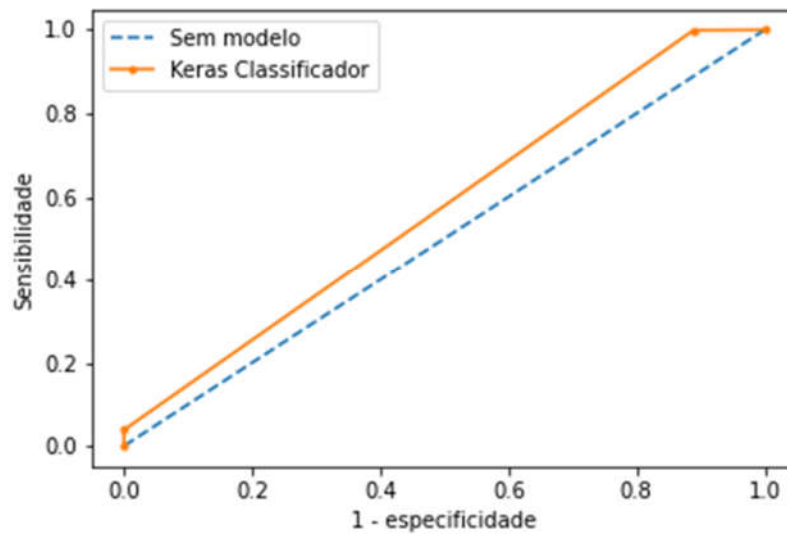


Figura 16: Curva ROC para o teste com Deep Learning sem balanceamento. Fonte: Autor.

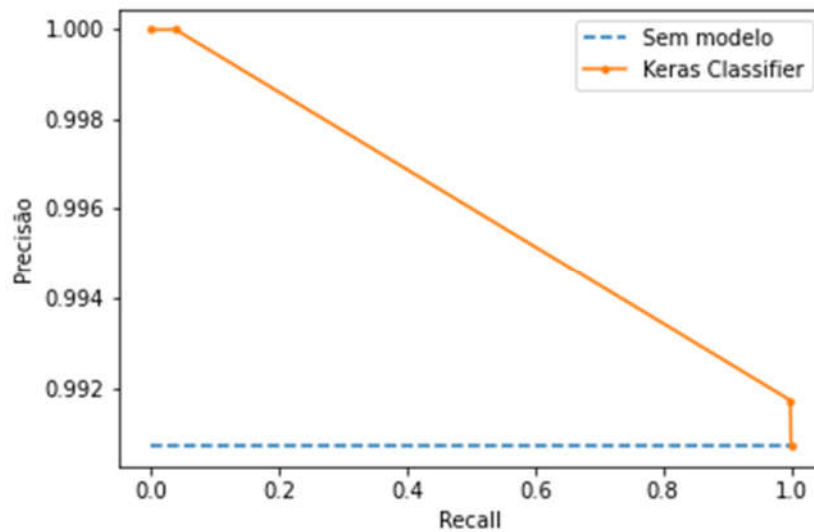


Figura 17: Curva PR para o teste com Deep Learning sem balanceamento. Fonte: Autor.

6. CONCLUSÃO

Com os resultados pode-se perceber uma melhora nos valores das médias dos *Cross Val Score* dos testes sem o procedimento de balanceamento do *dataset* em comparação dos testes onde foi feito o *undersampling* nos dados. Isso ocorre, pois, a métrica avaliada no *Cross Val Score* leva em consideração a taxa de acerto da rede que, com o desbalanceamento, torna os acertos para a classe majoritária tão altos que sobrepõem os acertos da classe minoritária. Para isso há a análise das curvas ROC E PRC, que ao considerar a quantidade de acertos da classe minoritária mostram que o desempenho dos testes não foi tão bom.

Com isso, nos próximos trabalhos busca-se melhorar tais métricas com mais testes e poder obter melhores comparações de modelos para que, em projetos futuros, por fim, aplicar tais

metodologias em *dataset* da própria universidade e com isso, obter conhecimento sobre o consumo de energia para poder inferir resultados que poupem energia e consequente gastos na universidade.

7. PUBLICAÇÕES TÉCNICO-CIENTÍFICAS. *Listar os bolsistas e nome do evento que participou; listar artigos publicados em periódicos, comunicações em congresso, capítulos em livros, livros, manuais, etc. Fornecer cópia, anexo, de cada item citado.*

Como autora principal e com a coautoria de João Vitor Rebelo Viana, Josecley Fialho Goes e Anderson Alvarenga de Moura Meneses submeti o trabalho intitulado “*Aplicação de Técnicas de Data Science e Machine Learning em um Problema de Classificação de um Dataset de Marketing Bancário*” no XIV Simpósio Brasileiro de Engenharia Física. Esse trabalho recebeu a premiação de melhor trabalho do evento.

Particpei do projeto Sistema de Monitoramento de Consumo de Energia Utilizando Internet das Coisas, Big Data e Machine Learning financiado pelo Banco da Amazônia (BASA).

Particpei na organização do evento Curso de Férias Ciência de Dados e nesse mesmo evento ministrei o minicurso de Introdução a Internet das Coisas juntamente com Dalton Felipe Silva Varão.

Com Clayton dos Santos, Dalton Varão, Davi Guimarães, João Bentes, Marla Geller, Mauro Moura e Anderson Meneses participei do desenvolvimento do Manual de Utilização e o Tutorial de Instalação do Software EnergySaver, sistema de coleta de dados e monitoramento de energia na Universidade Federal do Oeste do Pará.

8. PRINCIPAIS PROBLEMAS E DIFICULDADES PARA A REALIZAÇÃO DAS ATIVIDADES

A principal dificuldade técnica na realização da pesquisa foi a partir da escolha do *dataset* que possuía uma característica de muitos dados faltantes ou dados desbalanceados, o que motivou o redirecionamento dos estudos para técnicas de balanceamento ou de diferentes avaliações ao invés de um estudo profundo nas configurações do modelo e seu melhoramento.

9. REFERÊNCIAS

AGGARWAL, C. C. **Neural Networks and Deep Learning**. Yorktown Heights, New York: Springer International Publishing, 2018.

CIELEN, D.; MEYSMAN, A. D. B.; ALI, M. **Introducing Data Science**. Shelter Island: Manning Publications Co. , 2016.

DONG, Q.; XING, K.; ZHANG, H. **Artificial Neural Network for Assessment of Energy Consumption and Cost for Cross Laminated Timber Office Building in Severe Cold Regions.** Sustainability, 30 Dezembro 2017.

ELSALAMONY, H. A.; ELSAYAD, A. M. **Bank Direct Marketing Based on Neural Network.** International Journal of Engineering and Advanced Technology (IJEAT), Agosto 2013. 392-400.

ERTEL, W. **Introduction to Artificial Intelligence.** Weingarten: Springer International Publishing, 2017.

FAYAZ, M.; KIM, D. **A Prediction Methodology of Energy Consumption Based on Deep Extreme Learning Machine and Comparative Analysis in Residential Buildings.** Electronics Journal, 28 Setembro 2018.

KOSELEVA, N.; ROPAITE, G. **Big data in building energy efficiency: understanding of big data and main challenges.** Elsevier, p. 544-549, 2017.

KUBAT, M. **An Introduction to Machine Learning.** Coral Gables: Springer International Publishing, 2017.

LAST, F.; DOUZAS, G.; BACAO, F. **Oversampling for Imbalanced Learning Based on K-Means and SMOTE.** Information Sciences, 12 Dezembro 2017. 1-20.

LIU, X.-Y.; WU, J.; ZHOU, Z.-H. **Exploratory Undersampling for Class-Imbalance Learning.** IEEE Transactions on Systems, Man, and Cybernetics, 2 Abril 2009. 539-550.

MORÁN, A. et al. **Analysis of electricity consumption profiles in public buildings with dimensionality reduction techniques.** Engineering Applications of Artificial Intelligence, 15 Maio 2013. 1872-1880.

MURPHEY, Y. L.; GUO, H.; FELDKAMP, L. A. **Neural Learning from Unbalanced Data.** Applied Intelligence, Setembro 2004. 117-128.

NIKOLAOU, T.; KOLOKOTSA, D.; STAVRAKAKIS, G. **Review on methodologies for energy benchmarking, rating and classification of buildings.** Advances in Building Energy Research, 29 junho 2011. 53-70.

PATTERSON, J.; GIBSON, A. **Deep Learning.** Sebastopol: O'Reilly Media, Inc, 2017.

ROBINSON, C. et al. **Machine learning approaches for estimating commercial building energy consumption.** Applied Energy, 9 Outubro 2017. 889-904.

SAYADI, H. et al. **Machine Learning-Based Approaches for Energy-Efficiency Prediction and Scheduling in Composite Cores Architectures.** IEEE 35th International Conference on Computer Design, 23 Novembro 2017. 129-136.

SKANSI, S. **Introduction to Deep Learning.** Zagreb: Springer International Publishing, 2018.


10. ANEXOS

- Trabalho publicado no SBEF 2019.

11. PARECER DO ORIENTADOR

A aluna Yasmin Braga Teixeira desenvolveu suas tarefas e trabalhos de maneira satisfatória, tendo recebido inclusive o prêmio de melhor trabalho no SBEF 2019. A aluna demonstrou ser proativa e dedicada durante o período da bolsa.

Santarém, PA, 15 de setembro de 2020.



Prof. Anderson Alvarenga de Moura Meneses, D.Sc.
Orientador



Yasmin Braga Teixeira
Bolsista