



ANÁLISE DO CONSUMO DE ENERGIA ELÉTRICA RESIDENCIAL UTILIZANDO MACHINE LEARNING

ARTIGO ORIGINAL

MORAES, Valescla Aragão de ^[1], HILASACA, Liz Maribel Huancapaza ^[2], BELIZARIO, Ivar Vargas ^[3], LESSA, Carmen Lúcia Avelar ^[4]

MORAES, Valescla Aragão de. *et al.* **Análise do consumo de energia elétrica residencial utilizando machine learning.** Revista Científica Multidisciplinar Núcleo do Conhecimento. Ano 08, Ed. 12, Vol. 02, pp. 113-132. Dezembro de 2023. ISSN: 2448-0959, Link de acesso: <https://www.nucleodoconhecimento.com.br/engenharia-eletrica/consumo-de-energia-eletrica>, DOI: 10.32749/nucleodoconhecimento.com.br/engenharia-eletrica/consumo-de-energia-eletrica

RESUMO

O consumo de energia elétrica é uma preocupação global devido ao seu uso de forma ineficiente e descontrolado, resultando em gastos excessivos. Nesse contexto, o presente artigo propõe a aplicação de técnicas de inteligência artificial, em particular o *machine learning*, para analisar o consumo de energia residencial. Foram utilizados dados obtidos por meio de sensores que passaram por um pré-processamento para eliminar instâncias com dados incompletos. Uma categorização do consumo de energia é apresentada, permitindo que a inteligência artificial implementada realize a classificação correspondente. Também foi realizada uma análise de técnicas para dados desbalanceados referente ao consumo de energia, assim, de acordo com a quantidade de experimentos executados, são fornecidas recomendações para balanceamento com a menor perda de informação possível. A validação da metodologia foi feita comparando os rótulos verdadeiros com as categorias preditas. Os resultados demonstraram a eficácia da proposta na classificação, alcançando 98% de acurácia. Além disso, as visualizações por meio de UMAP, t-SNE e Coordenadas Paralelas destacaram a separação das instâncias de distintas categorias, indicando a eficácia da metodologia.



Palavras-chave: Consumo de energia elétrica, Consumo de energia, Classificação, Aprendizado de Máquina.

1. INTRODUÇÃO

O aprendizado de máquina (*Machine Learning*, ML) é um campo da inteligência artificial baseado em teoria da otimização e informação, estatística, ciência cognitiva e muitas outras disciplinas da ciência, engenharia e matemática (Mitchell, 1997). O ML é implementado para uma ampla gama de aplicações, o que trouxe grande impacto na ciência e na sociedade (Kim *et al.*, 2022).

A aplicação do ML vem crescendo em diversas áreas (Chagas, 2019). O estudo de Jin *et al.* (2014) utiliza algoritmos de agendamento e escalonamento para prever e ajustar de forma automatizada o consumo de eletrodomésticos. Murata e Onoda (2002) realizaram um trabalho com o objetivo de calcular uma previsão de consumo de aparelhos elétricos. Em seu trabalho, destacam que o algoritmo RBFN (*Radial Basis Function Neural Network*) apresentou um melhor resultado, com maior precisão, principalmente para altos consumos, quando comparado com o MLP (*Multilayer Perceptron*) e o SVR (*Support Vector Regression*).

O objetivo deste trabalho é utilizar o ML para classificar o Consumo de Energia Elétrica (CEE) residencial e, assim, reconhecer os padrões do consumo de uma residência. Este trabalho contribui de forma significativa, porque gera dados como resultado da execução de sua implementação. Também apresenta uma comparação entre diversas técnicas para o tratamento de dados, expondo as mais eficazes para a demonstração dos padrões de CEE através da utilização de balanceamento dos resultados obtidos.

A categorização do CEE foi feita através de níveis definidos como muito baixo, baixo, médio, alto e muito alto. Dessa forma, foi fornecido um novo conjunto de dados categorizado que poderá ser utilizado em trabalhos futuros na área de ML.



Com o objetivo de facilitar a tomada de decisões com relação ao consumo de energia em residências, foi utilizada a técnica de visualização, que permite a verificação do resultado de forma mais clara e rápida. Os consumidores, as empresas do setor de energia e o governo são beneficiados com essas informações, já que podem orientar estratégias e políticas eficazes relacionadas ao consumo de energia.

O conjunto de dados utilizado neste artigo foi o *Individual Household Electric Power Consumption*, desenvolvido por Hebrail e Berard (2012) e fornecido pela Universidade de Califórnia Irvine (UCI) por meio do *Machine Learning Repository*. Trata-se de dados referentes ao consumo de energia elétrica residencial. Esses dados incluem informações como medições de tensão, potências, corrente elétrica e submedições do consumo em determinados pontos da residência.

Para a realização dos experimentos, foi desenvolvida a implementação em Python através de *Kaggle*. O desenvolvimento da pesquisa foi dividido em duas etapas: primeiramente, foi feita a coleta e categorização dos dados; em sequência, foi desenvolvido o algoritmo para processar os dados e, então, obter os resultados necessários relacionados ao consumo de energia em residências.

2. DESENVOLVIMENTO

2.1 CONJUNTO DE DADOS

O conjunto de dados atual abrange 2.075.259 medições coletadas em uma residência e está organizado da seguinte forma:

Data: No formato de dia/mês/ano.

Hora: No formato hora:minuto:segundo.



global_active_power: Representa a potência ativa média global por minuto na residência, expressa em quilowatts.

global_reactive_power: É referente à potência reativa média global na residência por minuto, dada em quilowatts.

Tensão: Indica a tensão média por minuto na residência, expressa em volts.

global_intensity: Corresponde à intensidade de corrente média global por minuto na residência, medida em ampères.

sub_metering_1: Corresponde à submedição de energia número 1, relacionada à cozinha.

sub_metering_2: Refere-se à submedição de energia número 2, que envolve a lavanderia.

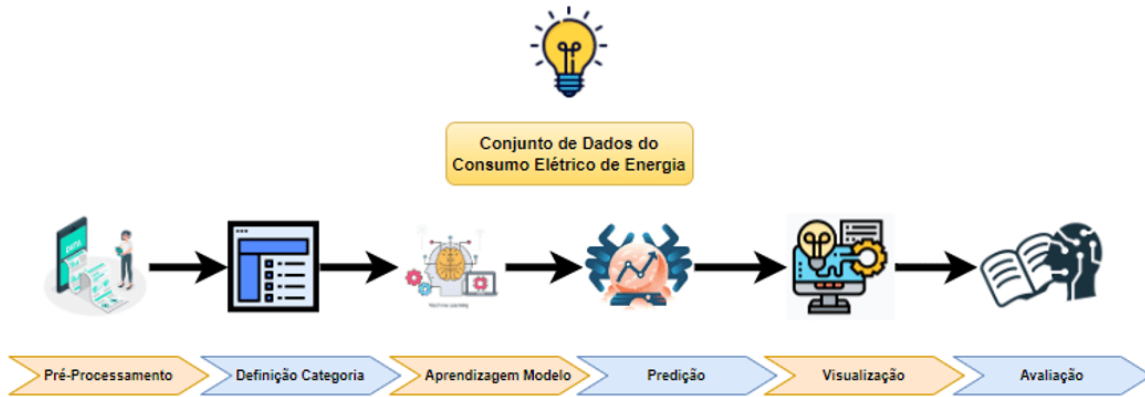
sub_metering_3: Referente à submedição de energia número 3, abrangendo dispositivos que consomem energia ativa.

Nota-se que o dado "*global_active_power**1000/60" determina a energia ativa consumida por equipamentos elétricos não reverenciados pelas submedições 1, 2 e 3. Esse valor é expresso em watt-hora por minuto.

Essas informações estabelecem a base necessária do conjunto de dados, permitindo uma análise ampla das medições de energia da residência em questão (Hebrail;Berard, 2012).

Para a análise do conjunto de dados a fim de classificar o CEE, foi seguida uma sequência de atividades como mostra a Figura 1. As mesmas foram divididas em seis etapas: pré-processamento; definição das categorias; aprendizagem do modelo; predição; visualizações e avaliação.

Figura 1 – Pipeline: resumo dos processos para a classificação de consumo elétrico residencial



Fonte: Os autores, 2023.

2.2 PRÉ-PROCESSAMENTO

Os dados foram obtidos por meio de sensores, por isso, o primeiro passo foi a eliminação de dados incompletos. A estratégia para extinguir as instâncias foi a seguinte: se qualquer registro possui pelo menos uma característica incompleta, a instância é anulada. Dessa forma, aproximadamente 1% do total de dados foi eliminado.

2.3 DEFINIÇÃO DE CATEGORIA

O cálculo para definir as classificações do CEE foi fundamentado no artigo de Fedrigo *et al.* (2009), em que foi feita uma análise de dados do CEE mensal de acordo com as estações do ano. O menor consumo foi de 81kW/mês, o médio 152,2kW/mês e o alto 310,6kW/mês. Com isso, foi possível obter informações para fazer o cálculo do projeto atual. Considerando que um mês com trinta dias equivale a 43200 minutos, os valores do consumo mensal foram divididos por minutos. O dado utilizado para a categorização baseou-se na coluna "active power" do conjunto de dados, que é representada pela potência ativa.



Menor consumo:

$$\frac{81000}{43200} = 1.87kW/min$$

Consumo médio:

$$\frac{152200}{43200} = 3.52kW/min$$

Maior consumo:

$$\frac{310600}{43200} = 7.18kW/min$$

Para definir as categorias, foi feita a diferença entre o maior e o menor consumo:

$$7.18 - 1.87 = 5.31kW/min$$

Foi possível obter as classificações dos consumos baixo, médio e alto dividindo por três:

$$\frac{5.31}{3} = 1.77kW/min$$

Depois, foi feita a soma dos valores do menor, médio e maior consumo:

$$1.77 + 1.87 = 3.64kW/min$$

$$1.77 + 3.64 = 5.41kW/min$$

$$1.77 + 5.41 = 7.18kW/min$$

Com base nesses resultados, as classificações foram definidas em cinco parâmetros:



- Muito baixo < 1.87 (kW/min)
- Baixo ≥ 1.87 e < 3.64 (kW/min)
- Médio ≥ 3.64 e < 5.41 (kW/min)
- Alto ≥ 5.41 e < 7.18 (kW/min)
- Muito alto ≥ 7.18 (kW/min)

2.3.1 BALANCEAMENTO

Foi verificado o número de instâncias nas classes que estavam desbalanceadas. Dessa forma, foi possível notar que algumas classes tinham um maior número de instâncias e outras apresentavam um menor número de instâncias. Portanto, foram realizadas tarefas para balancear as classes minoritárias para aproximar-se ao número de instâncias com a classe majoritária. Esse tipo de balanceamento é denominado como *Oversampling*, sendo utilizada a técnica de SMOTE (*Synthetic Minority Over-sampling Technique*) (Imbalanced-Learn, 2023), desenvolvida por Chawla *et al.* (2002), que realiza o aumento da amostragem para a classe positiva, gerando novas instâncias “sintéticas” em contraste à abordagem de seleção de amostras.

2.4. APRENDIZAGEM DE MODELO

Após gerar o conjunto de dados balanceado, foram construídos subconjuntos para executar tarefas de treinamento e avaliação do modelo. Esses subconjuntos gerados foram treino, validação e teste. A porcentagem foi de 80% para o subconjunto treino, 10% para validação e 10% para teste. A geração dos mesmos foi feita de forma que as instâncias para cada classe estivessem divididas de maneira proporcional.



Utilizando o subconjunto treino, foi treinado um modelo de aprendizagem supervisionado baseado em árvores. Assim, foi utilizado o modelo *random forest*, que pode ser empregado a uma variável de resposta categórica, ou contínua, que se refere como “regressão”. Esse modelo foi usado com a seguinte estrutura: $max_depth = 4$ e $n_estimators = 200$, onde o primeiro parâmetro assegura o máximo dos níveis por árvores e o segundo parâmetro garante o número total de árvores criadas (Cutler; Cutler; Stevens, 2012).

2.5 PREDIÇÃO

Para computar as etiquetas por predição, foi preciso utilizar os *features* ou características dos conjuntos de validação e teste. Essa tarefa foi feita para examinar o grau de aprendizagem que o modelo treinado teve. As especificações da validação são descritas na seção de avaliação.

2.6 VISUALIZAÇÕES

O UMAP (*Uniform Manifold Approximation and Projection*) e o t-SNE (*t-Stochastic Distributed Neighbor Embedding*) foram as técnicas utilizadas para reduzir a dimensionalidade dos gráficos.

O t-SNE utiliza a distribuição *t-student* para modelar a probabilidade que usa uma abordagem iterativa para melhorar a projeção dos dados, além de buscar similaridade entre os pares de pontos nos dados originais (Van Der Maaten; Hinton, 2008).

O UMAP foi desenvolvido como uma alternativa ao t-SNE, outro algoritmo popular de redução de proporção de gráficos. O artigo de McInnes e Healy (2018) menciona que o t-SNE pode ser caro computacionalmente e tende a agrupar amostras de maneira densa, dificultando a interpretação visual dos dados. O UMAP foi projetado para abordar essas limitações e fornecer uma técnica mais eficiente e flexível para



visualização. Ele oferece a capacidade de usar rótulos (ou rótulos parciais) para redução de dimensão supervisionada (ou semisupervisionada), além da capacidade de transformar novos dados invisíveis em um espaço de incorporação pré-treinado (McInnes; Healy, 2018).

Nessa perspectiva, foi gerado um gráfico de coordenadas paralelas para realizar a comparação entre os valores de dados, permitindo comparar dados de tipos de magnitudes diferentes em única visualização (Tibco Software, 2023).

2.7 AVALIAÇÃO

Foi necessário empregar um conjunto de dados que possuam rótulos verdadeiros e categorizados para validar a metodologia. É necessário ressaltar que a validação, nesta etapa, consiste na comparação dos rótulos e categorias verdadeiras com os determinados previamente na etapa de predição.

A validação dos dados foi realizada por meio de métricas de avaliação, como a acurácia, sensibilidade, precisão e *f1-score*.

A base para o cálculo e interpretação dessas métricas é a matriz de confusão, uma vez que trata-se de uma representação que avalia a qualidade de uma classificação digital de imagem, de forma a revelar a correlação entre os dados classificados e os dados de referência, considerados como verdadeiros (Prina; Tretin, 2015).

A acurácia reflete o desempenho geral de um modelo, indicando quantas classificações estão corretas em relação ao número total (Mariano, 2021).

$$\text{acurácia} = \frac{VP+VN}{VP+FN+VN+FP} \quad (1)$$

Considerando que VP é o valor positivo, VN é o valor negativo, FN é o valor falso negativo e FP é o valor falso positivo.



A precisão realiza a avaliação de previsões realizadas para a classe positiva e define as previsões corretas.

$$\text{precisão} = \frac{VP}{VP+FP} \quad (2)$$

A sensibilidade calcula os casos verdadeiros positivos em relação a todas as situações esperadas de classe positiva.

$$\text{sensibilidade} = \frac{VP}{VP+FN} \quad (3)$$

O F1-Score apresenta uma visão consolidada da performance, porque realiza a junção entre a precisão e a sensibilidade em uma única métrica. Apresenta sensibilidade a pontuações baixas em qualquer medida devido à sua característica de média harmônica, alertando possíveis situações em que a precisão ou a sensibilidade apresentam níveis inferiores.

$$F1 \text{ score} = 2 \times \frac{\text{sensibilidade} \times \text{precisão}}{\text{sensibilidade} + \text{precisão}} \quad (4)$$

Em resumo, a matriz de confusão e suas métricas derivadas são ferramentas cruciais para avaliar a qualidade e o desempenho de modelos de classificação, proporcionando uma visão abrangente e detalhada das previsões em relação aos resultados reais.

Foi empregado outro método de validação, denominado *rand_score*, que se baseia na avaliação da similaridade entre dois agrupamentos. Conforme documentado no site da *Scikit-Learn* (2023), o *rand_score* compara todos os pares de amostras e contabiliza aqueles que são atribuídos aos mesmos ou diferentes agrupamentos, tanto nos resultados previstos quanto nos verdadeiros. Essa métrica é calculada pela razão entre o número de pares de amostras que concordam e o número total



de pares avaliados, fornecendo uma medida de validação da qualidade da classificação obtida.

3. RESULTADOS

Empregando os critérios da metodologia descritos na seção 2.2, o conjunto foi subdividido em cinco classes, conforme ilustra a Tabela 1.

Tabela 1 - Número de instâncias do conjunto etiquetado para 5 classes

Muito Baixo	Baixo	Médio	Alto	Muito Alto
1704102	272940	60558	10375	1305

Fonte: Os autores, 2023.

É definido o número de instâncias por classe que será empregado para criar o subconjunto teste. Assim, foram definidas nove variações em porcentagens, considerando o número de instâncias da classe minoritária (Muito Alto = 1305). Os resultados são ilustrados na Tabela 2.

Tabela 2 - Tamanho de instâncias por classe para o subconjunto teste com base na classe minoritária

10%	20%	30%	40%	50%	60%	70%	80%	90%
130	261	391	522	652	783	913	1044	1174

Fonte: Os autores, 2023.

No balanceamento do conjunto de dados, primeiro, foi utilizada a técnica de *random undersampling* para reduzir o número de instâncias das duas classes majoritárias (Muito Baixo e Baixo). Foi considerada a aproximação da média (cem mil) do número de instâncias das duas classes majoritárias. A Tabela 3 ilustra os resultados de redução para essas classes.

Tabela 3 - Balanceamento das duas classes majoritárias



Muito Baixo	Baixo	Médio	Alto	Muito Alto
100000	100000	60558	10375	1305

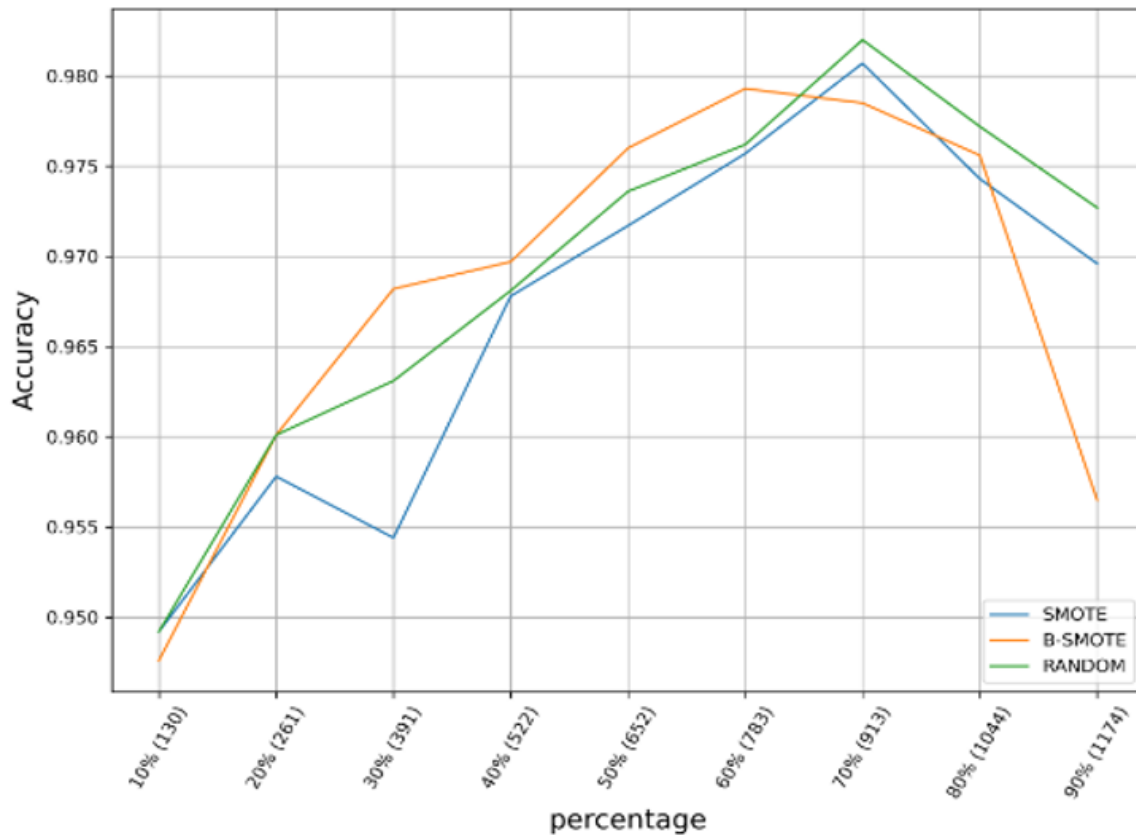
Fonte: Os autores, 2023.

Para balancear as classes minoritárias (Médio, Alto, e Muito Alto) até atingir cem mil instâncias por classe, foram avaliados três métodos de *oversampling* (SMOTE, *Borderline-SMOTE* e *RANDOM*). A ideia do experimento foi avaliar qual método de balanceamento é o mais adequado para executar as tarefas de classificação.

Dada a configuração da Tabela 2, a metodologia é executada nove vezes. Em cada execução do experimento, foram avaliados os três métodos de *oversampling* para balancear as classes minoritárias. Após o balanceamento, os subconjuntos de treino e validação são criados, de forma que o subconjunto de validação tem o mesmo número de instâncias por classe que o subconjunto teste. O restante de instância é atribuído ao conjunto de treino. Depois, o modelo de RF (*Random Forest*) é treinado com os conjuntos de treino e validação, em que o subconjunto validação é utilizado especificamente para determinar os parâmetros ótimos para o modelo.

Os melhores resultados para a criação do conjunto de teste, de acordo com a Figura 3, deram-se com a configuração de 70%. Por outro lado, vale ressaltar que o método *Borderline-SMOTE* é melhor até a configuração de 60%. Pode-se verificar também que os resultados para 80% e 90% têm tendência a diminuir. Isso pode ser resultante do subconjunto de treinamento ser balanceado com uma quantidade muito menor de instâncias, o qual é insuficiente para treinar o modelo.

Figura 3 – Resultados de classificação para as três técnicas de *oversampling*. No eixo Y, resultados do *score* de classificação. No eixo X, as variações do número de instâncias por classe para criar o subconjunto de teste



Fonte: Os autores, 2023.

Para uma mais adequada visualização dos resultados, a Tabela 4 e a Figura 4 ilustram os melhores resultados da configuração de 70%, onde o resultado superior foi obtido com o método de balanceamento *Random Oversampling*. O segundo melhor método foi obtido com o método SMOTE. Em termos da métrica *rand_score*, pode-se observar valores altos dessa métrica para os métodos *Random* e SMOTE *Oversampling*, isso poderia indicar que o agrupamento em cinco *clusters* dos dados balanceados tem similaridade adequada em relação aos dados originalmente etiquetados.

Tabela 4 – Melhores resultados de classificação obtidos com a técnica de *oversampling* empregando o 70% de instâncias por classe para o subconjunto teste considerando a classe minoritária

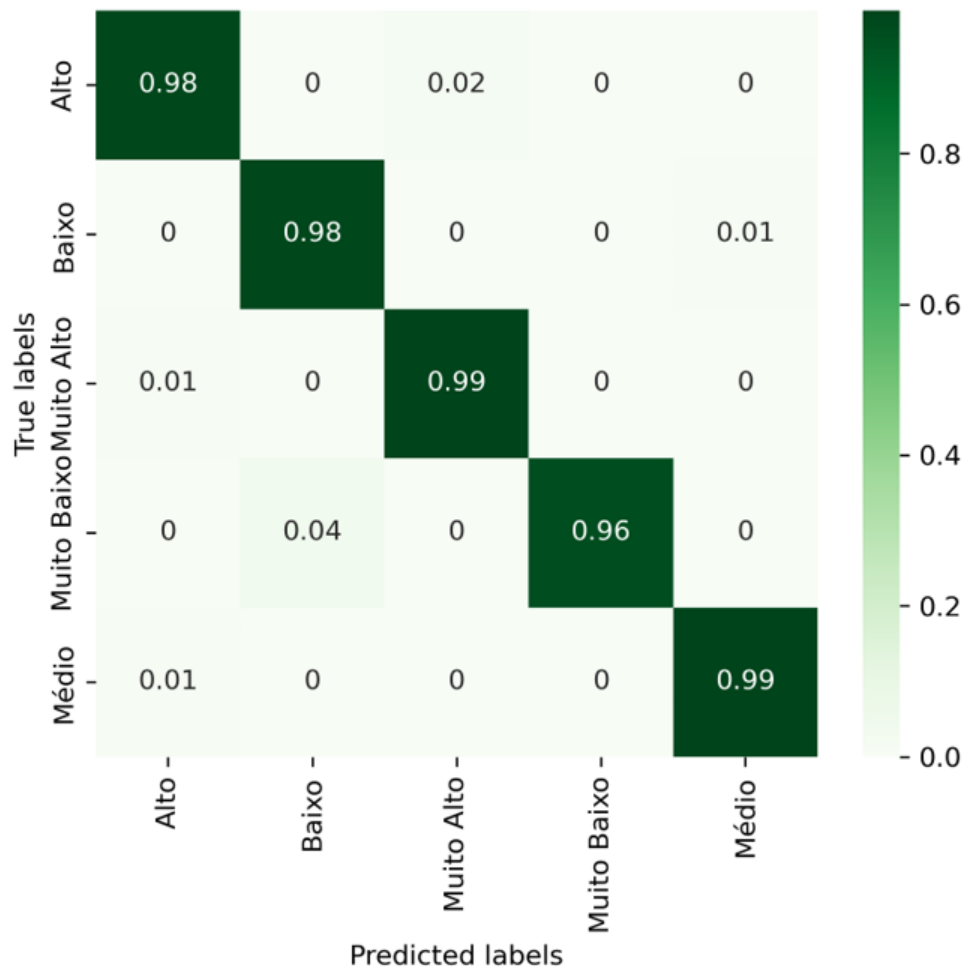
Oversampling	Rand Score	Acc Validação	Acc Teste
---------------------	-------------------	---------------	-----------



Random	0.7342	0.9776	0.9820
SMOTE	0.7349	0.9796	0.9807
B-SMOTE	0.7298	0.9789	0.9785

Fonte: Os autores, 2023.

Figura 4 – Resultados da matriz de confusão para a configuração do subconjunto de teste de 70% utilizando o método *random oversampling*



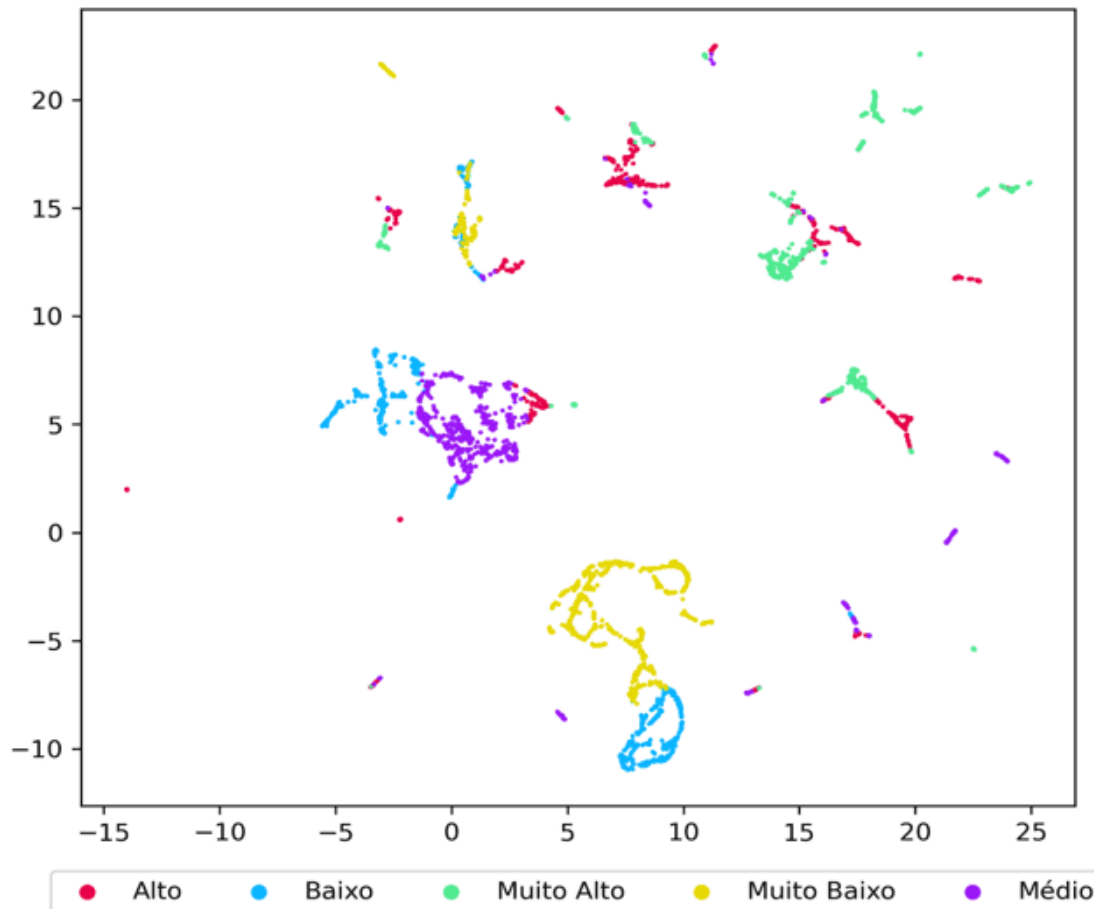
Fonte: Os autores, 2023.

As Figuras 5 e 6 ilustram as representações do subconjunto de teste com a configuração de 70%. As técnicas utilizadas foram as projeções UMAP e T-SNE, respectivamente. Nessas visualizações, é possível observar os exemplos das cinco classes, representados por cores distintas. Especificamente, a técnica T-SNE

126

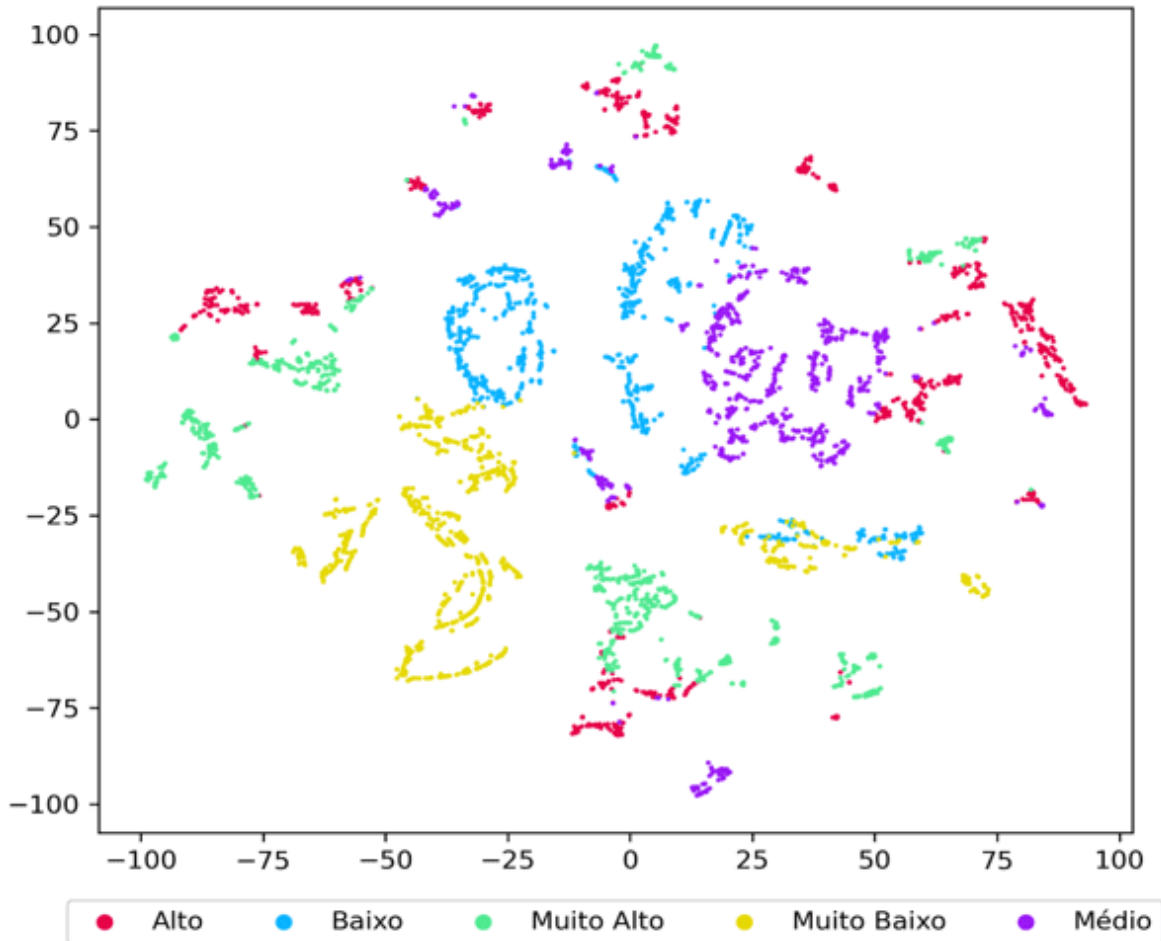
proporciona os melhores resultados em termos de exibição das amostras, permitindo apreciar as instâncias projetadas de forma mais dispersa, ao mesmo tempo em que mantém pequenos agrupamentos.

Figura 5 - Visualização do subconjunto teste utilizando o método UMAP. As cores representam as classes dos tipos de consumo



Fonte: Os autores, 2023.

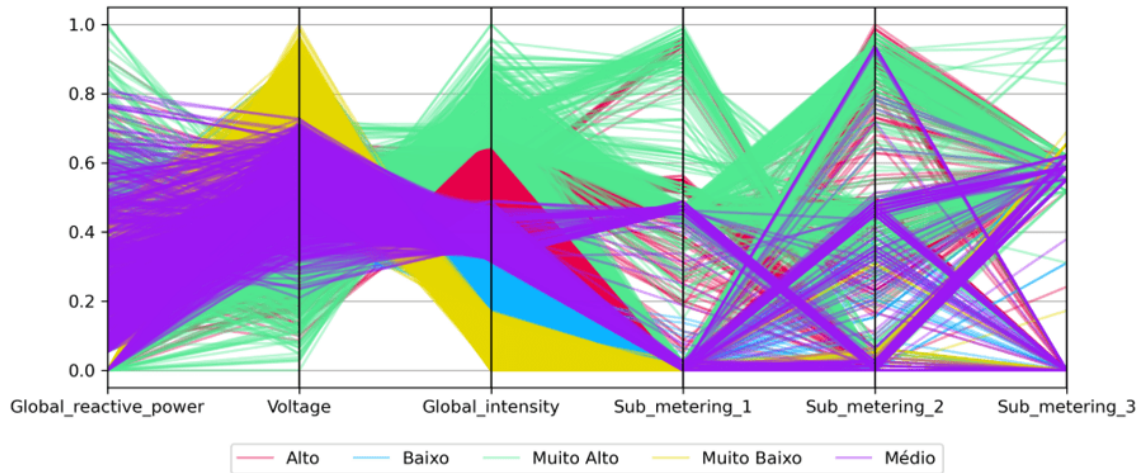
Figura 6 - Visualização do subconjunto teste utilizando o método t-SNE. As cores representam as classes dos tipos de CEE



Fonte: Os autores, 2023.

A Figura 7 mostra a visualização utilizando a técnica de coordenadas paralelas. Nela, cada linha representa uma instância, cada coluna representa um atributo e as cores representam os tipos de CEE. A visualização evidencia quais atributos ou características são as mais discriminatórias para a ótima classificação como demonstra a tabela de resultados (ver Tabela 4). É nítido observar que o atributo *Global Intensity* nas coordenadas paralelas foi o mais destacado para classificar os tipos de consumo CEE.

Figura 7 – Visualização por coordenadas paralelas do subconjunto teste. Cada coluna representa as características utilizadas na execução dos experimentos. As cores representam as classes dos tipos de CEE



Fonte: Os autores, 2023.

4. CONCLUSÃO

A análise realizada neste estudo concentra-se na classificação de diferentes níveis de CEE, utilizando técnicas de balanceamento de conjuntos de dados, especificamente o *oversampling*. O objetivo foi determinar qual método proporciona os melhores resultados de classificação, mas, ao mesmo tempo, com menor perda de informação. Os resultados apresentados indicam que, ao empregar 70% (com base na classe minoritária) das instâncias por classe para o subconjunto de teste, o método de balanceamento obteve o melhor desempenho em termos de métricas como *rand score* e acurácia. As visualizações utilizando os métodos UMAP, t-SNE e Coordenadas Paralelas ajudaram a mostrar como as instâncias das diferentes classes estão sendo agrupadas após o balanceamento e a classificação, além de destacar o atributo *Global Intensity* como o mais discriminatório dos tipos de consumo de energia elétrica.

Finalmente, os resultados e análises indicam que o uso do método de *random oversampling* leva a uma sólida classificação das diversas classes de CEE. Isso reforça a importância do balanceamento de conjuntos de dados e das técnicas de



oversampling para aprimorar o desempenho de modelos de classificação em cenários desbalanceados.

Este estudo fornece uma base consistente para futuras pesquisas e aplicações práticas. À medida que a demanda por análises de dados precisas continua a crescer, a compreensão desses métodos torna-se cada vez mais crucial para a tomada de decisões informadas e o avanço do conhecimento em várias áreas.

REFERÊNCIAS

CHAGAS, E. T. O. Deep Learning e suas aplicações na atualidade. **Revista Científica Multidisciplinar Núcleo do Conhecimento**, São Paulo, ano 04, ed. 05, v. 04, p. 05-26, mai. 2019. Disponível em: <https://www.nucleodoconhecimento.com.br/administracao/deep-learning>. Acesso em: 17 jul. 2023.

CHAWLA, N. V., *et al.* SMOTE: Synthetic minority over-sampling technique. **J. Artif. Intell. Res. (JAIR)**, n. 16, p. 321–357, 2002.

CUTLER, A.; CUTLER, D; STEVENS, J. R. Florestas Aleatórias. In: ZHANG, C., MA., Y. (eds) **Ensemble Machine Learning: methods and applications**. Springer. Nova York, NY, 2012. Disponível em: https://doi.org/10.1007/978-1-4419-9326-7_5. Acesso em: 06 nov. 2023.

FEDRIGO, N. S. *et al.* **Usos Finais de Energia Elétrica no Setor Residencial Brasileiro**. LabEEE, 2009. Disponível em: https://labeee.ufsc.br/sites/default/files/publicacoes/relatorios_ic/IC2009_Natalia.pdf. Acesso em: 25 mar. 2023.

HEBRIL, G.; BERARD, A. **Individual household electric power consumption**. UCI Machine Learning Repository, 2012. Disponível em: <https://doi.org/10.24432/C58K54>. Acesso em: 31 mar. 2023.

IMBALANCED-LEARN. **SMOTE Variants**. Imbalanced Learn, 2023. Disponível em: https://imbalanced-learn.org/stable/over_sampling.html#smote-variants. Acesso em: 27 mai. 2023.

JIN, W. *et al.* **A Household Electricity Consumption Algorithm with Upper Limit**. In: 2014 International Conference on Wireless Communication and Sensor Network.



IEEE: Wuhan, China, 2014. Disponível em: <https://ieeexplore.ieee.org/document/7061769>. Acesso em: 15 abr. 2023

KIM, I. *et al.* **Machine Learning para otimização de sistemas de energia**. MDPI, 2022. Disponível em: <https://www.mdpi.com/1996-1073/15/11/4116>. Acesso em: 25 mar. 2023.

MARIANO, D. **Métricas de avaliação em Machine Learning**. Diegomariano, 2021. Disponível em: <https://diegomariano.com/metricas-de-avaliacao-em-machine-learning/>. Acesso em: 22 ago. 2023.

MCINNES, L.; HEALY, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. **Journal of Open Source Software**, v. 3, n. 29, 2018, p. 861.

MITCHELL, T. **Machine Learning**, ed. 1. São Francisco: McGraw-Hill, 1997.

MURATA, H.; ONODA, T. **Estimation of power consumption for household electric appliances**. In: Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02. IEEE: Singapore, 2003. Disponível em: <https://ieeexplore.ieee.org/document/1201903>. Acesso em: 13 mai. 2023.

PRINA, B. Z; TRENTIN, R. GMC: Geração de Matriz de Confusão a partir de uma classificação digital de imagem do ArcGIS®. **Anais [...]**, XVII Simpósio Brasileiro de Sensoriamento Remoto–SBSR, p. 137, 2015.

SCIKIT-LEARN. Métricas: `sklearn.metrics.rand_score`, Scikit-Learn, 2023. Disponível em: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.rand_score. Acesso em: 17 ago. 2023.

TIBCO Software Inc. **Parallel Coordinate Plot**. TIBCO Software, c2023. Disponível em: https://docs.tibco.com/pub/sfire-cloud/12.4.0/doc/html/pt-BR/TIB_sfire_client/client/topics/pt-BR/parallel_coordinate_plot.html. Acesso em: 15 jul. 2023.

VAN DER MAATEN, L.; HINTON, G. Visualizing data using t-SNE. **Journal of machine learning research**, v. 9, n. 11, 2008.

Enviado: 10 de outubro de 2023.

Aprovado: 21 de novembro de 2023.

RC: 150643

Disponível em: <https://www.nucleodoconhecimento.com.br/engenharia-eletrica/consumo-de-energia-eletrica>



MULTIDISCIPLINARY SCIENTIFIC JOURNAL

**NÚCLEO DO
CONHECIMENTO**

REVISTA CIENTÍFICA MULTIDISCIPLINAR NÚCLEO DO
CONHECIMENTO ISSN: 2448-0959

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

