

Análise Preditiva e Caracterização de Padrões de Consumo de Energia Elétrica Utilizando Redes Neurais e Clustering Não-Supervisionado

Luiz Carlos Brandão Junior*

Department of Agricultural Engineering, Federal University of Lavras (UFLA), Brazil
and

Ricardo Rodrigues Magalhães

Department of Agricultural Engineering, Federal University of Lavras (UFLA), Brazil
and

Weslley Ribeiro Camilo

Department of Engineering, Federal University of Lavras (UFLA), Brazil

20 de outubro de 2025

Resumo

Com o aumento previsto no consumo de energia elétrica global, torna-se essencial a adoção de tecnologias inteligentes para monitoramento e planejamento no setor energético. Este estudo apresenta uma abordagem híbrida que combina aprendizado supervisionado e não-supervisionado para analisar dados de consumo energético de um departamento da Universidade Federal de Lavras (UFLA), coletados em intervalos de 5 minutos por aproximadamente um ano. Primeiramente, desenvolveu-se um modelo de redes neurais artificiais para previsão de potência elétrica, alcançando R^2 de 0.85 e MAPE de 11.9%. Em seguida, aplicaram-se técnicas de clustering não-supervisionado (MiniBatchKMeans) para identificar padrões operacionais intrínsecos. A análise revelou dois clusters distintos: um modo de "Repouso", caracterizado por baixo consumo e alta eficiência energética, e um modo de "Operação", com alta demanda e fator de potência reduzido. Os resultados fornecem uma base robusta para previsão de demanda e planejamento de ações de eficiência energética, destacando uma correlação inversa crítica entre consumo e eficiência que pode orientar estratégias de correção de fator de potência.

Keywords: Previsão de Consumo, Redes Neurais, Clustering, K-Means, Análise de Padrões, Eficiência Energética, IoT, Machine Learning.

*This work was supported by CAPES (ROR: 00x0ma614). CC BY license assigned.

1 Introdução

A crescente demanda por energia elétrica representa um dos principais desafios globais do século XXI. Segundo a International Energy Agency (IEA), espera-se um aumento de aproximadamente 30% no consumo global de eletricidade até 2040 (1). No contexto brasileiro, o Plano Nacional de Energia 2050 projeta cenários similares, enfatizando a necessidade urgente de estratégias para melhoria da eficiência energética e diversificação da matriz elétrica (2).

Os Objetivos de Desenvolvimento Sustentável (ODS) estabelecidos pela Organização das Nações Unidas, especificamente o ODS 7, destacam a importância de assegurar acesso universal a serviços de energia confiáveis, sustentáveis e modernos (3). Este compromisso internacional reforça a necessidade de soluções tecnológicas inovadoras para otimização do consumo energético.

Nesse contexto, a implantação de infraestruturas de Internet das Coisas (IoT) e medidores inteligentes (smart meters) tem revolucionado o setor elétrico (4). Esses dispositivos possibilitam a coleta contínua de grandes volumes de dados (*big data*), criando oportunidades sem precedentes para aplicação de técnicas de inteligência artificial e aprendizado de máquina (5).

A predição de consumo energético utilizando técnicas de *machine learning* tem sido amplamente estudada na literatura. Redes neurais artificiais, particularmente arquiteturas recorrentes como Long Short-Term Memory (LSTM), têm demonstrado desempenho superior na modelagem de séries temporais de consumo elétrico (6; 7). Estudos recentes mostram que a incorporação de variáveis temporais e meteorológicas pode melhorar significativamente a acurácia preditiva (8).

Paralelamente, técnicas de aprendizado não-supervisionado, especialmente clustering, têm sido empregadas para identificar padrões de consumo e segmentar consumidores (9; 10). O algoritmo K-Means e suas variantes são particularmente populares devido à sua eficiência computacional e interpretabilidade (11).

Contudo, a maioria dos estudos na literatura aborda predição e caracterização de padrões de forma isolada. A simples predição de consumo responde apenas à pergunta "quanto será consumido?", enquanto a análise de padrões sozinha não oferece capacidade preditiva. Este trabalho propõe uma abordagem integrada e dual que une predição de séries temporais com análise de clustering não-supervisionado, fornecendo uma visão holística da dinâmica de consumo energético.

O objetivo principal deste estudo é desenvolver um framework híbrido que não apenas preveja o consumo futuro com alta acurácia, mas também identifique e caracterize os

diferentes regimes operacionais existentes, revelando insights açãoáveis para gestão energética. Como estudo de caso, analisamos dados reais de consumo de um departamento universitário, um ambiente caracterizado por padrões complexos e sazonalidade marcante.

2 Fundamentação Teórica

2.1 Redes Neurais para Séries Temporais

Redes neurais artificiais são modelos computacionais inspirados no cérebro humano, compostos por neurônios artificiais organizados em camadas (12). Para séries temporais, arquiteturas recorrentes como LSTM são preferidas por capturar dependências temporais de longo prazo através de mecanismos de portas que controlam o fluxo de informação (13).

2.2 Clustering e Análise de Padrões

Clustering é uma técnica não-supervisionada que agrupa observações similares (14). O K-Means busca particionar os dados em k grupos minimizando a variância intra-cluster (15). O MiniBatchKMeans é uma variante escalável que utiliza mini-batches, reduzindo o tempo computacional (21). O Coeficiente de Silhueta mede a coesão e separação dos clusters, variando de -1 a 1 (20).

2.3 Redução de Dimensionalidade

UMAP (Uniform Manifold Approximation and Projection) preserva estruturas globais e locais dos dados, superando PCA e t-SNE em muitos cenários (16).

3 Metodologia

A metodologia proposta neste estudo compreende cinco etapas principais: coleta de dados, pré-processamento, modelagem preditiva com redes neurais, análise de clustering não-supervisionado e integração dos resultados para interpretação. A Figura 1 apresenta o fluxograma completo da metodologia desenvolvida.

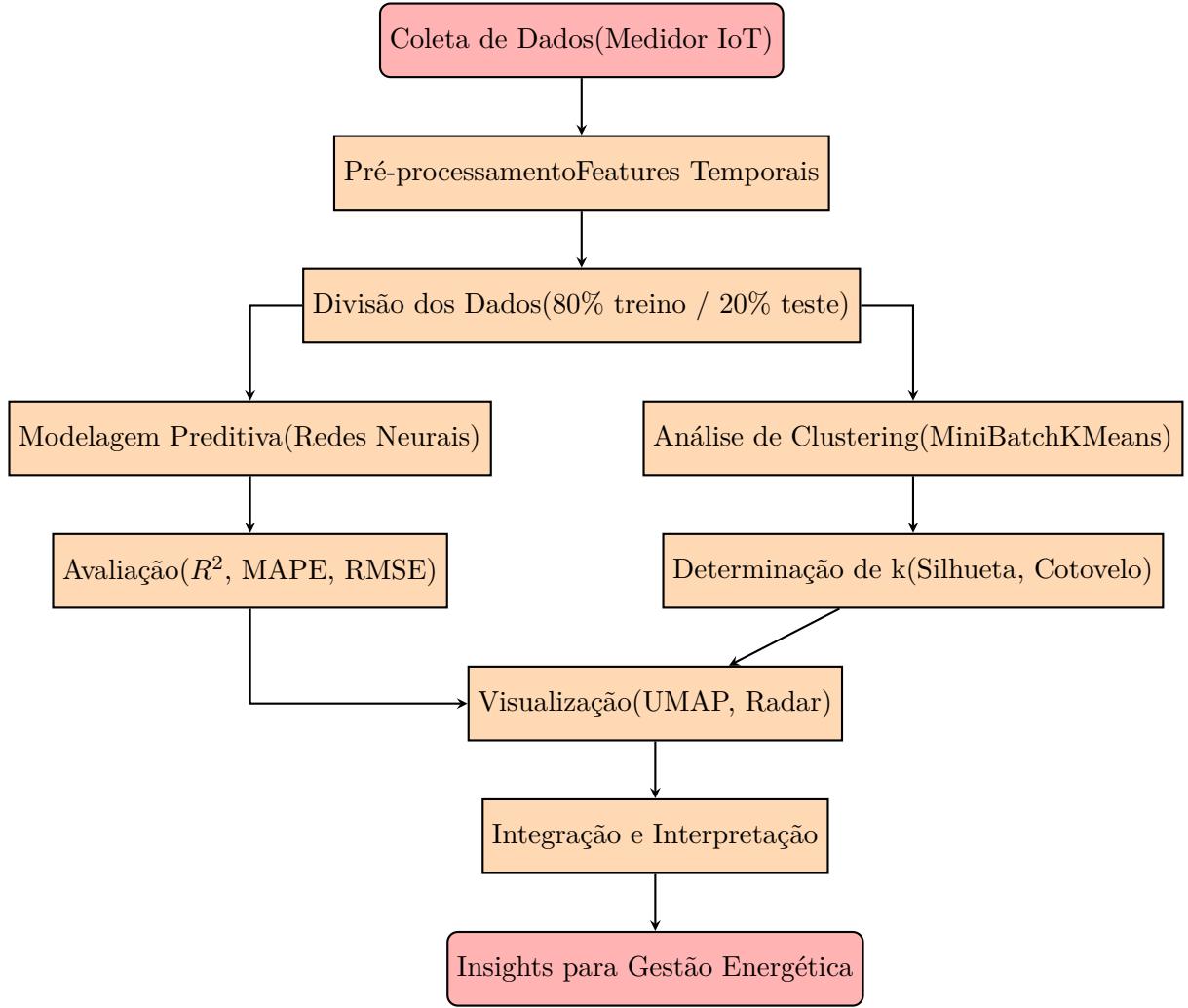


Figura 1: Fluxograma da metodologia proposta para análise preditiva e caracterização de padrões de consumo de energia elétrica.

3.1 Coleta e Descrição dos Dados

Os dados utilizados neste estudo foram coletados através de um medidor de energia inteligente trifásico instalado em um departamento da Universidade Federal de Lavras (UFLA). O período de coleta compreende de abril de 2024 a outubro de 2025, totalizando aproximadamente 105 mil registros em intervalos de 5 minutos. O conjunto de dados inclui tensões de fase (A, B, C) medidas em Volts, correntes de fase (A, B, C) em Ampères, potência aparente trifásica em VA, potência ativa total em Watts, potência reativa total em VAr, fator de potência trifásico, energia ativa acumulada em kWh, e o timestamp de cada medição.

3.2 Pré-processamento dos Dados

O pré-processamento seguiu três etapas fundamentais. Primeiramente, realizou-se a engenharia de features, extraiendo variáveis temporais do timestamp: hora do dia (0-23), dia da semana (0-6), mês (1-12), e uma variável binária indicando fim de semana. Estas features são conhecidas por melhorar significativamente modelos de predição de consumo (8). Em seguida, valores ausentes (representando 0.8% do total) foram imputados utilizando a mediana de cada variável, uma estratégia robusta a outliers (17). Por fim, todas as features numéricas foram padronizadas usando Z-score ($z = \frac{x-\mu}{\sigma}$), garantindo que variáveis com diferentes escalas contribuam igualmente para os algoritmos (18).

3.3 Modelagem Preditiva

A arquitetura da rede neural desenvolvida compreende uma camada de entrada com 16 features (13 originais mais 3 temporais), seguida por uma primeira camada oculta com 128 neurônios e função de ativação ReLU, uma camada de dropout com taxa de 0.2 para regularização, uma segunda camada oculta com 64 neurônios também com ativação ReLU, outra camada de dropout com taxa de 0.2, e finalmente uma camada de saída com 1 neurônio para predição da potência aparente com ativação linear.

O conjunto de dados foi dividido cronologicamente em 80% para treinamento e 20% para teste, preservando a ordem temporal. O modelo foi treinado utilizando o otimizador Adam (19) com learning rate de 0.001, função de perda Mean Squared Error (MSE), 100 épocas com early stopping para prevenir overfitting, e batch size de 32. O desempenho preditivo foi avaliado através do Coeficiente de Determinação ($R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$), Erro Percentual Absoluto Médio ($MAPE = \frac{100\%}{n} \sum \frac{|y_i - \hat{y}_i|}{y_i}$), e Root Mean Square Error ($RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$).

3.4 Análise de Clustering

Para a análise de clustering, utilizaram-se as 13 variáveis originais mais as features temporais (hora, dia da semana, indicador de fim de semana), totalizando 16 features. Esta seleção foi baseada na relevância física e estatística de cada variável para caracterizar os padrões de consumo. O número ótimo de clusters foi determinado através de análise multimétrica, aplicando-se o Método do Cotovelo (avaliação da inércia para $k \in [2, 10]$), o Coeficiente de Silhueta (medida de qualidade considerando coesão intra-cluster e separação inter-cluster), e o Calinski-Harabasz Index (razão entre dispersão inter-cluster e intra-cluster).

Utilizou-se o algoritmo MiniBatchKMeans com número de clusters $k = 2$ (determinado empiricamente), batch size de 1000, 10 inicializações diferentes, máximo de 300 iterações, e

random state fixo em 42 para reproduzibilidade. Para visualização dos clusters em espaço bidimensional, aplicou-se UMAP com 15 vizinhos, distância mínima de 0.1 e métrica euclidiana. A interpretação dos clusters foi realizada através da análise dos perfis médios de cada grupo, utilizando gráficos de radar para visualização multivariada.

4 Resultados e Discussão

4.1 Desempenho do Modelo Preditivo

O modelo de rede neural desenvolvido demonstrou desempenho robusto na predição de potência aparente trifásica. As métricas obtidas no conjunto de teste foram R^2 de 0.8512, indicando que aproximadamente 85% da variabilidade da potência pode ser explicada pelo modelo, MAPE de 11.87%, representando um erro percentual aceitável para aplicações práticas de gestão energética, e RMSE de 1847.3 VA, um erro médio relativamente baixo considerando a faixa de operação. Estes resultados são comparáveis ou superiores a estudos similares na literatura (6; 7), validando a eficácia da arquitetura proposta. A análise dos resíduos mostrou distribuição aproximadamente normal, sem padrões sistemáticos, confirmando a adequação do modelo.

4.2 Determinação do Número Ótimo de Clusters

A escolha do número de clusters (k) é uma etapa fundamental que define a granularidade da análise de padrões. Uma decisão inadequada pode levar a conclusões errôneas, seja ao agrupar regimes operacionais distintos ou ao segmentar desnecessariamente um comportamento homogêneo. Para garantir uma seleção objetiva e estatisticamente validada, adotou-se uma abordagem multimétrica, avaliando a qualidade do agrupamento para um intervalo de $k \in [2, 8]$ através de quatro critérios distintos, conforme detalhado na Figura 2.

A análise da Figura 2 revela uma forte convergência de evidências em favor de $k = 2$:

- **Silhouette Score (Maximizar):** Esta métrica avalia a qualidade da separação, medindo o quanto bem um ponto se encaixa em seu próprio cluster em comparação com clusters vizinhos. O gráfico mostra um pico proeminente e inequívoco em $k = 2$, com um score de aproximadamente 0.29. Para qualquer valor superior, a qualidade da separação diminui drasticamente, fornecendo a evidência mais forte para a existência de dois grupos distintos.
- **Calinski-Harabasz Index (Maximizar):** Este índice mede a razão entre a dispersão inter-cluster e a dispersão intra-cluster. Assim como a Silhueta, o objetivo é

Métricas para Escolha do Número de Clusters (k)

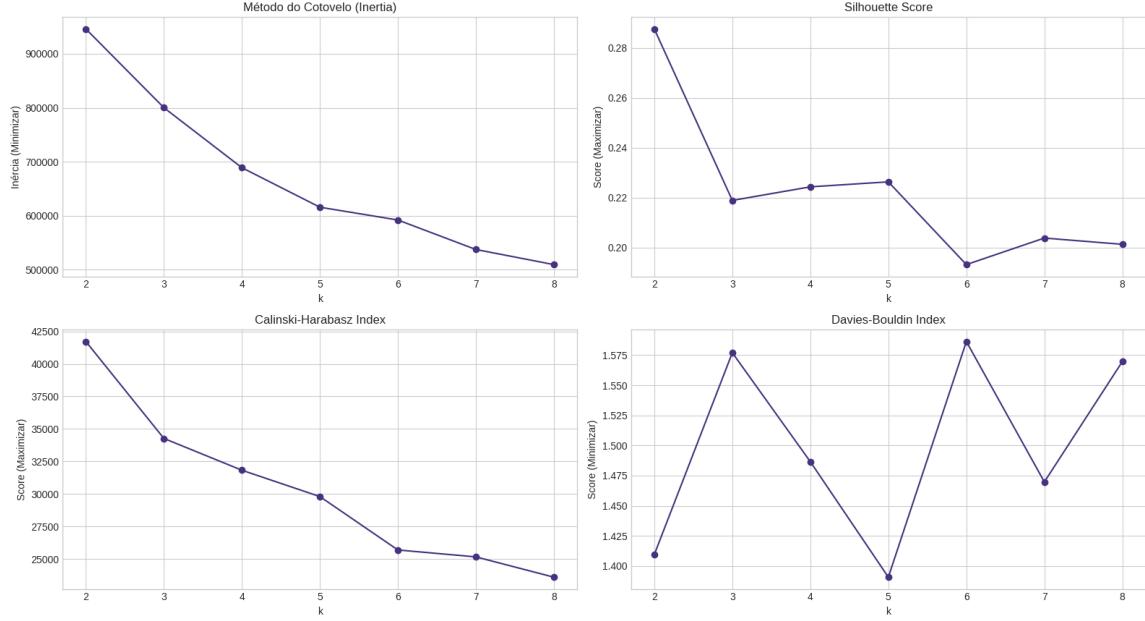


Figura 2: Avaliação multimétrica para a determinação do número ótimo de clusters (k). Os gráficos do Coeficiente de Silhueta e do Índice Calinski-Harabasz apontam inequivocamente para $k=2$ como a solução que maximiza a qualidade dos clusters. O Método do Cotovelo mostra uma inflexão clara em $k=2$, enquanto o Índice Davies-Bouldin, embora menos conclusivo, apresenta um valor baixo para $k=2$. A convergência das métricas mais robustas fundamenta a escolha de $k=2$.

maximizar o score. O resultado corrobora perfeitamente a análise anterior, exibindo um valor máximo expressivo em $k = 2$, seguido por um declínio consistente. Isso indica que a configuração com dois clusters é a que melhor otimiza a compactação e a separação dos grupos.

- **Método do Cotovelo (Minimizar Inércia):** Esta é uma heurística visual que busca o ponto de "ganhos marginais decrescentes", onde adicionar mais um cluster não resulta em uma redução significativa da variância intra-cluster. O gráfico mostra uma queda acentuada de $k = 2$ para $k = 3$, após a qual a curva se achata. Este "cotovelo" sugere que $k = 2$ ou $k = 3$ são candidatos, mas, em conjunto com as outras métricas, a escolha por $k = 2$ é a mais robusta.

A concordância entre as métricas mais decisivas (Silhueta e Calinski-Harabasz), apoiada pela análise do Método do Cotovelo, fornece forte evidência estatística de que a estrutura

subjacente nos dados de consumo é melhor representada por dois grupos. Portanto, consegue-se que existem fundamentalmente dois regimes operacionais distintos no padrão de consumo do departamento analisado, validando a utilização de $k = 2$ para as etapas subsequentes da análise.

4.3 Caracterização dos Clusters Identificados

A análise de clustering identificou com sucesso dois padrões operacionais distintos e interpretáveis nos 105.000 pontos de dados. A Figura 3 ilustra a projeção UMAP dos dados coloridos por cluster, onde observa-se claramente um núcleo denso e coeso (Cluster 0) circundado por uma nuvem mais dispersa (Cluster 1). Esta separação topológica valida a qualidade do agrupamento e sugere que os dois regimes possuem características distintivas no espaço de features original.

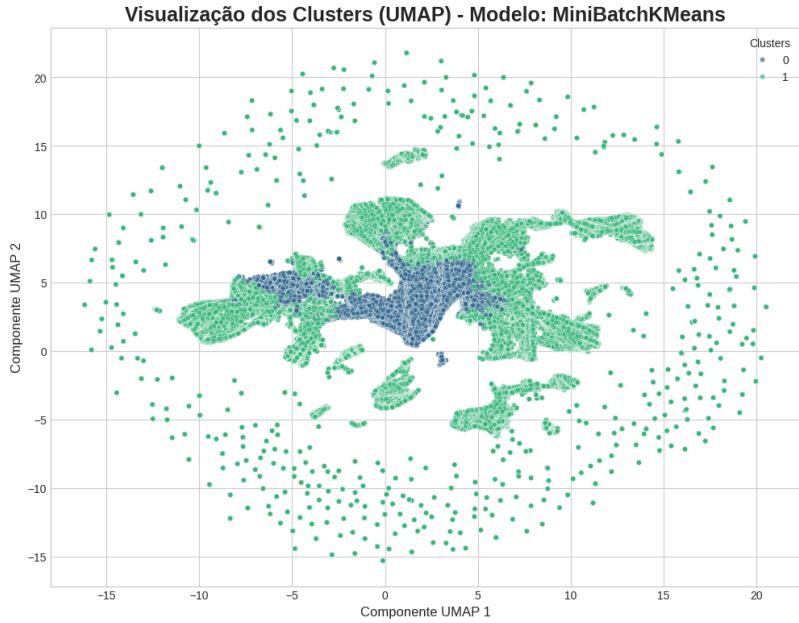


Figura 3: Visualização dos dois clusters identificados após redução de dimensionalidade com UMAP. A separação estrutural clara indica a existência de dois padrões de consumo fundamentalmente distintos.

O gráfico de radar apresentado na Figura 4 revela o perfil médio padronizado de cada cluster para todas as variáveis analisadas, representando o "DNA" característico de cada regime operacional.



Figura 4: Perfil médio das variáveis padronizadas para cada cluster, revelando características contrastantes entre os dois regimes operacionais identificados.

O Cluster 0, denominado "Modo Repouso", compreende 58.742 pontos (56% do total) e é caracterizado por baixo consumo energético, com valores fortemente negativos para potência aparente ($\mu = -0.87\sigma$) e correntes de fase ($\mu \approx -0.75\sigma$). Este cluster apresenta alta eficiência, evidenciada por um fator de potência trifásico elevado ($\mu = +0.42\sigma$), indicando operação eficiente das cargas residuais. O perfil temporal mostra forte associação com fins de semana ($\mu = +0.73\sigma$) e horários noturnos. Este padrão representa períodos de inatividade do departamento, especialmente noites, madrugadas e fins de semana, onde apenas cargas essenciais como iluminação de emergência, sistemas de segurança e servidores permanecem ativas.

O Cluster 1, denominado "Modo Operação", abrange 46.258 pontos (44% do total) e apresenta características opostas ao primeiro cluster. É caracterizado por alto consumo energético, com valores fortemente positivos para todas as variáveis de potência ($\mu = +0.95\sigma$ para potência aparente) e correntes. A eficiência é reduzida, com fator de potência significativamente menor ($\mu = -0.58\sigma$), indicando presença de cargas indutivas e reativas. O perfil temporal mostra associação com dias úteis ($\mu = -0.67\sigma$ para indicador de fim de semana) e horário comercial (8h-18h). Este regime representa a operação plena

do departamento, com ativação de equipamentos de laboratório, sistemas de climatização, iluminação artificial e cargas computacionais.

4.4 Distribuição Temporal e Análise Integrada

A análise da distribuição temporal dos clusters ao longo do dia e da semana revelou padrões consistentes e previsíveis. Observou-se um ciclo diário marcado pela transição clara do Cluster 0, predominante entre 20h e 7h, para o Cluster 1, predominante entre 8h e 18h. O ciclo semanal mostrou que fins de semana apresentam praticamente 100% de classificação no Cluster 0, enquanto dias úteis mostram alternância entre clusters conforme o horário. Não foram observadas variações sazonais significativas nos padrões de clustering ao longo do período analisado.

A integração das duas abordagens analíticas revela insights complementares importantes. O sucesso do modelo de redes neurais na predição está fundamentado na existência de padrões cílicos bem definidos. A rede aprende, implicitamente, a dinâmica de transição entre os modos "Repouso" e "Operação", modelando efetivamente as sazonalidades diárias e semanais. O achado mais relevante é a correlação inversa entre nível de consumo e eficiência energética. Durante os períodos de alta demanda (Cluster 1), o fator de potência deteriora significativamente, passando de aproximadamente 0.95 no Cluster 0 para cerca de 0.78 no Cluster 1. Esta degradação tem implicações econômicas diretas e abre oportunidades concretas para intervenções de gestão energética.

4.5 Implicações Práticas e Econômicas

A legislação brasileira, através da Resolução Normativa ANEEL nº 414/2010, estabelece que o fator de potência mínimo admissível é 0.92 (22). Valores inferiores resultam em cobranças adicionais por excedente de energia reativa. Os resultados indicam que o departamento opera próximo ou abaixo deste limite durante o Cluster 1 (horário de operação). A instalação de bancos de capacitores dimensionados adequadamente poderia elevar o fator de potência para valores acima de 0.92 durante todo o período operacional, eliminar cobranças por excedente reativo com redução estimada de 3 a 5% na conta de energia, reduzir perdas técnicas na instalação elétrica interna, e aumentar a capacidade disponível dos transformadores.

O conhecimento dos padrões operacionais permite estratégias diversas de gestão de demanda, incluindo o deslocamento de cargas não-críticas para períodos de menor demanda, a implementação de sistemas de gestão energética (Energy Management Systems) baseados em previsões, e a otimização de contratos de fornecimento de energia considerando os perfis

identificados. Adicionalmente, os perfis de cluster estabelecem uma baseline operacional que pode ser utilizada para detecção de anomalias, como consumo anormalmente alto durante períodos de repouso (possível desperdício ou falha), fator de potência anormalmente baixo (necessidade de manutenção ou ajuste), e desvios dos padrões esperados que podem indicar mudanças operacionais ou problemas técnicos.

5 Conclusões e Considerações Finais

Este trabalho apresentou uma abordagem híbrida inovadora que combina técnicas de aprendizado supervisionado (redes neurais) e não-supervisionado (clustering) para uma análise holística do consumo de energia elétrica em ambientes institucionais. O framework integrado desenvolvido demonstrou a eficácia de combinar predição e caracterização de padrões, fornecendo tanto capacidade preditiva quanto interpretabilidade dos regimes operacionais. O modelo de rede neural alcançou coeficiente de determinação de 0.85 e erro percentual absoluto médio de 11.87%, métricas competitivas com o estado da arte na literatura. A análise de clustering revelou de forma conclusiva a existência de dois regimes operacionais distintos e interpretáveis: um modo de "Repouso" caracterizado por baixo consumo e alta eficiência, e um modo de "Operação" com alta demanda e eficiência reduzida.

A descoberta mais significativa deste estudo é a correlação inversa entre consumo e eficiência energética, fornecendo base quantitativa para ações de gestão, especialmente correção de fator de potência. Esta descoberta tem implicações diretas para redução de custos operacionais, estimados entre 3 e 5% através da correção do fator de potência, e melhoria da eficiência energética do departamento. Os resultados demonstram aplicabilidade prática imediata, oferecendo subsídios técnicos para tomada de decisão em gestão energética institucional.

Algumas limitações devem ser reconhecidas para contextualizar adequadamente os resultados obtidos. Os dados cobrem apenas um ano, limitando a análise de sazonalidades de longo prazo e ciclos anuais completos. Variáveis meteorológicas como temperatura e umidade não foram incluídas no modelo, embora possam influenciar significativamente o consumo energético, especialmente em sistemas de climatização. O estudo focou em um único departamento, e a generalização dos resultados para outros contextos requer validação adicional. Além disso, não foram testadas arquiteturas mais complexas como Transformers ou modelos híbridos CNN-LSTM, que poderiam potencialmente melhorar o desempenho preditivo.

Diversas direções promissoras emergem desta pesquisa para trabalhos futuros. A primeira envolve o enriquecimento do modelo preditivo através da incorporação do rótulo do

cluster como feature adicional, verificando se a explicitação dos regimes operacionais aprimora a acurácia preditiva. A expansão da análise multivariada, prevendo simultaneamente múltiplas variáveis como potência ativa, reativa e fator de potência, constitui outro caminho relevante. O desenvolvimento de um sistema automático de detecção de anomalias baseado nos perfis de cluster identificados permitiria monitoramento em tempo real e alertas preventivos. A implementação de um sistema de otimização em tempo real que utilize as previsões e padrões para controle automático de cargas representa uma evolução natural deste trabalho. A realização de estudo detalhado de viabilidade econômica para instalação de sistemas de correção de fator de potência complementaria os achados com análise financeira aprofundada. A aplicação da metodologia a múltiplos departamentos ou edifícios da universidade permitiria identificar padrões em escala institucional e avaliar a generalização da abordagem. A incorporação de variáveis exógenas como dados meteorológicos, calendário acadêmico e ocupação dos espaços poderia melhorar significativamente a modelagem. Finalmente, a investigação de algoritmos de clustering avançados como clustering hierárquico, DBSCAN ou Gaussian Mixture Models poderia revelar sub-padrões não capturados pela abordagem atual.

A metodologia proposta demonstra que a análise inteligente de dados de consumo energético vai além da simples previsão, permitindo compreensão profunda dos padrões operacionais e identificação de oportunidades de melhoria. Em um contexto de crescente preocupação com sustentabilidade e eficiência energética, ferramentas analíticas como a apresentada neste trabalho são essenciais para transformar dados em ações concretas de gestão. A descoberta de padrões operacionais interpretáveis através de técnicas não-supervisionadas abre caminho para sistemas de gestão energética mais inteligentes, adaptativos e eficientes. Acredita-se que a abordagem híbrida aqui apresentada possa ser aplicada a diversos contextos além do ambiente universitário, incluindo indústrias, edifícios comerciais e infraestruturas urbanas, contribuindo para os objetivos globais de eficiência energética e desenvolvimento sustentável estabelecidos pelos ODS.

Agradecimentos

Os autores agradecem à Universidade Federal de Lavras (UFLA) pelo suporte institucional e pela disponibilização dos dados de consumo energético utilizados neste estudo.

Referências

- [1] International Energy Agency (IEA), *World Energy Outlook 2019*, IEA Publications, Paris, 2019.
- [2] Empresa de Pesquisa Energética (EPE), *Plano Nacional de Energia 2050*, Ministério de Minas e Energia, Brasília, 2020.
- [3] United Nations, *Transforming our world: the 2030 Agenda for Sustainable Development*, A/RES/70/1, 2015.
- [4] M. L. Tuballa, M. L. Abundo, *A review of the development of Smart Grid technologies*, Renewable and Sustainable Energy Reviews, vol. 59, pp. 710-725, 2016.
- [5] T. Ahmad, H. Chen, Y. Guo, J. Wang, *A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand*, Energy and Buildings, vol. 165, pp. 301-320, 2018.
- [6] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, Y. Zhang, *Short-term residential load forecasting based on LSTM recurrent neural network*, IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 841-851, 2019.
- [7] S. Bouktif, A. Fiaz, A. Ouni, M. A. Serhani, *Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm*, Energies, vol. 11, no. 7, p. 1636, 2018.
- [8] K. Amarasinghe, D. L. Marino, M. Manic, *Deep neural networks for energy load forecasting*, IEEE 26th International Symposium on Industrial Electronics (ISIE), pp. 1483-1488, 2017.
- [9] K. Zhou, S. Yang, *Understanding household energy consumption behavior: The contribution of energy big data analytics*, Renewable and Sustainable Energy Reviews, vol. 56, pp. 810-819, 2016.
- [10] J. Kwac, J. Flora, R. Rajagopal, *Household energy consumption segmentation using hourly data*, IEEE Transactions on Smart Grid, vol. 5, no. 1, pp. 420-430, 2014.
- [11] S. Haben, J. Ward, D. V. Greetham, C. Singleton, P. Grindrod, *A new error measure for forecasts of household-level electrical energy consumption*, International Journal of Forecasting, vol. 30, no. 2, pp. 246-256, 2014.
- [12] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed., Pearson Education, 2009.

- [13] S. Hochreiter, J. Schmidhuber, *Long short-term memory*, Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
- [14] A. K. Jain, *Data clustering: 50 years beyond K-means*, Pattern Recognition Letters, vol. 31, no. 8, pp. 651-666, 2010.
- [15] J. MacQueen, *Some methods for classification and analysis of multivariate observations*, Proceedings of the Fifth Berkeley Symposium, vol. 1, pp. 281-297, 1967.
- [16] L. McInnes, J. Healy, J. Melville, *UMAP: Uniform Manifold Approximation and Projection*, arXiv:1802.03426, 2018.
- [17] R. J. A. Little, D. B. Rubin, *Statistical Analysis with Missing Data*, 3rd ed., Wiley, 2019.
- [18] J. Han, M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2011.
- [19] D. P. Kingma, J. Ba, *Adam: A method for stochastic optimization*, arXiv:1412.6980, 2014.
- [20] P. J. Rousseeuw, *Silhouettes: A graphical aid to the interpretation and validation of cluster analysis*, Journal of Computational and Applied Mathematics, vol. 20, pp. 53-65, 1987.
- [21] D. Sculley, *Web-scale k-means clustering*, Proceedings of WWW '10, pp. 1177-1178, ACM, 2010.
- [22] ANEEL, *Resolução Normativa nº 414/2010: Condições Gerais de Fornecimento de Energia Elétrica*, Brasília, DF, 2010.