

Uso de Machine Learning para Otimização da Renovação do Parque de Medidores em Distribuidora de Energia

Kleber Silva de Medeiros (Orientando)

Especialização em Data Science & Analytics
Centro Universitário SENAI CIMATEC, Salvador-BA
E-mail: kleber.medeiros@aln.senaicimatec.edu.br

Prof. Dr. Márcio Freire Cruz (Orientador)

Centro Universitário SENAI CIMATEC, Salvador-BA
E-mail: marciofreire@gmail.com

Resumo — O presente estudo propõe uma nova abordagem para a renovação do parque de medidores de energia elétrica considerados obsoletos do ponto de vista regulatório. A quantidade expressiva de medidores antigos ainda em uso, aliada às abordagens atualmente empregadas, tem gerado altos custos operacionais e investimentos exorbitantes. Torna-se, portanto, imprescindível buscar alternativas mais eficientes. Nesse sentido, o objetivo do estudo é aprimorar a eficiência dos custos operacionais, a partir do aumento da agregação de energia, possibilitada pela identificação precisa dos equipamentos que necessitam de substituição. A abordagem atuará como um complemento às métricas já utilizadas, integrando modelos de *Machine Learning* para identificar oportunidades neste processo de renovação do parque de medidores. Os modelos incorporaram uma série de dados de entrada e selecionam de forma mais precisa e estratégica os medidores a serem substituídos, aprimorando as tomadas de decisões. Os resultados obtidos demonstram melhorias na gestão de recursos e um aumento na eficiência da agregação de energia. Esta abordagem foi implementada utilizando uma base histórica de quatro anos (2020 – 2023), permitindo a realização de uma análise comparativa entre os algoritmos de aprendizado de máquina: Floresta Aleatória, Regressão Logística e *Gradient Boosting*, em comparação com a Rede Neural Multicamadas. Cada um desses métodos foi analisado em termos de desempenho, precisão e capacidade de generalização. Nos resultados comparativos, a Rede Neural Artificial apresentou uma eficácia superior, alcançando uma acurácia de aproximadamente 75% e uma AUC de 73%. Este desempenho supera significativamente a metodologia atualmente utilizada, que não ultrapassa 39% de acurácia. Assim, será possível otimizar os processos de atualização tecnológica, impulsionando a excelência operacional e promovendo um uso mais eficiente e sustentável da energia elétrica.

Palavras-chave — Redes neurais multicamadas, Agregação de energia elétrica, Medidores eletromecânicos, Medidores Eletrônicos, Histórico de consumo de energia elétrica.

I. INTRODUÇÃO

Atualmente, o parque metrológico brasileiro possui ainda uma grande diversidade de tipos de medidores de energia elétrica, equipamentos estes responsáveis pelo faturamento de

energia elétrica das unidades consumidoras. Estes equipamentos podem ser classificados em três tipos: eletromecânicos, considerado o mais antigo, possuindo seu princípio de funcionamento na indução eletromagnética; os eletrônicos, que realizam a medição de forma digital a partir dos circuitos integrados; e o mais atuais, os medidores inteligentes, que além das funcionalidades dos medidores do tipo eletrônico podem se integrar a redes inteligentes, proporcionando uma maior integração com a rede elétrica.

Segundo os dados do estudo [1], em 2009, os medidores eletromecânicos correspondiam à 92,61% do parque instalado, enquanto os medidores eletrônicos à 7,39%, no cenário nacional, porém hoje ainda existem quantidades expressivas tanto para os medidores eletromecânicos, quanto para os medidores eletrônicos, visto que, uma das premissas para atualização tecnológica do parque se dará pelo tempo de vida útil regulatória/econômica, que, segundo definido pelo Manual de Controle Patrimonial do Setor Elétrico – MCPSE, os medidores eletromecânicos possuem um tempo de vida útil econômica de 25 anos e os medidores eletrônicos 13 anos [2]. Este tempo não se refere à vida metrológica, ou seja, quanto tempo o medidor ficará dentro da classe de exatidão, e sim ao processo de permanência do medidor em operação. Desta forma, o capital investido pela distribuidora de energia é remunerado em sua totalidade pela tarifa de energia [2].

A distribuição de medidores eletromecânicos e eletrônicos na concessionária de energia elétrica que disponibilizou os dados históricos para este trabalho está numa proporção de 30/70, respectivamente. Essa disparidade representa um desafio considerável, pois implica uma substituição em grande escala, dada a magnitude dos custos envolvidos, tanto em termos de investimento quanto de gastos operacionais. Diante dessa realidade, torna-se claro que uma troca em larga escala é inviável. O processo atual de renovação do parque baseia-se em critérios técnicos e na idade dos equipamentos.

Para enfrentar esse desafio, este estudo utilizou uma base histórica de substituição de medidores de energia elétrica para identificar padrões baseados na eficiência observada na troca desses equipamentos. Foram experimentados diversos algoritmos de aprendizado de máquina, incluindo redes neurais,

observando o mais eficaz e a melhor adequação à base de dados. A partir do modelo final encontrado, pretende-se criar uma abordagem para a seleção dos próximos equipamentos considerados obsoletos a serem substituídos, visando a eficiência dos custos operacionais.

II. TRABALHO RELACIONADOS/REFERENCIAL TEÓRICO

A renovação do parque de medidores de energia elétrica é um tema recorrente e de grande interesse para as distribuidoras de energia. Isso se deve à necessidade de atualizar o parque para garantir a eficiência nas medições de energia elétrica, além de ser um dos critérios para a remuneração dos investimentos [2], conforme as regras estabelecidas pelo órgão regulador. Contudo, o volume de medidores que precisam ser substituídos é expressivo, o que gera um desafio significativo: como realizar essa atualização sem incorrer em custos exorbitantes? Os estudos [3, 4] discutem a confiabilidade e desempenho dos medidores para garantir a precisão na medição do consumo de energia elétrica. Já a pesquisa discutida em [5], traz uma visão a respeito da previsão de falha dos medidores, ou seja, a antecipação da substituição deles. No entanto, essas pesquisas se concentram em critérios técnicos e na idade dos medidores, sem considerar a eficiência na troca baseada em análise de desempenho histórico.

Algoritmos de Aprendizado de Máquina em Processos de Industriais

Nos processos industriais, algoritmos de aprendizado de máquina são frequentemente utilizados para identificar padrões que podem indicar falhas nos processos [6]. Técnicas como redes neurais e florestas aleatórias são capazes de detectar padrões complexos em grandes conjuntos de dados, fornecendo insights valiosos para a tomada de decisão. Essas técnicas avançadas permitem uma análise mais precisa e eficiente, melhorando a manutenção preditiva e a otimização dos processos industriais, resultando em maior eficiência e redução de custos operacionais.

Pode-se observar que existe uma lacuna na literatura no que diz respeito a integração da análise de eficiência histórica na seleção dos medidores para substituição, pois a maioria das pesquisas realizadas não considera a eficiência observada na troca de equipamentos anteriores como critério complementar para futuras substituições. Essa lacuna justifica a necessidade de uma abordagem que utilize os dados históricos e algoritmos de aprendizado de máquina para torna a renovação mais eficiente.

A teoria da decisão baseada em dados, enfatiza o uso de análises quantitativas e dados históricos para informar decisões estratégicas [7]. Aplicando essa teoria ao contexto de renovação de medidores, podemos utilizar algoritmos de aprendizado de máquina para analisar dados históricos de substituição e identificar padrões que torne o processo e a tomada de decisão mais eficientes.

Aprendizado de Máquina e Otimização

O aprendizado de máquina é uma subárea da inteligência artificial focada em desenvolver algoritmos que aprendem com dados e fazem previsões ou decisões baseadas nesses dados. Em

particular, técnicas supervisionadas, como redes neurais e florestas aleatórias, entre outros, são eficazes para problemas de classificação e podem ser aplicadas para prever a eficiência da substituição de medidores. Esses algoritmos são capazes de capturar relacionamentos complexos nos dados, tornando-os ideais para a tarefa proposta [8].

Floresta Aleatória

A Floresta Aleatória pode ser compreendida como um conjunto de Árvores de Decisão, empregando o método de *ensemble* [9] que consiste em uma técnica utilizada em aprendizado de máquina que combina as previsões de múltiplos modelos (normalmente fracos ou de desempenhos mediano), e assim, formar um modelo mais robusto e com desempenho desejado. A seguir podemos destacar alguns dos tipos métodos de *ensemble*:

- *Bootstrap Aggregating* que é uma técnica onde os vários modelos são treinados separadamente, utilizando-se de subconjunto de dados, e por fim, as previsões são obtidas a partir das médias para a regressão, pelo voto majoritário para a classificação, o *Random Forest* é um exemplo disso;

- Já o *Boosting* é uma técnica que ajusta iterativamente o modelo ao erro do conjunto de dados de treinamento, e a cada iteração, os dados com as piores previsões recebem o maior peso, e um novo modelo é treinado para efetuar a correção desses erros. Na concepção final é uma combinação ponderada dos modelos, o algoritmo *Gradient Boosting* utiliza-se dessa técnica;

- A técnica do *Stacking* utiliza-se de meta modelos para combinar suas previsões, com dado de entrada as previsões dos modelos de base, e assim, aprendendo a melhor forma de combinar essas previsões;

- E por último o *Voting* que se utiliza da votação na classe de saída e da classe com maior número de votos é a escolhida como previsão final, aplica-se tanto na regressão (pela média de previsões), quanto nos problemas de classificação.

Com base nos métodos discutidos anteriormente, fica claro que o desempenho da Floresta Aleatória depende de seus dois componentes principais: as árvores de decisão e o método de *ensemble*. O desafio principal ao utilizar a Floresta Aleatória é garantir uma boa variabilidade entre as árvores, o que resulta em um maior poder de generalização do modelo. Um dos métodos eficazes para alcançar essa diversidade é o uso do algoritmo de Florestas Aleatórias [10].

Regressão Logística

Trata-se de uma técnica muito disseminada para desenvolvimento de modelos de *credit scoring* [11], onde não se exige a normalidade das variáveis independentes, e assim, tornando-a mais robusta quando ela não é atendida [12]. A regressão logística diz respeito a probabilidade de um determinado evento ocorrer, ou seja, limita-se entre 0 e 1.

Gradient Boosting

Também muito utilizado em problemas de classificação, onde produz um modelo preditivo na forma de um conjunto de modelos de previsão, o modelo *Gradient Boosting* funciona com base na otimização de uma função de perda arbitrária diferenciável, pois combina em sua estrutura vários modelos mais fracos de maneira adaptativa para produzir um modelo altamente preciso. Pode-se dizer que as árvores da Floresta Aleatória são formadas em paralelo, já as *Boosting* em série de forma sequencial [13]. Ou seja, cada uma trabalha para corrigir o erro da anterior e assim sucessivamente. Os resultados não são independentes, e por consequência o resultado também precisará passar por cada árvore [14]. Segue abaixo, de forma ilustrativa, na figura 01, como se dá a disposição das árvores em série:

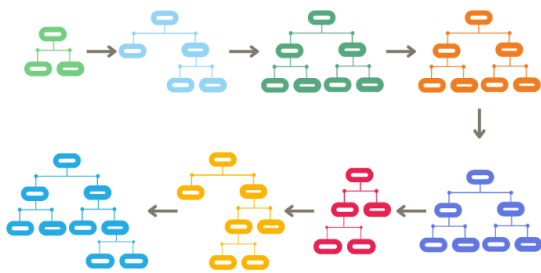


Figura 01 : Gradient Boosting – Árvores em Série
Fonte: Autor (2024)

Rede Neural Artificial

Por se tratar de um problema de classificação binária, a aplicação de RNAs tem evidenciado bons resultados em diversos segmentos, entre os quais a detecção de falhas [15], diagnósticos médicos [16], reconhecimento de voz [17], entre outras. Existem diversos tipos de RNAs, porém para a tarefa de classificação tem se destacado rede neural *Multilayer Perceptron* (MLP), que é um modelo de aprendizado de máquina composto por várias camadas de neurônios, e conectadas por meio de pesos que se ajustam ao longo do treinamento. A rede consiste na seguinte configuração: uma camada de entrada, uma ou mais camadas ocultas e por fim uma camada de saída, conforme apresentado na figura 02.

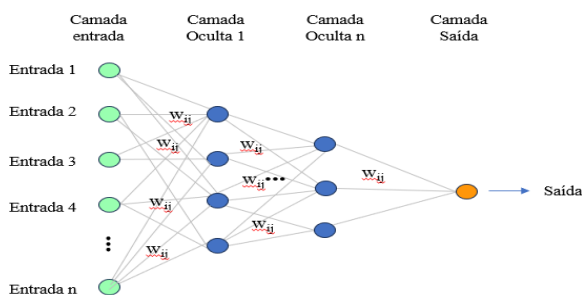


Figura 02: Rede Neural – MLP
Fonte: Autor (2024)

Métricas para Modelos de Classificação

Após a descrição dos algoritmos utilizados neste estudo, é importante reconhecer que nenhum modelo consegue representar os dados com total exatidão ou alcançar o resultado ideal. No entanto, o objetivo é que o modelo se aproxime o máximo possível desse resultado. Na literatura, existem diversas métricas para avaliar o quão próximas ou distantes estão as previsões feitas por um modelo de classificação. Essas métricas geralmente se baseiam em uma matriz que organiza as previsões do modelo, conhecida como matriz de confusão. Em um modelo de classificação binária, que envolve duas classes possíveis, a distribuição das previsões é ilustrada na Figura 03, onde VP representa os verdadeiros positivos, VN os verdadeiros negativos, FN os falsos negativos e FP os falsos positivos. Em uma matriz de confusão, cada linha corresponde a uma classe real, enquanto cada coluna representa a classe prevista pelo modelo. Assim, os valores nos quadrantes de verdadeiros positivos e verdadeiros negativos indicam que o modelo classificou corretamente, enquanto os quadrantes de falsos positivos e falsos negativos indicam erros de classificação [18].

		Predito	
		Positivo	Negativo
Real	Positivo	VP	FN
	Negativo	FP	VN

Figura 03 : Matriz de Confusão
Fonte: Autor (2024)

Neste trabalho, serão adotadas algumas das métricas mais conhecidas para a avaliação de problemas de classificação binária, onde cada uma delas desempenham um papel crucial na avaliação do desempenho do modelo sob diferentes aspectos [19], são elas:

Acurácia

Representa a proporção de previsões corretas (tanto positivas quanto negativas) em relação ao total de casos avaliados. É útil para medir o desempenho geral do modelo, mas pode ser enganosa em conjuntos de dados desbalanceados.

Precisão (*Precision*)

Mede a proporção de verdadeiros positivos em relação ao número total de previsões positivas feitas pelo modelo. É crucial quando o custo de falsos positivos é alto.

Revocação (*Recall*)

Também conhecida como sensibilidade, mede a proporção de verdadeiros positivos identificados corretamente em relação ao total de instâncias que realmente são positivas. É importante quando o custo de falsos negativos é alto.

F1-score

É a média harmônica entre precisão e revocação, proporcionando uma métrica equilibrada, especialmente útil em situações em que há um trade-off entre precisão e revocação.

AUC-ROC (Área Sob a Curva ROC)

Representa a capacidade do modelo em distinguir entre classes, com valores mais próximos de 1 indicando melhor desempenho.

A revisão da literatura indica que, embora existam estudos sobre confiabilidade, desempenho e previsão de falhas em medidores de energia, além da aplicação de aprendizado de máquina para otimização de processos, ainda há uma lacuna significativa na integração dessas áreas, especialmente com foco na eficiência histórica das substituições de medidores. Este estudo busca preencher essa lacuna, aplicando algoritmos de aprendizado de máquina para desenvolver uma abordagem mais eficiente na seleção de medidores obsoletos para substituição. A eficácia da abordagem proposta será medida utilizando métricas de avaliação como acurácia, precisão, revocação, e F1-score e a AUC-ROC, além de uma análise comparativa entre os modelos. A otimização do processo será avaliada com base na redução do número de substituições desnecessárias e na maximização da vida útil dos medidores, garantindo assim uma gestão mais eficiente dos recursos e uma melhoria nos resultados operacionais.

III. METODOLOGIA

Para conduzir a pesquisa, inicialmente, os dados foram extraídos de múltiplas plataformas *System Applications and Products in Data Processing* (SAP) [20], e armazenados em um gerenciador de banco de dados utilizando o *software Microsoft Access* [21]. Isso permitiu a criação das conexões com os arquivos gerados, bem como o tratamento dos dados sensíveis localmente. Após o tratamento inicial, foi gerado um arquivo final que foi submetido ao ambiente de desenvolvimento *Google Colab* [22]. Utilizando a biblioteca *Pandas*[23] do *Python* [24], os dados foram organizados em conjuntos, permitindo assim a continuidade do processo de Descoberta de Conhecimento em Bases de Dados (KDD - Knowledge Discovery in Databases) [25], conforme ilustrado na figura 04.

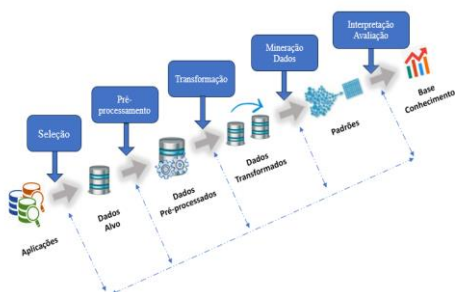


Figura 04: Processo extração de informações de base de dados - KDD

Fonte: Autor (2024)

a) Seleção:

A base histórica de atualizações já realizadas pela distribuidora no parque de medidores dispõe de dados referentes há quatro anos – 2020 a 2023. Diante dessa disponibilidade de dados, tornou-se necessário estruturar o processo de seleção, extração e pré-tratamento das informações correspondentes e tratamento de dados sensíveis, conforme ilustrado a seguir na figura 05.

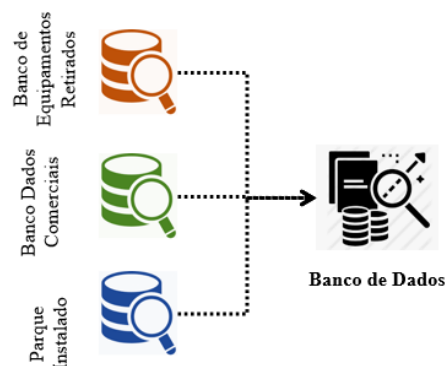


Figura 05: Processo de Extração e pré-processamento

Fonte: Autor (2024)

Todos os registros dos medidores retirados das unidades consumidoras de energia elétrica na região de concessão da distribuidora, assim como os dados comerciais, são armazenados em sistemas de múltiplas plataformas *System Applications and Products in Data Processing* (SAP). Esses sistemas contêm informações detalhadas sobre o processo de substituição de cada equipamento, incluindo dados técnicos como fabricante, modelo, ano de fabricação, tipo de equipamento, número de fases e tempo de operação. Nos bancos de dados comerciais, estão contidas informações sobre a energia faturada antes e após a troca, o consumo médio de energia, o tempo de agregação de energia após a substituição do equipamento, a localização física detalhada e a categoria de tarifa, regida pelos normativos em [26, 27]

Além disso, acrescenta-se uma variável relacionada ao Código Nacional de Atividade Econômica (CNAE) e outros dados considerados relevantes antes do processo de tratamento dos dados. As informações contábeis, também oriundas da mesma plataforma, incluem detalhes como o valor residual do equipamento retirado, custo operacional e valor total investido na renovação tecnológica do respectivo equipamento conforme regulamentação [2, 27].

Programas desenvolvidos especificamente permitiram a extração desses dados para alimentar um Sistema de Gerenciamento de Banco de Dados (SGBD) no *Microsoft Access*, facilitando a análise e tratamento preliminar dos dados sensíveis dados.

Após a consolidação das variáveis, o banco resultante totalizou 428.764 registros e 24 variáveis.

b) Pré-processamento:

Na primeira fase do pré-processamento, as variáveis foram identificadas e separadas por tipo - categóricas, de ponto flutuante, de objeto e inteiras. Em seguida, foram aplicadas as abordagens adequadas para consolidar o conjunto de dados, preparando-o para as próximas etapas do processo de pré-processamento [28]. Após identificar e separar as variáveis, realizou-se a verificação de dados duplicados e tratamento de valores ausentes ou inconsistentes nos registros, resultando na remoção desses dados, o que retornou uma quantidade menor nos registros, porém foi mantida a quantidade de variáveis, inicialmente.

Durante a análise de dados, observou-se uma redução notável no número de substituições de equipamentos de medição ao longo dos anos, juntamente com uma leve melhoria nos resultados agregados. No entanto, durante essa análise, identificou-se a falta de informações sobre a origem de alguns fabricantes e modelos, o que pode ser atribuído a inconsistências na manutenção dos registros. Para garantir a precisão das métricas dos modelos a serem avaliados, foi realizada uma limpeza desses dados. Além disso, notou-se a ausência de informações sobre o consumo médio antes e após a substituição do equipamento, aspecto fundamental para a análise. Portanto, esses registros foram excluídos do estudo.

Em uma das etapas do pré-processamento de dados, aplicou-se a matriz de correlação de *Pearson* [29] ao conjunto de dados. A análise das correlações entre diversas variáveis e a agregação total de energia revelou os seguintes resultados:

- Para agregação média a correlação apresentada (0.869699) muita alta (próxima de 1) indicando que a agregação média e a agregação total são praticamente a mesma coisa ou estão muito fortemente relacionadas;
- A variável quantidade de ocorrência durante o ano da agregação de energia apresentou uma correlação positiva (0.122506) fraca sugere que um aumento em agregação total tende a acompanhar um pequeno aumento na quantidade de ocorrências, mas a relação não é forte;
- Para as variáveis a seguir as correlações são positivas porém muito fracas, indicando que elas não tem relação linear significativa, são elas: média de energia medida no ano corrente da troca (0.116481), total de energia medida no ano corrente da troca (0.097598), classificação da tensão de fornecimento (0.070584), número de fases (0.068990), tensão de medição (0.068614), ano de fabricação do equipamento de medição (0.057460), referência histórica de falhas nos equipamentos de medição (0.051770), tipo de equipamento de medição (0.049303), ano da ação de retirada do equipamento de medição (0.009431);
- Média de energia medida no ano anterior à troca do equipamento (-0.048397), total de energia medida no ano anterior à troca do equipamento. (-0.054325), tempo de vida útil do equipamento de medição(-0.057126), mês da retirada do equipamento(-0.086139), possuem correlações negativas muito fracas com agregação total sugerindo que, à medida que a agregação total aumenta, essas variáveis tendem a diminuir ligeiramente, mas a relação não é significativa.

A análise detalhada dessas variáveis revelou que muitas delas eram redundantes ou complementares em termos de informação. Com o objetivo de evitar possíveis problemas na performance dos modelos de aprendizado de máquina, como a amplificação dos erros ou superestimação da importância das variáveis, optou-se por eliminar essas variáveis correlacionadas, resultando na exclusão de duas variáveis do conjunto de dados.

c) Transformação:

As análises realizadas durante a exploração dos dados destacaram a importância de definir critérios claros para a seleção dos medidores a serem substituídos, com o intuito de otimizar a eficácia dessas substituições. Diante desse contexto, optou-se por enriquecer o conjunto de dados com uma nova variável, fundamentada na análise histórica das substituições. Essa variável classifica cada substituição com base em sua eficiência: atribui-se o valor 1 (um) aos casos em que a substituição resultou em melhorias significativas no processo (indicando eficiência), ou seja, houve retorno de energia agregada que compensou todos os custos de investimento e gastos operacionais com a substituição desse equipamento; atribui-se o valor 0 (zero) aos casos em que a substituição não trouxe valores de agregação de energia suficiente para cobrir todos os custos atrelados a troca do equipamento (baixa eficiência). Entende-se por energia agregada a energia acrescida após a troca do equipamento de medição, indicando que o equipamento anterior passou por algum tipo de falha em sua concepção, seja ela pelo processo de desgaste, classe de exatidão [30] do medidor diferente do anterior, ou por outros motivos.

Para definir precisamente o modo de falha dos medidores, é importante destacar que, no caso dos equipamentos de medição do tipo eletromecânico, há um viés ainda mais pronunciado. Determinados tipos de falhas são difíceis de detectar, uma vez que a maioria dos erros de medição, causados principalmente pelo envelhecimento dos componentes, não altera a qualidade da energia fornecida nem interrompe o seu fornecimento. Portanto, é necessário realizar ensaios desses medidores em laboratórios específicos para identificar essas falhas de maneira eficaz seguindo os métodos de ensaios definidas em [30].

Retornando ao ponto da eficiência na troca do equipamento, tal abordagem permite uma distinção clara entre as substituições benéficas e aquelas que não atenderam às expectativas de eficiência, facilitando a tomada de decisões estratégicas sobre futuras ações de substituição. Dessa forma, criou-se uma variável-alvo, permitindo a aplicação de modelos de classificação.

Em seguida, consolidou-se algumas variáveis, seguindo o protocolo estabelecido pela concessionária de energia, que se alinha ao critério de avaliação de lotes defeituosos conforme a norma NBR – 5426- Plano de amostragem e procedimentos na inspeção por atributos [31]. Como resultado, três variáveis foram combinadas em uma única, designada como "lote", são elas: Fabricante do medidor, modelo do medidor e ano de construção do medidor.

A partir dessa segmentação, foi realizada uma análise do comportamento da nova variável "lote" em relação à agregação total e ao desvio padrão. O objetivo foi identificar possíveis *outliers* [32] que pudessem distorcer as estatísticas descritivas

dos dados, prejudicando a interpretação adequada do conjunto de dados. Dada a necessidade de testar e treinar vários tipos de modelos de aprendizado de máquina, foi crucial evitar que esses *outliers* comprometessem modelos sensíveis a valores extremos ou afetassem a capacidade de generalização dos padrões nos dados, o que poderia resultar em um desempenho inferior em conjuntos de dados novos ou não vistos. Portanto, optou-se por realizar uma limpeza nos dados que apresentavam valores muito discrepantes em relação aos "lotes" remanescentes.

Antes de avançar para as próximas etapas, foi observado que o conjunto de dados atual ainda possuía uma quantidade considerável de variáveis - 22 no total. Foi, então, aplicado a Análise de Componente Principais – PCA [33], com o objetivo de reduzir essa dimensionalidade. No entanto, foi priorizada a interpretabilidade das variáveis originais durante esse processo, realizando uma cuidadosa avaliação em conjunto com a área de negócios da distribuidora de energia para determinar a variável deveria ser mantida ou excluída. Após várias iterações, foi possível obter um conjunto de dados mais refinado, permitindo identificar padrões de maneira mais robusta.

Por fim, reduziram-se as variáveis que deram origem às novas propostas e, após a exclusão das respectivas observações previamente identificadas, o conjunto de dados final ficou com um total de 248.380 registros e 12 variáveis, conforme quadro 01.

Quadro 01: Variáveis pós consolidação

ID	Variável	Descrição Breve
00	mes	Mês base
01	fases	N. de fases
02	tipo	Tipo Equipamento
03	lote	lote equipamento
04	tvu	Tempo de vida útil
05	loc	Localidade
06	ctg_taf	Tarifação Energia
07	tt_aa	Total energia ano anterior
08	md_aa	Média energia ano anterior
09	tt_ac	Total energia ano corrente
10	md_ac	Média energia ano corrente
11	target	Troca eficiente ("0" ou "1")

Fonte: Autor (2024)

Para a variável alvo (*target*), observou-se um desbalanceamento, com 61% para a classe 0 e 39% para a classe 1. Esse desequilíbrio é um ponto crucial a ser considerado nas ponderações dos próximos passos que definirão os modelos de aprendizado de máquina.

Dado que este é um problema de classificação binária, onde a classe 1 representa a retirada do equipamento de medição com eficiência e a classe 0 representa a retirada sem eficiência, prosseguiremos com as próximas etapas, que envolvem a experimentação dos algoritmos de aprendizado de máquina.

d) Mineração dos Dados:

Na Modelagem e Desenvolvimento, foram aplicados técnicas e algoritmos de aprendizado de máquina ao conjunto de dados final para construir modelos preditivos classificadores. O processo envolveu a seleção e ajuste dos modelos, seguido pela avaliação de sua performance por meio de métricas apropriadas. O desenvolvimento dos modelos foi iterativo, exigindo experimentação cuidadosa para assegurar a robustez e generalização dos resultados.

O primeiro passo na etapa de modelagem envolveu a transformação das variáveis em formatos numéricos (cinco variáveis), ordinais (uma variável) e categóricas (cinco variáveis). Após essa transformação, procedeu-se com a divisão do conjunto de dados em treino e teste, assegurando a aleatoriedade da seleção e devido ao volume de dados estabeleceu-se um percentual de 20% para o conjunto de teste.

O passo seguinte foi aplicar um *pipeline* [34] simples, que implementa uma sequência linear de etapas. Este pipeline inclui a normalização dos dados, a transformação das variáveis, o tratamento da classe desbalanceada.

Por fim, procedeu-se à experimentação de diversos algoritmos de aprendizado de máquina, incluindo Floresta Aleatória, Regressão Logística, *Gradient Boosting* e Redes Neurais do tipo MLP (*Perceptron* Multicamadas). Cada um desses algoritmos foi aplicado ao conjunto de dados em análise com o objetivo de identificar o modelo que proporcionasse o melhor desempenho. A comparação entre esses diferentes modelos permitiu avaliar a eficácia de cada abordagem e selecionar aquela que melhor se adequasse às características dos dados, garantindo assim a robustez e a precisão dos resultados obtidos.

IV. RESULTADOS E DISCUSSÃO

Utilizando uma abordagem comparativa, foi analisado o desempenho de quatro modelos de aprendizado de máquina: Floresta Aleatória, *Gradient Boosting*, Regressão Logística e RNA MLP. Cada modelo foi construído e avaliado com base em um conjunto de métricas de desempenho detalhadas, incluindo acurácia, precisão, *recall*, F1 e AUC-ROC. Essas métricas foram aplicadas para avaliar a eficácia dos modelos tanto na identificação de casos em que houve eficiência na troca do equipamento de medição (classe 1), bem como os casos em que não houve eficiência na substituição do medidor (classe 0).

A partir do conjunto de dados, realizou-se a separação em conjuntos de treino e teste, reservando 20% dos dados para teste. Após a divisão, cada modelo foi analisado com base nos dados de teste.

Os resultados obtidos com o modelo baseado em Floresta Aleatória são apresentados na figura 06, destacando a Matriz de Confusão correspondente. No processo de otimização do modelo baseado no algoritmo de Floresta Aleatória, foram meticulosamente ajustados certos parâmetros chave para alcançar a melhor performance. Dentre estes, destacam-se: o número de estimadores, fixado em 100, e o parâmetro *max_features*, também estabelecido em 100. Adicionalmente, considerando o desbalanceamento observado nas classes do

conjunto de dados, optou-se pela aplicação do parâmetro *class_weight* configurado como *balanced*. Os parâmetros restantes foram mantidos em seus valores padrão, seguindo as recomendações da literatura e as práticas comuns na modelagem de dados complexos.

		Previsão Modelo	
		0	1
Real	0	26.338	3.901
	1	8.502	10.935

Figura 06: Matriz Confusão – Floresta Aleatória
Fonte: Autor (2024)

Para a avaliação do modelo que utiliza Regressão Logística, conforme a figura 07, realizou-se alguns ajustes nos parâmetros e o que melhor se adequou ao conjunto de dados foi o padrão, com exceção da *class weight* que se aplicou o parâmetro *balanced*, devido ao desbalanceamento das classes presentes ao conjunto de dados.

		Previsão Modelo	
		0	1
Real	0	21.119	9.120
	1	7.148	12.289

Figura 07: Matriz Confusão – Regressão Logística
Fonte: Autor (2024)

O terceiro modelo examinado foi o baseado em *Gradient Boosting*. Os desempenhos observados são sumarizados na figura 08. Na experimentação dos parâmetros para o modelo final utilizando o algoritmo *Gradient Boosting*, verificou-se que a configuração com 50 estimadores foi a que melhor se adequou ao conjunto de dados.

		Previsão Modelo	
		0	1
Real	0	28.402	1.837
	1	12.752	6.685

Figura 08: Matriz Confusão – *Gradient Boosting*
Fonte: Autor (2024)

Finalmente, foi avaliado o modelo construído sobre uma RNA – MLP, onde a rede neural foi estruturada de acordo com a metodologia proposta, ou seja, uma camada de entrada contendo 469 variáveis, seguida por 6 camadas ocultas. Para as primeiras camadas ocultas, foi aplicada a função de ativação "RELU", junto com *batch normalization* e *dropout* com taxa de

0.3. A rede foi finalizada com uma camada de saída contendo dois neurônios e utilizando a função de ativação "SIGMOID".

Durante a construção do modelo da rede neural, foi implementado um otimizador com taxa de aprendizado de 0.006 e devido ao desbalanceamento das classes no conjunto de dados, foi necessário atribuir pesos às classes (0.812 para a classe 0 e 1.29 para a classe 1). Além disso, foram empregadas as funções de *checkpoint* e *early stopping* para monitorar e salvar o melhor modelo do treino ao longo das 100 épocas, evitando possíveis ajustes ou *overfitting* [35]. É importante ressaltar que foram realizadas diversas experimentações para os ajustes dos hiperparâmetros, com a intenção de obter a melhor performance. A seguir os resultados alcançado, conforme figura 09.

		Previsão Modelo	
		0	1
Real	0	25.193	5.046
	1	7.414	12.023

Figura 09: Matriz Confusão – MLP
Fonte: Autor (2024)

Considerando a quantidade de acertos na Classe 1 (Verdadeiros Positivos) e erros na Classe 0 (Falsos Positivos), o modelo Redes Neurais (MLP) apresenta-se até o momento como o melhor modelo. Ele oferece um bom equilíbrio com alta quantidade de acertos na Classe 1 e menos erros na Classe 0, indicando uma melhor performance geral para o objetivo desejado.

A seguir, avalia-se de forma comparativa as métricas de desempenho para cada modelo, determinou-se o mais eficiente para o nosso conjunto de dados específico. Esta avaliação foi essencial para compreender a aplicabilidade prática dos modelos em situações reais, contribuindo assim para o avanço na eficiência da troca de equipamentos em ambientes.

Na análise detalhada das métricas de desempenho, cujos resultados são consolidados na tabela 01, identificou-se que, dos modelos avaliados, a Rede Neural *Multilayer Perceptron* (MLP) também se destacou. Apesar de algumas métricas apresentarem valores similares aos obtidos pelo modelo de Floresta Aleatória, a MLP sobressaiu com um AUC-ROC, atingindo o valor de 73%, evidenciando uma maior confiabilidade nas suas previsões. Notavelmente, a eficácia do modelo em identificar corretamente casos da classe minoritária 1 - crucial no contexto de renovação do parque de medidores - demonstra sua importância prática. Este resultado é particularmente significativo, visto que a correta identificação de situações que requerem a troca de equipamentos é fundamental para maximizar a eficiência e o retorno financeiro do processo de renovação.

Tabela 01: Comparativo das Métricas – Consolidado

Modelo	Acurácia	Precisão	Revocação	FI	AUC_ROC
MLP	0,749	0,704	0,619	0,659	0,730
RF	0,750	0,737	0,563	0,638	0,717
RL	0,673	0,574	0,632	0,602	0,665
GB	0,706	0,784	0,344	0,478	0,642

Fonte: Autor (2024)

O valor de 0,73 na métrica AUC-ROC da MLP sugere um modelo com boa capacidade de distinguir entre as duas classes em questão, ou seja, uma boa medida que combina a taxa de verdadeiros positivos e falsos positivos em diferentes limiares, oferecendo uma visão holística do desempenho do modelo de rede neural, independentemente de qualquer ponto de corte específico. Isso indica que, na maioria dos casos, o modelo pôde identificar corretamente a classe positiva sem incorrer em um número excessivo de falsos positivos, o que também pode ser observado na curva AUC-ROC apresentado na figura 10. Este desempenho, embora não excepcional, é consideravelmente melhor do que o acaso (AUC-ROC = 0,5) e destacando o potencial da MLP em contextos de classificação binária. Contudo, o resultado também aponta para oportunidades de otimização, seja através do ajuste fino dos hiperparâmetros, ou outras implementações da engenharia de recursos, com o objetivo de elevar a capacidade discriminatória do modelo.

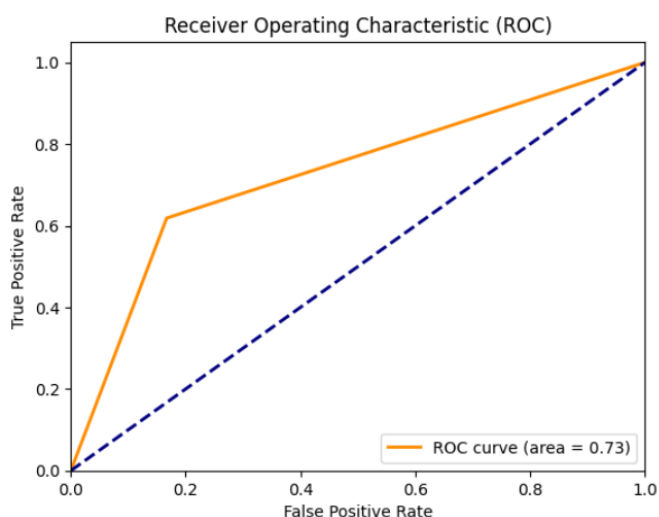


Figura 10 : Curva AUC – ROC da RNA MLP

Fonte: Autor (2024)

V. CONCLUSÕES

Com base nos resultados e nas considerações apresentadas neste estudo, verificou-se que a utilização da Rede Neural Artificial (RNA) do tipo Perceptron Multicamadas (MLP), levando em consideração o conjunto de métricas e comparado com demais modelos, atingiu o objetivo esperado. Os dados

utilizados como base para o estudo indicam que o procedimento convencional para a atualização dos medidores é ineficaz, com uma acurácia histórica que não ultrapassa 39% (considerando apenas os equipamentos corretamente substituídos). Essa ineficácia acarreta investimentos e custos operacionais elevados, além de resultar em mais de 60% de substituições desnecessárias (considerando os equipamentos removidos sem trazer benefícios), implicando na retirada de medidores que ainda poderiam estar em operação.

Ao contrastar essa prática com o desempenho do modelo de RNA MLP desenvolvido, observa-se um salto na acurácia para aproximadamente 75%. Este aumento justifica a implementação do modelo pela eficácia superior na identificação dos medidores que efetivamente necessitam ser substituídos.

Um ponto interessante é que, ao aplicar o novo modelo em toda a base de dados disponível e compará-lo com o modelo atualmente em uso, observou-se um ganho financeiro teórico de 120,8%. Essa comparação evidencia o significativo potencial econômico da nova abordagem em relação ao método tradicional, destacando a eficiência e os benefícios financeiros que podem ser obtidos com sua implementação. Assim, reforça a importância deste estudo ao apresentar uma nova estratégia para a gestão de ativos, que integra precisão analítica com considerações econômico-financeiras.

Dessa forma, ao adotar este modelo RNA MLP, as concessionárias de energia de energia não apenas podem aprimorar sua capacidade de decisão em relação à renovação de medidores, mas também alavancam significativamente a eficiência operacional e financeira, ou seja, quanto menos equipamentos são retirados desnecessariamente, menor é a geração de resíduo e maior é o ganho em termos de sustentabilidade, sugerindo um caminho promissor para futuras investigações.

Uma das propostas para a implementação em ambiente de produção, de maneira mais versátil e prática, é a experimentação do modelo estudado em cenários variados e experimentais. Devido à diversidade de tipologias de equipamentos no parque de medidores, existe uma constante necessidade de retreinar os modelos, o que dificulta a inserção direta do modelo em produção. No entanto, essa abordagem permitirá avaliar a maturidade da ferramenta proposta, auxiliando na tomada de decisões sem depender exclusivamente dela como única fonte decisória.

Considerando o cenário de potenciais melhorias, observa-se que as métricas alcançadas com redes neurais podem ainda ter espaço para aprimoramento. Neste sentido, diversas sugestões podem ser consideradas para estudos futuros. Entre elas, destaca-se a exploração de outros modelos de aprendizado de máquina não abordados até o momento, bem como uma análise mais aprofundada na seleção de *features*. A inclusão de uma avaliação amostral dos equipamentos de medição retirado considerando o estudo do modo de falha do equipamento, também pode ser benéfica para otimizar a precisão no processo de retirada dos medidores, reforçando a avaliação da parte técnica e unido este último a investigação de bibliotecas de aprendizado de máquina com enfoque em aprendizado por reforço, a partir dos modelos já implementados em produção. Esses direcionamentos sugeridos corroboram para enriquecer

ainda mais a abordagem proposta e contribuir para avanços significativos na atualização tecnológica do parque de medidores das distribuidoras de energia.

AGRADECIMENTOS

A conclusão deste trabalho não foi apenas um esforço acadêmico e profissional, mas também um desafio pessoal. Gostaria de expressar minha profunda gratidão a Deus, cuja presença foi constante, iluminando a minha trajetória e fortalecendo meu espírito nas horas dos maiores desafios. À minha família, ofereço meu mais sincero agradecimento por seu amor incondicional, paciência que foram fundamentais na construção desse projeto. Agradeço imensamente ao meu orientador, pela paciência, incentivo constante e orientação precisa, sem os quais este trabalho não alcançaria seu verdadeiro potencial. Cada um de vocês é parte essencial desta conquista.

REFERÊNCIAS

- [1] M. R. da Silva, **Modelação e análise da vida útil (metrológica) de medidores tipo indução de energia elétrica ativa**, M.S. Faculdade de Engenharia, Universidade Estadual Paulista (UNESP), Ilha Solteira, SP, Brazil, Aug. 2010. [Online]. Disponível em: [Dissertação Final \(unesp.br\)](#). Acesso em: 30 Mai 2024, p. 45.
- [2] BRASIL. Agência Nacional de Energia Elétrica – ANEEL. **Manual de Controle Patrimonial do Setor Elétrico – MCPSE**. Disponível em: [Manual de Controle Patrimonial - MCPSE \(aneel.gov.br\)](#). Acesso em: 30 Mai 2024.
- [3] M. R. da Silva, **Modelação e análise da vida útil (metrológica) de medidores tipo indução de energia elétrica ativa**, M.S. Faculdade de Engenharia, Universidade Estadual Paulista (UNESP), Ilha Solteira, SP, Brazil, Aug. 2010. [Online]. Disponível em: [Dissertação Final \(unesp.br\)](#). Acesso em: 30 Mai 2024.
- [4] D. P. Neto, **Avaliação de Desempenho de Medidor Trifásico de Energia Elétrica tipo Indução Operando com Carga Não-Lineares**. 2010. 142 f. Dissertação (Mestrado em Engenharia) - Universidade Federal de Goiás, Goiânia, 2010.
- [5] R. M. Nogueira, **Metodologia para Ensaio de vida acelerada dos medidores eletrônicos de energia elétrica em instalações de baixa tensão**. 2016. 80 f. Graduação Engenharia Elétrica - Universidade Federal do Espírito Santo, 2016.
- [6] D. A. Lobão, **Técnicas de aprendizado de máquinas aplicadas ao diagnóstico de falhas em equipamentos industriais**. 2020. 89 f. Dissertação (Mestrado em Engenharia de Controle e Automação) - Instituto Federal do Espírito Santo, Campus Serra, Vitória, 2020.
- [7] T. H. Davenport, & J.G. Harris, **Competing on Analytics: Updated, with a New Introduction: The New Science of Winning** (Revised). Boston, Massachusetts: Harvard Business Press, 2017.
- [8] I. Goodfellow, Y. Bengio, A. Courville, **Deep Learning**. MIT Press, 2016.
- [9] A. Géron, **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras e TensorFlow**, 2nd ed., Rio de Janeiro: Alta Books, p. 147, 2021.
- [10] L. Breiman, **Random Forests**, *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
- [11] P. T. M. Rosa, **Modelos de Credit Scoring: Regressão Logística, CHAID e REAL**. 125 f. Dissertação (Mestrado em Estatística) – Curso de Pós-graduação em Estatística, Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, 2000.
- [12] J. F. Hair JR., **Análise multivariada de dados**. São Paulo: Bookman, 2009.
- [13] A. Keprate, R. M. C. Ratnayake, **Using Gradient Boosting Regressor to Predict Stress Intensity Factor of a Crack Propagating in Small Bore Piping**. IEEE Int. Conf. on Industrial Engineering and Engineering Management, p. 1331-1336, 2017.
- [14] J. H. Friedman, **Greedy Function Approximation: A Gradient Boosting Machine**. *Annals of Statistics*, 1189–1232, 2001.
- [15] C. M. Bishop, **Pattern Recognition and Machine Learning**. New York, NY, USA: Springer, pp. 235-280, 2006.
- [16] X.-D. Li, J.-S. Wang, W.-K. Hao, M. Wang, and M. Zhang, **Multi-layer perceptron classification method of medical data based on biogeography-based optimization algorithm with probability distributions**, *Applied Soft Computing*, vol. 121, p. 108766, 2022.
- [17] C. Weng, D. Yu, S. Watanabe, and B.-H. F. Juang, **Recurrent deep neural networks for robust speech recognition**. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.
- [18] A. Géron, **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras e TensorFlow**, 2nd ed., Rio de Janeiro: Alta Books, p. 73. 2021.
- [19] A. Géron, **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras e TensorFlow**, 2nd ed., Rio de Janeiro: Alta Books, p. 74. 2021.
- [20] **System Applications and Products in Data Processing (SAP), SAP SE. [Software]**. Disponível em: [SAP Brasil | Software e Soluções Inteligentes na Nuvem](#). Acesso em: 30 Mai 2024.
- [21] **Microsoft Access, Microsoft Corporation. [Software]**. Disponível em: <https://www.microsoft.com/access>. Acesso em: 30 Mai 2024.
- [22] **Google Colaboratory, Google LLC. [Software]**. Disponível em: <https://colab.research.google.com>. Acesso em: 30 Mai 2024.
- [23] **Pandas: Python Data Analysis Library, NumFOCUS. [Software]**. Disponível em: <https://pandas.pydata.org> . Acesso em: 30 Mai 2024.
- [24] **Python Programming Language, Python Software Foundation. [Software]**. Disponível em: <https://www.python.org> , Acesso em: 30 Mai 2024.
- [25] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth and R. Ilturusamy, **Advances in Knowledge Discovery in Databases**, Cambridge, MA: MIT Press. 1996.
- [26] BRASIL. Agência Nacional de Energia Elétrica – ANEEL. **Resolução Normativa ANEEL N° 1000, DE 7 DEZEMBRO 2021**. Disponível em: www2.aneel.gov.br/cedoc/ren20211000.html . Acesso em: 30 Mai 2024.

- [27] BRASIL. Agência Nacional de Energia Elétrica – ANEEL. Procedimento Regulação Tarifária – Proret. Disponível em:
[Procedimentos de Regulação Tarifária - Proret — Agência Nacional de Energia Elétrica \(www.gov.br\)](#) . Acesso em: 02 Jun 2024.
- [28] S. Garcia, J. Luengo, F. Herrera, **Tutorial on practical tips of the most influential data preprocessing algorithms in data mining**, Knowledge-Based Systems, vol. 98, pp. 1-29, 2016.
- [29] J. Benesty, J. Chen, Y. Huang, and I. Cohen, **Pearson correlation coefficient," in *Noise Reduction in Speech Processing*, Topics in Signal Processing**, vol. 2, Berlin, Heidelberg: Springer, 2009, p.37. [Online]. Available: https://doi.org/10.1007/978-3-642-00296-0_5, Acesso em: 02 Jun 2024.
- [30] BRASIL. Agência Nacional de Energia Elétrica – ANEEL. **Portaria Inmetro nº 587, de 05 de novembro de 2012**. Disponível em:
<http://www.inmetro.gov.br/legislacao/rtac/pdf/RTAC001929.pdf> , Acesso em: 02 Jun 2024.
- [31] BRASIL. Associação Brasileira de Normas Técnicas – ABNT. **Plano de Amostragem e procedimentos na inspeção por atributos – NBR – 5426**. Disponível em : <https://www.normas.com.br/>. Acesso em: 02 Jun 2024.
- [32] J. Wang, **Encyclopedia of Business Analytics and Optimization** . Disponível em : [Outlier Detection in Big Data \(whiterose.ac.uk\)](#), 2014, p. 1762, Acesso em: 08 Jun 2024.
- [33] F. De la Torre, M. J. Black, **Robust principal component analysis for computer vision**, IEEE, 2001, vol. 1, pp. 362-369, doi: 10.1109/ICCV.2001.937541.
- [34] R. S. Olson, N. Bartley, R. J. Urbanowicz, and J. H. Moore, **Evaluation of a tree-based pipeline optimization tool for automating data science**, pp.485–492,[Online], Disponível em: [Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science | Proceedings of the Genetic and Evolutionary Computation Conference 2016 \(acm.org\)](#) , Acesso em: 30 Mai 2024.
- [35] J. M. Kernbach, V. E. Staartjes, **Foundations of Machine Learning-Based Clinical Prediction Modeling: Part II - Generalization and Overfitting**, 2022, pp. 35–50, Disponível em: https://doi.org/10.1007/978-3-030-85292-4_3, Acesso em: 20 Mai 2024.

**CENTRO UNIVERSITÁRIO SENAI CIMATEC
ESPECIALIZAÇÃO EM DATA SCIENCE & ANALYTICS****ATA DE APRESENTAÇÃO DE PROJETO FINAL DE CURSO**

Ata de apresentação do Projeto Final de Curso **Especialização em Data Science & Analytics**, submetido pelo aluno **Kleber Silva de Medeiros**, como parte dos requisitos para obtenção do Certificado de **Especialista em Data Science & Analytics** pelo Centro Universitário SENAI CIMATEC, às 17h00 do dia 17 de julho de 2024. Reuniu-se remotamente pela plataforma Google Meet, a Banca Examinadora designada pelo Prof. Dr. Márcio Freire Cruz - Orientador, constituída pelo Prof. Dr. Márcio Freire Cruz e pelo Prof. Msc. Flávio Santos Conterato. O orientador deu início aos trabalhos com as devidas orientações, e a exposição foi realizada pelo estudante dentro do prazo de tempo estabelecido. Ao final da apresentação a banca reuniu-se atribuindo a seguinte nota: 8,2 (oito pontos e dois décimos).

A banca de avaliadores decidiu pela:**(X) Aprovação do trabalho**

Caberá ao aluno apresentar em no máximo em 30 (trinta) dias a contar da data de assinatura desta Ata, uma cópia do trabalho em PDF, constando as considerações pontuadas pela banca. A Ata de Apresentação do Projeto Final de Curso deve ser digitalizada e inserida na terceira página do TCC ou como anexo do artigo.


() Reprovação do trabalho

O aluno terá que se matricular novamente no TCC – Trabalho de Conclusão de Curso e ser submetido a uma banca avaliadora no semestre seguinte.


As ações consequentes ao status de Aprovação deverão obedecer ao prazo proposto acima sob pena do parecer final ser modificado para o status de Reprovado automaticamente e sem possibilidade de recurso.

Para constar, lavrou-se a presente ata que vai assinada por todos os membros da Banca. Por estarem cientes de suas obrigações estão de acordo com os termos desse documento:


Salvador, 17 de julho de 2024

Documento assinado digitalmente
 **MÁRCIO FREIRE CRUZ**
Data: 17/07/2024 18:41:27-0300
Verifique em <https://validar.jf.gov.br>

Márcio Freire Cruz
Professor Orientador

Documento assinado digitalmente
 **FLAVIO SANTOS CONTERATO**
Data: 17/07/2024 19:03:18-0300
Verifique em <https://validar.jf.gov.br>

Flávio Santos Conterato
Membro da banca

Assinado eletronicamente por:
Patricia Freitas Tourinho
CPF: ***.733.265-**
Data: 16/08/2024 19:22:41 -0300


Patricia Freitas Tourinho
Coordenadora do Pós-Graduação Lato Sensu