

Comparação de Classificadores na Análise de Gases Dissolvidos

Oliveira Neto. José Jorge de* Xavier. George Victor Rocha**
Cardoso. Diego Oliveira*** Canuto. Jânio Coutinho****

* Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Sergipe, SE, (e-mail: neto.oliveira.273@gmail.com)

** Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Campina Grande, PB, (e-mail: george.xavier@ee.ufcg.edu.br)

*** Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal da Bahia, BA, (e-mail: diego.olicard@gmail.com)

**** Departamento de Engenharia Elétrica, Universidade Federal de Sergipe, SE, (e-mail: janio.canuto@gmail.com)

Abstract: Dissolved Gas Analysis (DGA) is a powerful technique for diagnosing incipient failure in oil insulated transformers. The use of intelligent classifiers has become an increasingly recurring practice in this area, being the Artificial Neural Network (ANN) one of the most popular algorithms. This work aim to compare three fast-implementing classifiers to ANN, an established but more complex method. The classifiers are: the k-nearest neighbor (KNN), one based on Fisher's Linear Discriminant and one on the Mahalanobis distance. Among these, the best results were obtained by KNN and ANN, with 81.97% and 85.25% hits during validation, followed by Fisher and Mahalanobis Discriminants, both with 70.05%. All classifiers performed well, showing superiority to some classical techniques such as Rogers and Doernenburg for the analyzed data. Thus, the three algorithms proved to be an alternative to ANN, with faster implementation and lower computational weight, with emphasis on KNN.

Resumo: A Análise de Gases Dissolvidos (AGD) é uma poderosa técnica para diagnóstico de falhas incipientes em transformadores isolados a óleo. A utilização de classificadores inteligentes tem se tornado uma prática cada vez mais recorrente nesta área, sendo a Rede Neural Artificial (RNA) um dos algoritmos mais populares. Neste trabalho, busca-se comparar três classificadores de rápida implementação com a RNA, método já estabelecido, porém mais complexo. Os classificadores são: o k-ésimo vizinho mais próximo (KNN), um baseado no Discriminante Linear de Fisher e um na métrica de Mahalanobis. Dentre esses, os melhores resultados foram obtidos pelo KNN e a RNA, com 81,97% e 85,25% de acertos durante a validação, seguidos pelo Discriminante de Fisher e Mahalanobis, ambos com 70,05%. Todos os classificadores tiveram um bom desempenho, mostrando superioridade a algumas técnicas clássicas como a de Rogers e Doernenburg para os dados analisados. Dessa forma, os três algoritmos se mostraram uma alternativa à RNA, com implementação mais rápida e com menor peso computacional, cabendo dar destaque ao KNN.

Keywords: Dissolved Gas Analysis; Pattern Recognition; Transformer Failures; Fisher Linear Discriminant; Artificial Neural Networks; Comparison of Classifiers.

Palavras-chaves: Análise de Gases Dissolvidos; Reconhecimento de Padrões; Falhas em Transformadores; Discriminante Linear de Fisher; Redes Neurais Artificiais; Comparação de Classificadores.

1. INTRODUÇÃO

Os transformadores são de extrema importância para o sistema elétrico. Naturalmente, devido às consequências graves associadas a sua falha, como perda de equipamentos, multas e risco à segurança, esse componente se torna centro de diversas pesquisas para garantir o seu funcionamento em condições normais. Por meio de diferentes técnicas de monitoramento, busca-se o momento ideal para sua manutenção em que os defeitos ainda são incipientes (Abu-

Elanien and Salama, 2010). Dentre as diferentes técnicas, a Análise de Gases Dissolvidos (AGD) é uma das mais tradicionais, sendo estudada há muitos anos, remontando à primeira guerra mundial (Kelly, 1980).

A AGD possui diversas ferramentas que auxiliam na realização do diagnóstico, muitas delas dispostas em guias e normas baseadas em investigações estatísticas de uma extensa base de dados ao longo dos anos, como os métodos de Rogers, Doernenburg, Duval, dentre outros (Std-C57.104, 2009; Duval, 2006). Aliada a estas ferramentas,

diversos sistemas inteligentes vêm sendo desenvolvidos, comumente utilizando Redes Neurais Artificiais (RNA), Lógica Fuzzy e Máquina de Vetores de Suporte (Bakar et al., 2014; Senoussaoui et al., 2018; Zirbes et al., 2005). No entanto, tais algoritmos possuem maior custo computacional devido ao ajuste de parâmetros necessários para seu funcionamento ideal, o que por sua vez demanda tempo de acordo com o poder computacional disponível. Como alternativa, discriminantes lineares podem ser aplicados nesse tipo de tarefa de forma eficiente, seja utilizando tal classificador de forma independente (Zaini, 2019) ou combinada a outros algoritmos (Liang et al., 2008). Quando utilizado de forma independente, são construídos a partir de cálculos e operações matriciais simples, possibilitando uma implementação direta. Desta maneira, numa situação de urgência em que tempo e poder computacional não estejam a disposição, tais classificadores podem auxiliar na avaliação do equipamento ao lado de outras técnicas, por exemplo.

Neste trabalho, é realizada uma comparação de quatro classificadores, a saber: RNA, k-ésimo vizinho mais próximo, Discriminante Linear de Fisher e um discriminante baseado na métrica Mahalanobis. Assim, busca-se investigar o quão eficiente seria a aplicação destas técnicas comparadas a um classificador bem estabelecido como a RNA.

2. METODOLOGIA

Para o desenvolvimento deste estudo foi utilizado o conjunto de dados disponibilizado por Morais et al. (2004). Essa base de dados é composta por três conjuntos, dos quais dois são analisados neste trabalho: (A) disponibilizado pelo Centro de Pesquisa em Energia Elétrica (Cepel) e (B) disponibilizadas pelo *International Electrotechnical Commission* (IEC). O conjunto A é constituído pela concentração de gases dissolvidos em transformadores com e sem comutadores de carga (*On Load Tap Changer* - OLTC). O conjunto B é puramente composto por equipamentos sem comutadores. Cada amostra contém a concentração em ppm de cinco gases (H_2 , CH_4 , C_2H_2 , C_2H_4 , C_2H_6), com uma rotulação de acordo com as suas condições, a saber:

- Conjunto A (Cepel) - 83 amostras de Condições Normais (N), 80 de Falha Térmica (T) e 61 de Falha Elétrica (E);
- Conjun B (IEC) - 16 amostras de Condições Normais (N), 14 de Falha Térmica (T) e 22 de Falha Elétrica (E).

Dois experimentos foram realizados com diferentes objetivos. O primeiro experimento, buscou comparar o desempenho dos classificadores entres si. Para tanto, uni-se o Conjunto A (após pré processamento) e B separando 70% para treinamento e 30% para validação. O segundo experimento utilizou o Conjunto A para treinamento e o B para validação. Neste ponto, o intuito é analisar o desempenho dos algoritmos com relação aos métodos clássicos de Rogers e Doernenburg aplicados aos dados IEC.

2.1 Pré-Processamento

As amostras passaram por um pré processamento, etapa em que as entradas são normalizadas. Durante a implementação direta dos classificadores no conjunto de dados A , foi percebido que a presença de dois tipos de transformadores poderia influenciar nas decisões devido as suas diferenças operacionais. Dessa forma, realiza-se um procedimento de agrupamento dos dados buscando obter um conjunto mais puro. O algoritmo utilizado foi o *k-means*, para k igual 2, implementada na classe de amostras N e T , em que os dois tipos de transformadores se tornam mais evidentes. Após a separação, observou-se um grupo de dados com concentrações de hidrogênio relativamente altas quando comparadas aos demais gases, o que foi interpretado como um indicativo de arco elétrico na operação dos transformadores. Cabe ressaltar que o hidrogênio é um dos gases chave de falhas elétricas como descargas parciais e formação de arco, como relatado no Guia IEEE (Std-C57.104, 2009). Ademais, a formação de arco elétrico é comum na operação normal dos transformadores OLTC (Duval, 2008). Nas Figuras 1 e 2 é possível observar a média dos dois grupos obtidos para as classes N e T , respectivamente. Dessa forma, um novo conjunto de dados A , sem comutadores de carga, é formado, composto por 38 amostras N , 52 T e 61 E .

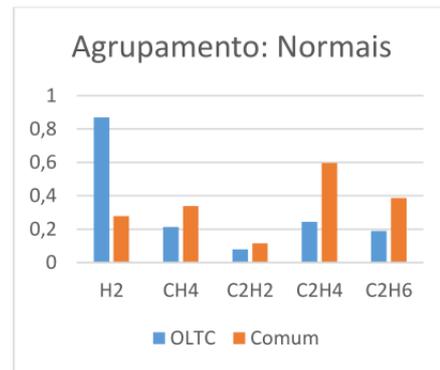


Figura 1. Média das concentrações na classe de dados N .

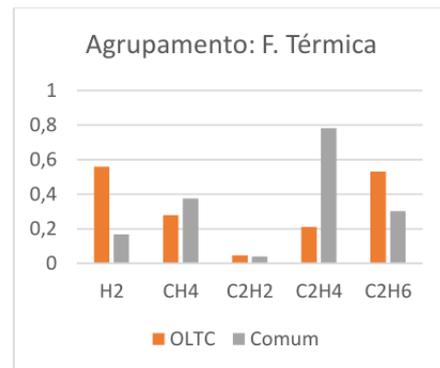


Figura 2. Média das concentrações na classe de dados T .

A partir desse ponto, inicia-se o treinamento dos modelos. No Experimento 1, uni-se o novo conjunto A e o conjunto B e se divide 70% para treinamento e o restante para validação. No Experimento 2, o conjunto A é utilizado para o treinamento e o conjunto B para validação. Os algoritmos de reconhecimento de padrões foram implementados

via *software* MATLAB. A avaliação é realizada por meio de tabelas de acurácia e de confusão dos classificadores. Ademais, uma representação gráfica dos dados em duas dimensões também é realizada utilizando Análise de Discriminante Linear (ADL) (Marcialis and Roli, 2002) para redução de dimensão.

2.2 Classificador 1

O Classificador 1 consiste numa cascata de outros dois baseados na função discriminante linear de Fisher. Tal função é uma combinação linear dos dados, buscando maximizar a separação entre duas populações (Varella, 2008). A função discriminante é dada pela equação (1), enquanto o ponto médio entre as duas populações é dado pela equação (2):

$$D(x) = [\mu_1 - \mu_2] \cdot \Sigma^{-1} \cdot x = w \cdot x \quad (1)$$

$$m = \frac{1}{2} [\mu_1 - \mu_2] \cdot \Sigma^{-1} \cdot [\mu_1 + \mu_2] \quad (2)$$

em que x é a amostra a ser classificada, μ_i é a média da classe i , Σ é a matriz de covariância das amostras. A amostra x é classificada de acordo com o valor de $D(x)$, caso esse seja maior ou igual a m , x é classificado como pertencente a população 1, caso contrário, pertence a população 2.

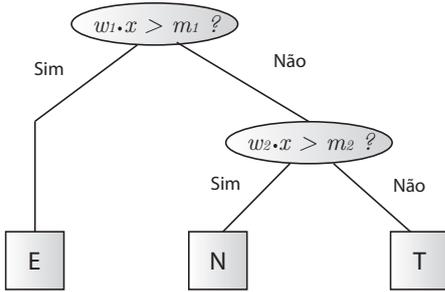


Figura 3. Cascata utilizada na classificação.

O algoritmo desenvolvido realiza a primeira separação em duas populações de amostras, uma composta por amostras E e outra por amostras N e T . Em seguida, separam-se as classes N e T . Na Figura 3 pode ser visto o fluxograma da cascata utilizada. A escolha da separação foi realizada observando a disposição dos dados, onde foi reconhecido uma maior distinção das falhas elétricas das demais. Para auxiliar a visualização, foi feita uma redução de dimensão, de 5 para 2, utilizando ADL e pode ser observada na Figura 4.

Após o treinamento, o classificador é submetido à validação, onde é possível observar seu desempenho diante de novos dados (também dispostos nas mesmas três classes, N , T e E).

2.3 Classificador 2

O segundo classificador a ser implementado é baseado na distância de Mahalanobis. Esse classificador realiza a discriminação das diferentes classes de acordo com a métrica descrita em (3):

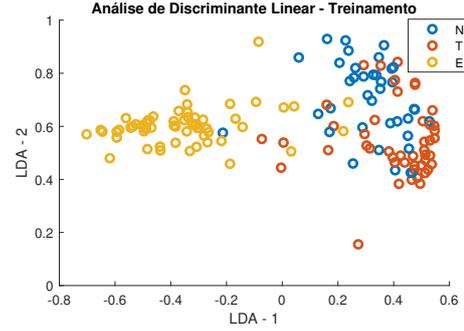


Figura 4. Análise gráfica da fase de treinamento.

$$Dm_i = \sqrt{(x - \mu_i) \Sigma^{-1} (x - \mu_i)}, \quad (3)$$

em que x é a amostra a ser classificada, μ_i é a média da classe i , Σ é a matriz de covariância das amostras. A amostra x é classificada de acordo com o menor valor de Dm_i , pertencendo ao grupo mais próximo de acordo com essa métrica. Supondo um conjunto de amostras distribuídas como uma gaussiana, esse tipo de métrica pode oferecer uma melhor delimitação dos conjuntos que a distância euclidiana, por exemplo. Adicionalmente, possui uma implementação simples e rápida.

2.4 Classificador 3

A terceira metodologia utilizada para classificação das amostras contou com um algoritmo mais robusto conhecido como Redes Neurais Artificiais. A rede neural busca emular o aprendizado a partir do modelo do neurônio (Figura 5). Assim como o neurônio humano, o modelo está sujeito a diferentes estímulos, os quais são amplificados ou atenuados de acordo com os pesos sinápticos w , inibidos ou não pelos vieses b . Esses estímulos são capazes de provocar determinado nível de excitação, representada no modelo pela função de ativação $g(\cdot)$. A partir da organização, do número de neurônios e das técnicas utilizadas para aprendizado é possível estruturar diferentes arquiteturas. A arquitetura utilizada neste trabalho utiliza uma camada escondida, conforme a Figura 6, em que $n = 5$, $k = 90$ e $m = 3$.

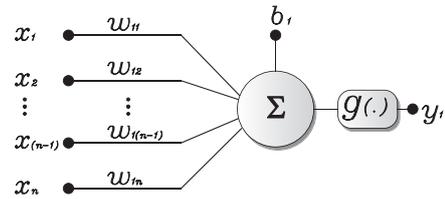


Figura 5. Modelo do neurônio artificial.

O treinamento do algoritmo foi realizado após a separação do conjunto A (mencionada na subseção anterior). A rede tem como entrada as cinco concentrações de gases e como saída um vetor de três dimensões. As amostras normais apresentam sua saída como $[1,0,0]$, as de falha elétrica $[0,1,0]$ e térmica $[0,0,1]$. Tais saídas são obtidas reconhecendo o máximo valor do vetor y .

Após a realização de diferentes testes variando o número de neurônios na camada escondida se obteve a configuração apresentada na Tabela 1. Como método de treinamento,

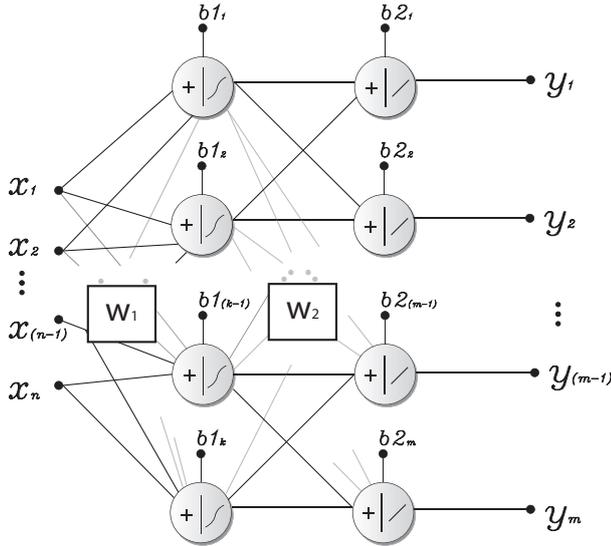


Figura 6. Arquitetura de RNA utilizada.

Tabela 1. Configuração da RNA implementada.

Número de Épocas	800.000
Número de Neurônios (camada escondida)	90
Número de Camadas	2
Função de Ativação 1	Tangente Hiperbólica
Função de Ativação 2	linear
Taxa de Aprendizado	0,001

utilizou-se o *backpropagation* (Theodoridis and Koutroumbas, 1999) com gradiente estocástico, objetivando um ajuste dos parâmetros uniforme com relação a todos os pontos. Como critério de parada foi utilizado o número de épocas (800.000) e o valor da maior atualização dos parâmetros (pesos sinápticos e vieses). Quando a maior atualização fosse inferior a $0,3 \times 10^{-15}$. A etapa de validação se deu de maneira análoga ao Classificador 1.

2.5 Classificador 4

O quarto classificador é baseado no K-ésimo Vizinho mais Próximo, do inglês *K-Nearest Neighborhood - KNN*. Esse algoritmo rotula a amostra de teste de acordo com a classe mais frequente dentre as k mais próximas, utilizando algum tipo de métrica para avaliar a distância. Nesse trabalho, o valor de k utilizado foi 10, enquanto que a métrica observada foi a distância de Mahalanobis. O número das amostras de cada classe também foi limitado ao tamanho da menor classe.

3. RESULTADOS E DISCUSSÕES

3.1 Experimento 1

Durante a construção dos classificadores, foram observados os seus desempenhos por meio da taxa de acertos na fase de treinamento e validação. Nas Tabelas 2 e 3 se encontram listados a acurácia de cada um dos classificadores com relação a cada classe e o total.

No intuito de mensurar detalhadamente o desempenho dos classificadores, foram montadas as Matrizes de Confusão de cada um, dadas pelas Tabelas 4 a 7. Considerando os

Tabela 2. Parcela de Acertos durante a fase de treinamento.

	N (%)	T (%)	E (%)	Total (%)
Clasf. 1	65,79	71,74	91,38	78,17
Clasf. 2	65,79	71,74	89,00	77,46
Clasf. 3	73,68	78,26	96,55	84,51
Clasf. 4	73,68	78,26	89,67	81,69

Tabela 3. Parcela de Acertos durante a fase de validação.

	N (%)	T (%)	E (%)	Total (%)
Clasf. 1	50,00	75,00	96,00	77,05
Clasf. 2	50,00	75,00	96,00	77,05
Clasf. 3	62,50	90,00	96,00	85,25
Clasf. 4	68,75	80,00	92,00	81,97

dados de forma binária, em condições normais ou condições de falha (T e E), é possível realizar algumas observações. Os classificadores 1, 2 e 4 tiveram 11,11% das amostras faltosas classificadas como normais (falsos positivos). Nesse tipo de problema, a classificação de um transformador defeituoso como saudável pode ter consequências graves, trazendo riscos de perda total do equipamento, risco a segurança das equipes de manutenção, etc. Ainda, os classificadores 1 e 2 obtiveram o maior número de falsos negativos, 50,00% das amostras normais classificadas como falha. Essa condição pode ser prejudicial ao sistema, acarretando intervenções desnecessárias.

Tabela 4. Matriz de Confusão do Classificador 1.

		Clasf. 1		
		N	T	E
Real	N	8	6	2
	T	5	15	0
	E	0	1	24

Tabela 5. Matriz de Confusão do Classificador 2.

		Clasf. 2		
		N	T	E
Real	N	8	6	2
	T	5	15	0
	E	0	1	24

Tabela 6. Matriz de Confusão do Classificador 3.

		Clasf. 3		
		N	T	E
Real	N	10	4	2
	T	2	18	0
	E	0	1	24

De modo geral, O ideal é que o classificador não posua ou tenha valores reduzidos de ambos, semelhante ao Classificador 3, conforme Tabela 6. A RNA obteve um

Tabela 7. Matriz de Confusão do Classificador 4.

		Clasf. 4		
		<i>N</i>	<i>T</i>	<i>E</i>
Real	<i>N</i>	11	3	2
	<i>T</i>	4	16	0
	<i>E</i>	1	1	23

menor número de falsos positivos diante dos demais, cerca de 4,44% na validação, e a melhor acurácia dentre os quatro, 85,25%. Em seguida, o *KNN* apresentou 81,97% de acurácia nesta etapa, se posicionando como o segundo colocado. O Classificador 4 apresentou o menor valor de falsos negativos, 3,12%. Cabe ressaltar que, dentre os quatro classificadores, a RNA é sem dúvidas o que exige maior esforço computacional. Tal algoritmo demanda tempo durante seu treinamento, necessitando de diversas iterações, enquanto os demais, por serem técnicas baseadas em estatísticas, são de rápida implementação, consistindo na execução de pequenas decisões. Ademais, os Classificadores 1, 2 e 4 apresentaram taxas de acerto próximas da RNA, tornando essas ferramentas atraentes em situações de pouco recurso computacional e tempo disponíveis.

3.2 Experimento 2

Ainda é possível comparar os resultados obtidos neste estudo com os de métodos clássicos disponíveis na literatura, como Rogers e Doernenburg (Std-C57.104, 2009). Morais et al. (2004) trazem no seu trabalho tabelas com a acurácia dessas técnicas perante os dados disponibilizados pela IEC. Dessa forma, realizou-se o treinamento dos quatro classificadores apenas com os dados Cepel (Tabela 8). A Tabela 9 é uma adaptação listando os resultados de ambos métodos, onde se observa o desempenho baixo com relação aos classificadores trazidos neste trabalho. Os autores também salientam quanto ao índice de falsos positivos no método Doernenburg, em torno de 13,46% de todo o conjunto, 36,84% sobre o total de erros. O método de Rogers não foi capaz de identificar certas amostras, alcançando uma parcela de 32,69% de amostras não identificadas.

Tabela 8. Parcela de Acertos durante a fase de treinamento- Dados Cepel.

	<i>N</i> (%)	<i>T</i> (%)	<i>E</i> (%)	Total (%)
Clasf. 1	52,63	71,15	91,80	74,83
Clasf. 2	50,00	71,15	88,52	72,85
Clasf. 3	63,15	78,84	95,08	81,45
Clasf. 4	63,15	82,69	88,52	80,13

Na Tabela 9 é possível notar o alto desempenho dos classificadores com relação ao treinamento. Isso pode ser explicado devido a significativa sobreposição dos dados Cepel (Figura 7), não tão evidente nos dados IEC (Figura 8). A superposição tende a dificultar a tarefa dos classificadores, principalmente se tratando de métodos lineares. Desta maneira, ao serem aplicados em um conjunto de dados onde as classes estão mais separadas, o desempenho tende a apresentar melhores resultados. Ainda neste experimento, o *KNN* se mostrou superior a RNA durante a validação, atingindo 96,15%.

Tabela 9. Parcelas de Acertos de métodos clássicos diante dos dados IEC apresentadas por Morais et al. (2004).

	<i>N</i> (%)	<i>T</i> (%)	<i>E</i> (%)	Total (%)
Rogers	—	86,36	64,28	53,84
Doern.	43,75	72,73	71,43	63,46
Clasf. 1	93,75	71,42	100	90,38
Clasf. 2	87,50	71,43	100	88,46
Clasf. 3	93,75	78,57	100	92,31
Clasf. 4	100,00	85,71	100	96,15

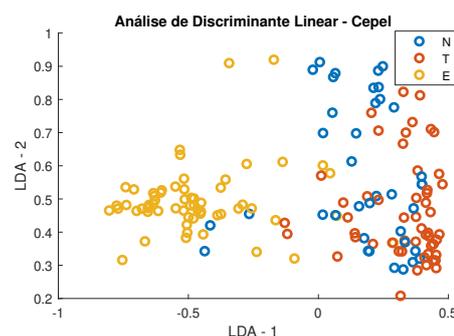


Figura 7. Análise gráfica dos dados Cepel.

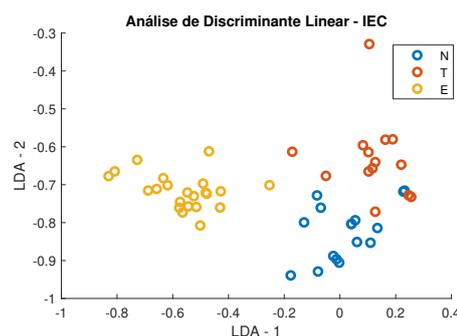


Figura 8. Análise gráfica dos dados IEC.

3.3 Resultados: Pré-Processamento

Na Tabela 10 se encontram listados os desempenhos de cada classificador antes e após separação dos dados, a qual é realizada no pré-processamento. Todos apresentaram um melhor desempenho quando treinados com o novo conjunto de dados, evidenciando o benefício da metodologia. Ainda é possível observar que a RNA e o *KNN* tiveram as alterações mais significativas, saltando de 75,00 para 92,11% e de 82,69 para 96,15% respectivamente. Essa melhoria era esperada, dado que as possíveis amostras de transformadores com OLTCs possuem um perfil diferente daqueles que não os possui.

Tabela 10. Acurácia dos classificadores antes e depois da separação de dados.

Clasf.	1 (%)	2 (%)	3 (%)	4 (%)
Antes	86,54	82,69	75,00	82,69
Depois	90,38	88,46	92,11	96,15

4. CONCLUSÃO

A AGD é uma das técnicas para diagnóstico de falhas incipientes mais estabelecida ao longo dos anos. Neste trabalho, foram desenvolvidos e comparados quatro classificadores de falhas de transformadores a partir desse tipo de análise.

O Classificador 1, o qual consiste em dois discriminantes lineares de Fisher em cascata, obteve 78,17% e 77,05% de acertos durante o treinamento e validação, respectivamente. O segundo classificador, baseado na distância de Mahalanobis, alcançou taxas de acerto de 77,46% e 77,05%, treinamento e validação. Ambos apresentaram o maior número de falsos positivos, equipamentos defeituosos classificados como normais, conforme Tabelas 4 e 5. Nesta aplicação é desejável que ocorra o menor número possível desse tipo de erro devido aos riscos associados a não detecção da falha do transformador.

A RNA, Classificador 3, obteve uma acurácia de 84,51% e 85,25%, treinamento e validação respectivamente. Ademais, observou-se valores reduzidos de falsos positivos e negativos, conforme Tabela 6. Dessa maneira, é possível verificar a capacidade desse algoritmo na identificação de falhas na AGD.

O quarto classificador, *KNN*, apresentou valores próximos a RNA, 81,51% e 81,97%, treinamento e validação.

Ainda foi observado que todos os classificadores obtiveram resultados superiores à aplicação de técnicas clássicas no conjunto de dados *B* (IEC) de acordo com o apresentado por Moraes et al. (2004). Nesta etapa, o Classificador 4 apresentou uma acurácia superior até mesmo a RNA.

Como pode ser notado, todos os classificadores obtiveram bons desempenhos, dentre os quais a RNA e o *KNN* se destacaram. Os três classificadores apresentados são de rápida implementação, podendo ser utilizados como uma alternativa com pouco custo computacional, com suas restrições.

AGRADECIMENTOS

Os autores deste trabalho agradecem a CAPES e a Universidade Federal de Sergipe pelo apoio financeiro e estrutural recebido na realização deste estudo.

REFERÊNCIAS

- Abu-Elanien, A.E. and Salama, M. (2010). Asset management techniques for transformers. *Electric power systems research*, 80(4), 456–464.
- Bakar, N.A., Abu-Siada, A., and Islam, S. (2014). A review of dissolved gas analysis measurement and interpretation techniques. *IEEE Electrical Insulation Magazine*, 30(3), 39–49.
- Duval, M. (2006). Dissolved gas analysis and the duval triangle. *TechCon Asia Pacific, Sydney, Australia*, 1–20.
- Duval, M. (2008). The duval triangle for load tap changers, non-mineral oils and low temperature faults in transformers. *IEEE Electrical Insulation Magazine*, 24(6), 22–29.
- Kelly, J.J. (1980). Transformer fault diagnosis by dissolved-gas analysis. *IEEE Transactions on Industry Applications*, (6), 777–782.
- Liang, Y., Sun, X., Liu, Q., Bian, J., and Li, Y. (2008). Fault diagnosis model of power transformer based on combinatorial k-fda. In *2008 International Conference on Condition Monitoring and Diagnosis*, 956–959. IEEE.
- Marcialis, G.L. and Roli, F. (2002). Fusion of lda and pca for face verification. In *International Workshop on Biometric Authentication*, 30–37. Springer.
- Moraes, D.R. et al. (2004). Ferramenta inteligente para detecção de falhas incipientes em transformadores baseada na análise de gases dissolvidos no óleo isolante.
- Senoussaoui, M.E.A., Brahami, M., and Fofana, I. (2018). Combining and comparing various machine-learning algorithms to improve dissolved gas analysis interpretation. *IET Generation, Transmission & Distribution*, 12(15), 3673–3679.
- Std-C57.104, I. (2009). Ieee guide for the interpretation of gases generated in oil-immersed transformers. *IEEE Std C57.104-2009*. doi:10.1109/IEEESTD.2009.4776518.
- Theodoridis, S. and Koutroumbas, K. (1999). Pattern recognition and neural networks. In *Advanced Course on Artificial Intelligence*, 169–195. Springer.
- Varella, C.A.A. (2008). Análise multivariada aplicada as ciencias agrárias.
- Zaini, H.G. (2019). Diagnosis of the power transformer faults based on dga using intelligent classifier. *International Journal of Engineering Research and Technology*, 12(8), 1964–1971.
- Zirbes, R., Rolim, J.G., and Zürn, H.H. (2005). Metodologias para avaliação e diagnóstico do estado de isolamentos de papel impregnado com óleo mineral. *Sba: Controle & Automação Sociedade Brasileira de Automação*, 16(3), 318–331.