

Uma análise heurística baseada em protótipos para o monitoramento de estados de motores de indução

Diego P. Sousa

Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará
Fortaleza, Brasil
diegoperdigao@gmail.com

Guilherme A. Barreto

Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará
Fortaleza, Brasil
gbarreto@ufc.br

Charles C. Casimiro

Departamento de Engenharia de Teleinformática
Universidade Federal do Ceará
Fortaleza, Brasil
charles@ufc.br

Cláudio M. S. Medeiros

Departamento de Indústria
Instituto Federal de Educação, Ciência e Tecnologia do Ceará
Fortaleza, Brasil
claudiosa@ifce.edu.br

Resumo—Neste artigo, nós propomos uma metodologia de projeto almejando a definição da quantidade adequada de protótipos e suas respectivas localizações iniciais confiáveis para a conhecida família de classificadores LVQ com o objetivo da elevação das taxas de acurácia em tarefas de detecção/classificação. Além dos exemplos de funcionamento normal de um motor trifásico, uma bancada de testes é construída para a geração de exemplos de curto-circuito de diferentes níveis de impedância. Para evitar a perda do motor (em decorrência de altas correntes de curto-circuito), os níveis de corrente são limitados através da utilização de resistores. Os sinais de corrente analisados são de difícil classificação entre normal e faltoso em natureza da fácil interpretação incorreta de falhas de alta impedância como amostras não faltosas. Com o objetivo da obtenção do número adequado de protótipos rotulados e suas corretas iniciações que reduzam erros de classificação, nós utilizamos *K-means* e técnicas de validação de agrupamentos. Através de simulações computacionais abrangentes, nós mostrados que a metodologia proposta é capaz de eventualmente gerar altas taxas de desempenho na aplicação de engenharia citada acima quando aplicada aos vários classificadores LVQ considerados. Especificamente, o GLVQ (*Generalized Learning Vector Quantization*) apresentou o melhor desempenho máximo de 96,6% quando utilizado em conjunto com os índices *Dunn* e *Silhouettes* como os mais eficientes na determinação da quantidade ideal de protótipos.

Keywords—motores de indução; detecção de falhas de curto-circuito; família de classificadores LVQ; técnicas de validação de agrupamentos

I. INTRODUÇÃO

Como detalhadamente discutido em [1], as pesquisas em aprendizado de máquina têm por objetivo principal o desenvolvimento de ferramentas computacionais eficientes capazes de fornecer soluções satisfatórias, senão as melhores, a problemas complexos de classificação de padrões, tais como os de natureza não linear, incerta e de alta dimensionalidade. Entre as propriedades almejadas, pode-se exemplificar a capacidade de construção de fronteiras de decisão arbitrárias onde se tenha o mínimo, ou até mesmo nenhuma, superposição acerca dos

dados disponíveis e, também, a capacidade de generalização do conhecimento para novos dados.

Considerando a motivação acima, o presente trabalho apresenta um estudo direcionado ao monitoramento de estados de motores de indução, pois, conforme justificado nos trabalhos [2] e [3], apesar de todos os estudos sobre esse tipo de máquina e a sua robustez comprovada, o motor de indução é, infelizmente, sujeito à diferentes tipos de falhas que são causadas devido ao envelhecimento da máquina, condições adversas do ambiente, aplicações inadequadas, ou falta de um programa de manutenção. Portanto, estudos sobre o monitoramento das diversas fontes faltosas que ocorrem nestes equipamentos são efetivamente indispensáveis.

Dentre os diversos tipos de falhas que estes equipamentos estão sujeitos, os autores [4] e [5] expõem a relevância daquelas ocasionadas por curto-circuito das espiras presentes nos enrolamentos do estator.

Diante desta problemática, uma bancada de testes fora construída no Laboratório de Processamento de Energia, do Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), e, em parceria com a Universidade Federal do Ceará (UFC), diversos estudos foram realizados em busca da modelagem mais eficiente com o objetivo da produção de um produto embarcado em um futuro próximo. Por exemplo, os trabalhos [6] e [7] investigaram soluções a partir de classificadores mais conhecidos (ex: MLP e SVM) e outras ferramentas recentes como a rede MLM. Há, também, estudos a partir de técnicas baseadas em rede SOM como os trabalhos [8] e [9]. Por fim, também pode-se listar o estudo [10] como o primeiro a partir de algoritmos baseados em LVQ.

Motivado pela busca de soluções que demandem menor custo computacional, o presente trabalho apresenta um estudo sobre a influência da escolha de métricas de validação de agrupamentos no procedimento de definição da quantidade de protótipos considerada adequada a ser utilizada em algoritmos baseados em protótipos.

As próximas seções estão organizadas da seguinte forma: na seção 2, as técnicas de validação de agrupamentos utilizadas são apresentadas; na seção 3, os classificadores utilizados da família LVQ são apresentados; na seção 4, descreve-se o procedimento realizado para a geração dos dados; já a metodologia proposta é detalhada e os resultados obtidos são discutidos na seção 5; por fim, as conclusões e trabalhos futuros são descritos na seção 6.

II. TÉCNICAS DE VALIDAÇÃO DE AGRUPAMENTOS

A diversas técnicas para validação de agrupamentos são ferramentas utilizadas *a posteriori* com o propósito de avaliar os resultados obtidos por algoritmos de *clustering*. Contudo, deve-se considerar que cada técnica possui seu próprio conjunto de suposições para interpretar a coesão e separabilidade de conjuntos e, portanto, os resultados finais podem divergir entre as técnicas escolhidas.

Nesta seção, são descritos alguns dos índices de validação de agrupamentos mais conhecidos na literatura. Na exposição destes, nós representados o número de agrupamentos por K e K_{opt} como o número ótimo de agrupamentos para representar uma coleção de dados de entrada. É importante enfatizar que estas métricas são nulas para o agrupamento trivial e, portanto, devem ser computadas apenas quando cada agrupamento contiver um número razoável (ou seja, maior do que 1) de padrões de entrada.

1) *Índice Calinski-Harabasz (CH)*: O índice CH, proposto por [11], é representado pela seguinte expressão:

$$CH(K) = \frac{\text{trace}(\mathbf{B})_K / (K - 1)}{\text{trace}(\mathbf{W})_K / (N - K)}, \quad (1)$$

em que $\mathbf{B}_K = \sum_{l=1}^K \#C^l (\mathbf{x}^l - \mathbf{x})(\mathbf{x}^l - \mathbf{x})'$ é a soma dos quadrados entre *clusters*, N é total de amostras, K é o número de *clusters*, $\mathbf{W}_K = \sum_{l=1}^K \sum_{i=1}^{\#C^l - 1} \sum_{j=i+1}^{\#C^l} (\mathbf{x}i^l - \mathbf{x}j^l)(\mathbf{x}i^l - \mathbf{x}j^l)'$ é a soma dos quadrados *intracluster*, $\mathbf{x}i^l$ e $\mathbf{x}j^l$ são, respectivamente, os i -ésimos e j -ésimos itens do *cluster* C^l , \mathbf{x}^l é o centróide do l -ésimo *cluster*, $\#C^l$ é a cardinalidade do l -ésimo *cluster* e \mathbf{x} é o centróide do banco de dados inteiro.

A escolha da quantidade ótima de *clusters* K_{opt} é dada pelo argumento que maximiza o índice CH.

2) *Família de Índices Dunn*: A família de índices Dunn, proposta por [12], é representada genericamente pela seguinte expressão:

$$D(K) = \frac{\min_{i \neq j} \{\delta(S_i, S_j)\}}{\max_{1 \leq l \leq k} \{\Delta(S_l)\}}, \quad (2)$$

dado que $\delta(S_i, S_j)$ denota uma função de dissimilaridade (ex: distância euclidiana) entre os agrupamentos S_i e S_j , e $\Delta(S_i)$ é uma medida da dispersão dos dados dentro do agrupamento S_i .

No índice Dunn original, $\delta(S_i, S_j)$ é definido pela equação:

$$\delta(S_i, S_j) = \min_{\mathbf{x} \in S_i, \mathbf{y} \in S_j} \{d(\mathbf{x}, \mathbf{y})\}, \quad (3)$$

e $\Delta(S_i)$ é dada pela equação:

$$\Delta(S_i) = \max_{\mathbf{x}, \mathbf{y} \in S_i} \{d(\mathbf{x}, \mathbf{y})\}. \quad (4)$$

O ponto máximo no gráfico de $D(K)$ contra K pode ser uma indicação do número de agrupamentos que mais se ajusta aos dados fornecidos.

De acordo com [13], o índice Dunn apresenta bons resultados na identificação de agrupamentos compactos e bem separados. As principais limitações do índice Dunn original são a sua complexidade e sensibilidade ao ruído.

3) *Índice Davies-Bouldin (DB)*: O índice Davies-Bouldin (DB), proposto por [14], é formulado a partir da razão entre a soma da dispersão dentro dos agrupamentos e a dispersão entre os agrupamentos.

No índice DB, a métrica $R_{i,j}$ representa uma medida de similaridade entre dois agrupamentos S_i e S_j formulada através da seguinte equação:

$$R_{j,k} = \frac{e_{S_j} + e_{S_k}}{d(S_j, S_k)}, \quad (5)$$

nesta medida, e_{S_j} e e_{S_k} são, respectivamente, os erros médios para os agrupamentos S_j e S_k , e $d(S_j, S_k)$ denota a distância euclidiana entre os protótipos dos dois agrupamentos.

Seja o índice para o k -ésimo agrupamento dado indicado por:

$$R_k = \max_{j \neq k} \{R_{j,k}\}. \quad (6)$$

O índice de DB é representado pela seguinte expressão:

$$DB(K) = \frac{1}{K} \sum_{k=1}^K R_k. \quad (7)$$

O melhor agrupamento é definido através do menor valor do índice DB para uma sequência de valores calculados para diferentes quantidades de agrupamentos.

De acordo com [16], o índice DB é mais indicado para dados que se organizam em agrupamentos hiperesféricos, não sendo apropriado para agrupamentos de formas arbitrárias.

4) *Silhouettes*: O índice *silhouettes*, proposto por [17], é calculado para cada padrão de entrada. Esta métrica quantifica a validade dos agrupamentos com base na proximidade entre os padrões de um agrupamento e na distância dos padrões de um agrupamento ao agrupamento mais próximo.

Formalmente, denota-se por $a(x_i)$ a dissimilaridade média do padrão x_i em relação aos demais padrões de seu respectivo agrupamento. Já a variável $b(x_i)$ representa a menor dissimilaridade média do i -ésimo vetor de características à qualquer outro agrupamento a qual ele não pertença.

Assim, a silhueta de um dado padrão, $\delta(x_i)$, empregando uma medida de dissimilaridade, é dada por:

$$\delta(x_i) = \begin{cases} 1 - a(x_i)/b(x_i), & a(x_i) < b(x_i) \\ 0, & a(x_i) = b(x_i) \\ b(x_i)/a(x_i) - 1, & a(x_i) > b(x_i) \end{cases} \quad (8)$$

Então, o índice *silhouettes* $S(K)$ é representado pela seguinte expressão:

$$S(K) = \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_i). \quad (9)$$

O critério *silhouettes* visa indicar quais padrões estão bem situados dentro dos seus agrupamentos e quais estão fora de um agrupamento apropriado.

Este método de validação pode ser aplicado tanto a agrupamentos de formatos hiperesféricos, quanto a agrupamentos de formas arbitrárias. A principal limitação deste método está no elevado custo computacional.

III. CLASSIFICADORES LVQ

Seja um conjunto de padrões de entrada e saída de treinamento dado por $\{(\mathbf{x}_l, y_l)\}_{l=1}^N$, onde $\mathbf{x}_l \in \mathbb{R}^p$ representa o l -ésimo padrão de entrada e $y_l \in \mathcal{C}$ representa sua classe correspondente onde deve assumir apenas um valor de M do conjunto finito $\mathcal{C} = \{c_1, c_2, \dots, c_M\}$.

Dado um conjunto de vetores rotulados $\mathbf{m}_i \in \mathbb{R}^p$, $i = 1, \dots, K$, a associação ao novo padrão de entrada $\mathbf{x}(t)$ é formulada a partir do seguinte critério de decisão:

$$\begin{aligned} \text{Classe de } \mathbf{x}(t) &= \text{Classe de } \mathbf{m}_c(t), \\ \text{onde } c &= \arg \min_{i=1, \dots, K} d(\mathbf{x}(t), \mathbf{m}_i(t)), \end{aligned} \quad (10)$$

em que $d(\cdot, \cdot)$ representa a medida de dissimilaridade específica à variante LVQ considerada e que c é o índice do protótipo mais próximo ao padrão de entrada entre os K protótipos disponíveis.

Para todos os classificadores baseados em LVQ, tem-se que o número de protótipos (K) é maior que o número de classes (M). Consequentemente, diferentes protótipos dividirem uma mesma classe é uma característica comum nesta família de classificadores.

A. LVQ1 [18]

Neste algoritmo, o índice c é definido como mostrado na Eq. (10) para um novo padrão de entrada $\mathbf{x}(t)$. Posteriormente, este protótipo \mathbf{m}_c é atualizado como

$$\mathbf{m}_c(t+1) = \mathbf{m}_c(t) + s(t)\alpha(t)[\mathbf{x}(t) - \mathbf{m}_c(t)], \quad (11)$$

dado que $s(t) = +1$ para caso de classificação correta ou $s(t) = -1$ caso esteja incorreta, e α é a taxa de aprendizagem.

B. LVQ2.1 [18]

No algoritmo LVQ2.1, os protótipos \mathbf{m}_i e \mathbf{m}_j são os vizinhos mais próximos ao padrão $\mathbf{x}(t)$ e estes são atualizados simultaneamente. Um deles (utiliza-se comumente \mathbf{m}_i) deve pertencer à classe correta e o outro à classe incorreta, respectivamente. Então, as regras de aprendizagem para esta variante LVQ são modeladas como

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + \alpha(t)[\mathbf{x}(t) - \mathbf{m}_i(t)], \quad (12)$$

$$\mathbf{m}_j(t+1) = \mathbf{m}_j(t) - \alpha(t)[\mathbf{x}(t) - \mathbf{m}_j(t)], \quad (13)$$

caso $\mathbf{x}(t)$ satisfaça a seguinte restrição:

$$\min \left(\frac{d_i}{d_j}, \frac{d_j}{d_i} \right) > s, \quad \text{where } s = \frac{1-w}{1+w}, \quad (14)$$

em que d_i e d_j são as distâncias euclidianas de $\mathbf{x}(t)$ para \mathbf{m}_i e \mathbf{m}_j , respectivamente. É recomendado atribuir valores entre 0,2 e 0,3 para a variável w que representa uma "janela relativa".

C. LVQ3 [18]

O algoritmo LVQ3 é uma extensão da versão LVQ2.1 que, diferentemente do algoritmo anterior, apresenta uma abordagem que também considera cenários em que os protótipos \mathbf{m}_i e \mathbf{m}_j possuem o mesmo rótulo. Aplica-se a seguinte regra de atualização:

$$\mathbf{m}_k(t+1) = \mathbf{m}_k(t) + \epsilon s(t)\alpha(t)[\mathbf{x}(t) - \mathbf{m}_k(t)], \quad (15)$$

para $k \in \{i, j\}$, condicionado ao \mathbf{x} respeitar a restrição da "janela". Através de seus experimentos, [18] reporta que é adequado utilizar valores de ϵ entre 0,1 e 0,5 enquanto w deve ser utilizado considerando o intervalo entre 0,2 e 0,3.

D. Generalized LVQ [19]

O algoritmo GLVQ é a primeira variante LVQ a não ser categorizada como heurística. A função custo deste algoritmo é dada por:

$$E_{GLVQ} = \sum_{i=1}^N \phi(\mu(\mathbf{x})), \quad \mu(\mathbf{x}) = \frac{d^+ - d^-}{d^+ + d^-}, \quad (16)$$

visto que $\phi(\cdot)$ é a função identidade para o GLVQ linear ou é a função logística para o GLVQ logístico, a variável μ é denominada distância relativa, e $d^+ = d(\mathbf{x}, \mathbf{m}^+)$ é a distância euclidiana quadrática entre o padrão de entrada $\mathbf{x}(t)$ e seu respectivo protótipo mais próximo $\mathbf{m}^+(t)$ de mesmo rótulo, e $d^- = d(\mathbf{x}, \mathbf{m}^-)$ representa a distância euclidiana quadrática entre o padrão de entrada $\mathbf{x}(t)$ e seu respectivo protótipo $\mathbf{m}^-(t)$ de rótulo distinto. Considerando estas variáveis, são aplicadas as seguintes regras de atualização:

$$\begin{aligned} \mathbf{m}^+(t+1) &= \mathbf{m}^+(t) + \\ &\alpha(t)\phi'(\mu(\mathbf{x}))[4d^-/(d^+ + d^-)^2][\mathbf{x}(t) - \mathbf{m}^+(t)], \end{aligned} \quad (17)$$

$$\begin{aligned} \mathbf{m}^-(t+1) &= \mathbf{m}^-(t) - \\ &\alpha(t)\phi'(\mu(\mathbf{x}))[4d^+/(d^+ + d^-)^2][\mathbf{x}(t) - \mathbf{m}^-(t)]. \end{aligned} \quad (18)$$

E. Generalized Relevance LVQ [20], [21]

Já para o algoritmo GRLVQ, a função custo é definida como:

$$E_{GRLVQ} = \sum_{i=1}^N \phi(\mu(\mathbf{x})), \quad \mu(\mathbf{x}) = \frac{d_{\lambda}^+ - d_{\lambda}^-}{d_{\lambda}^+ + d_{\lambda}^-}, \quad (19)$$

em que $\phi(\cdot)$ denota a função logística, $d_{\lambda}^+ = d(\mathbf{x}, \mathbf{m}^+)$ representa a distância euclidiana ponderada quadrática entre o padrão de entrada $\mathbf{x}(t)$ e seu respectivo protótipo mais próximo $\mathbf{m}^+(t)$ que possui mesmo rótulo, $d_{\lambda}^- = d(\mathbf{x}, \mathbf{m}^-)$ representa a distância euclidiana ponderada quadrática entre o padrão de entrada $\mathbf{x}(t)$ e seu respectivo protótipo mais próximo $\mathbf{m}^-(t)$ de rótulo distinto. Por fim, o vetor de ponderações pode ser determinado através do método gradiente descendente como

$$\boldsymbol{\lambda}(t+1) = \boldsymbol{\lambda}(t) - \epsilon_{\lambda}(t)\phi' \left(\frac{(\mathbf{x} - \mathbf{m}^+)^2 d_{\lambda}^- - (\mathbf{x} - \mathbf{m}^-)^2 d_{\lambda}^+}{(d_{\lambda}^+ + d_{\lambda}^-)^2} \right), \quad (20)$$

onde ϵ_{λ} é o fator de ganho e que $\lambda_a \geq 0$ para $a = 1, \dots, p$. Instabilidades numéricas são evitadas ao utilizar a

normalização $\sum_{a=1}^p \lambda_a = 1$. Por fim, as regras de atualização são mostradas abaixo:

$$\mathbf{m}^+(t+1) = \mathbf{m}^+(t) + \alpha(t)\phi'(\mu(\mathbf{x})) [4d_{\lambda}^- / (d_{\lambda}^+ + d_{\lambda}^-)^2] [\mathbf{x}(t) - \mathbf{m}^+(t)], \quad (21)$$

$$\mathbf{m}^-(t+1) = \mathbf{m}^-(t) - \alpha(t)\phi'(\mu(\mathbf{x})) [4d_{\lambda}^+ / (d_{\lambda}^+ + d_{\lambda}^-)^2] [\mathbf{x}(t) - \mathbf{m}^-(t)]. \quad (22)$$

IV. BANCO DE DADOS UTILIZADO

Neste trabalho, utilizou-se um motor de indução trifásico fabricado pela WEG¹. Suas principais características são 0,75 kW (potência), 220/380 V (voltagem nominal), 3,02/1,75 A (corrente nominal), 79,5% (eficiência), 1720 rpm (velocidade de rotação nominal), $I_p/I_n = 7,2$ (pico de razão de corrente nominal), e 0,82 (fator de potência). O conjunto de dados foi gerado considerando diferentes condições de operação que serão brevemente descritas a seguir. A bancada de testes montada é mostrada na Figura 1.

Primeiramente, um sistema de frenagem de Foucault é utilizado para aplicar 3 níveis de carga: 0% (sem carga), 50% de carga nominal e 100% (carga total). Para variar a velocidade do motor, utilizou-se um inversor de frequências de modelo WEG CFW-09 com 7 diferentes frequências: 30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz e 60 Hz.

Além disso, o motor fora exposto à 3 níveis de curto-circuito. No nível mais ameno (nível 1), 5 espiras entraram em curto-circuitado onde representam 1,41% do total de espiras de uma fase. Já no nível intermediário (nível 2), 17 espiras (4,8% do total de espiras) entraram em curto-circuito. Por fim, no nível mais severo (nível 3), 32 espiras (9,26%) foram colocadas em curto-circuito.

Finalmente, foram considerados dois tipos de curto-circuito: alta-impedância (com o objetivo de emular o estágio inicial de um curto-circuito com corrente atenuada) e baixa impedância. Ambos tipos de falhas foram controladas por resistores para evitar que o motor fosse danificado permanentemente.

Logo, combinando os dois tipos de curto-circuito e os três níveis de falha, são obtidos seis cenários faltosos. Então, o conjunto de dados é formado por um exemplo de cada condição de operação considerada: 3 fases do motor; 3 níveis de carga; 7 velocidades do inversor; 7 condições de operação do motor (normal mais 6 diferentes tipos de anomalias). Consequentemente, o banco de dados é formado por 441 amostras (147 amostras por fase).

Dado que o motor fora conectado em delta, duas correntes de linha foram diretamente conectadas à fase do motor em curto. Como o objetivo final é o desenvolvimento de um sistema de monitoramento em uma única fase, apenas uma entre essas duas são utilizadas almejando a redução de redundância. Assim, foram utilizadas 294 amostras: 147 da primeira fase (conectada diretamente à corrente de falha) e 147 da terceira fase (indiretamente conectada à corrente de falha).

A partir destas informações, pode-se inferir que a tarefa de projeto em interesse pode ser abordada por 3 maneiras distintas: (i) classificação em 7 classes (normal, alta impedância 1,

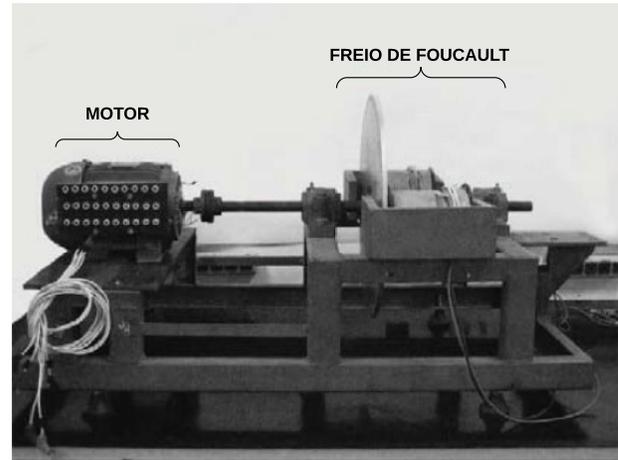


Figura 1: Bancada de testes utilizada.

alta impedância 2, alta impedância 3, baixa impedância 1, baixa impedância 2 e baixa impedância 3); (ii) classificação em 3 classes ao mesclar os 3 níveis de severidade em cada tipo de falha (normal, alta impedância e baixa impedância); e (iii) classificação binária ao mesclar todas as diferentes faltas em uma única classe faltosa (normal e falha).

Para este banco de dados, cada amostra representa um sinal de corrente coletado que possui 100.000 pontos que são obtidos através de 10 segundos de aquisição com 10 kHz de frequência de amostragem. Utilizou-se a FFT (do inglês, *Fast Fourier Transform*) como técnica de extração de atributos e aplicou-se o procedimento descrito abaixo:

- Etapa 1** - Seleção da condição de carga do motor;
- Etapa 2** - Seleção da frequência fundamental (f_c) do inversor de frequências;
- Etapa 3** - Leitura do sinal de corrente por 10 segundos utilizando taxa de amostragem de 10 KHz;
- Etapa 4** - Aplicar FFT ao sinal de corrente e calcular o módulo dos números complexos coletados;
- Etapa 5** - Buscar a frequência correspondente ao máximo valor do espectro calculado. Representar-lo como \hat{f}_c , dado que é um valor estimado de f_c (ver Etapa 2).
- Etapa 6** - Construir o vetor de características correspondente aos da saída da FFT para as seguintes harmônicas de \hat{f}_c : $\{0.5\hat{f}_c, 1.5\hat{f}_c, 2.5\hat{f}_c, 3\hat{f}_c, 5\hat{f}_c \text{ e } 7\hat{f}_c\}$;
- Etapa 7** - Para retirar o viés da frequência fundamental selecionada na Etapa 2, normalizar o vetor de atributos obtido na etapa anterior pela harmônica de \hat{f}_c .

Para concluir, este conjunto de dados é composto por 294 vetores de características rotulados e cada vetor é formado 6 atributos onde representam os valores da FFT de 6 harmônicas da frequência fundamental. Entre as 3 formas distintas de rotulação, fora adotada a abordagem de classificação ternária no desenvolvimento deste trabalho.

¹<http://www.weg.net/institucional/BR/en/>

V. RESULTADOS E DISCUSSÕES

Nesta seção, nós apresentamos e avaliamos a metodologia proposta para encontrar a melhor métrica de validação de agrupamentos que forneça a quantidade de protótipos, assim como suas posições iniciais, para as 3 diferentes classes existentes no conjunto de dados, que são representadas pelos rótulos N (normal), A (alta impedância) e B (baixa impedância).

Os resultados apresentados nesta seção foram realizados considerando o tratamento do problema em 3 classes a fim da obtenção da quantidade adequada de protótipos de cada classe. Porém, para melhor avaliar os desempenhos encontrados, as taxas de classificação ternária se encontram mapeadas para classificação binária visto que a aplicação de interesse é o correto diagnóstico do motor entre os estados normal e falha.

Para cada classificador LVQ considerado, foram realizadas 10 rodadas independentes do método utilizado. A cada rodada, executa-se a seguinte metodologia proposta dividida em 4 etapas: (i) divisão aleatória dos dados entre treino (80%) e teste (20%) a ser utilizada em todos classificadores; (ii) para cada classe, realiza-se a determinação do K_{opt} e a respectiva iniciação dos protótipos através da aplicação de *clustering* e técnicas de validação de agrupamentos; por fim, (iii e iv) treinamento e teste de cada variante LVQ. Ao final de cada realização, armazenam-se as taxas de classificação obtidas.

A segunda etapa da metodologia é realizada em dois passos. No primeiro passo, nós aplicamos o algoritmo *K-means* em cada classe individualmente considerando o intervalo $K = 2, 3, \dots, K_{max} = 10$ protótipos. Nós executamos 10 realizações do *K-means* para cada valor de K presente no intervalo e selecionamos o conjunto $\{\mathbf{p}_j\}_{j=2}^K$ que gerou o menor valor de MSE_{QQ}^2 de cada classe. Utilizando estes conjuntos selecionados, nós calculamos os correspondentes valores de cada métrica de validação de agrupamentos com o objetivo de selecionar o número ótimo de protótipos por classe. Almejando encontrar a melhor métrica, utilizamos cada técnica de validação separadamente assim como o voto majoritário destes. Já na segunda etapa, nós iniciamos os classificadores LVQ utilizando os K_{opt} protótipos selecionados.

Inicialmente, o desempenho da abordagem de votação majoritária fora aferida com o objetivo de verificação do rendimento da classificação de cada classificador LVQ. Os desempenhos encontrados são mostrados na Tabela I. Através da análise desta tabela, verifica-se que apenas o LVQ2.1 não obteve os melhores resultados de máxima acurácia. Para a média, o GLVQ logístico apresentou o melhor resultado. Enquanto para a mediana, os melhores algoritmos foram o GLVQ linear, GLVQ logístico e GRLVQ. Quanto às taxas mínimas, os melhores algoritmos foram as duas versões do GLVQ. Já para o desvio padrão, o algoritmo mais adequado é o GLVQ linear.

Além disso, verifica-se que o algoritmo LVQ1 apresentou desvio padrão bastante acima aos encontrados pelos demais e que seu valor mínimo de acurácia fora demasiadamente

indesejável. Portanto, este fora o algoritmo de pior desempenho. Considerando as demais variantes LVQ, o GLVQ linear não apresentou as melhores estatísticas somente para a média (perdendo para sua versão logística que possui maior custo computacional).

Após a definição do algoritmo GLVQ linear como o vencedor dentre os modelos LVQ avaliados, comparou-se o desempenho de cada métrica de validação de agrupamentos quando aplicadas individualmente na metodologia de treinamento proposta.

Os resultados das diferentes técnicas de validação são mostrados na Tabela II. Verifica-se que as estatísticas média e mediana encontradas são as mesmas em todos os casos considerados. Além disso, verifica-se que, excetuando-se o valor máximo, o voto majoritário apresentou as melhores estatísticas. Também a partir da análise desta tabela, observa-se que as métricas CH e DB obtiveram as melhores estatísticas apenas para os casos em que todos empataram. Já as técnicas Dunn e *Silhouettes* forneceram as melhores taxas máximas. Consequentemente, as decisões por voto majoritário, Dunn e *Silhouettes* demandam análise mais rígida.

Considerando cenários de recursos não escassos, dada a grande diversidade de técnicas não consideradas neste tra-

	LVQ1	LVQ2.1	LVQ3	GLVQ linear	GLVQ logístico	GRLVQ
máxima	94,9153	93,2203	94,9153	94,9153	94,9153	94,9153
média	83,4237	85,7288	85,8644	88,3898	88,4746	87,7288
mediana	84,7458	85,5932	86,4407	88,1356	88,1356	88,1356
mínima	8,4746	74,5763	74,5763	79,6610	79,6610	77,9661
std	13,2576	3,9230	3,7338	3,1073	3,2410	3,3478

Tabela I: Estatísticas de desempenho obtidas através do voto majoritário.

	Voto majoritário	CH	DB	Dunn	Silhouettes
máxima	94,9153	94,9153	94,9153	96,6102	96,6102
média	88,3898	87,7797	88,2712	88,2712	88,3729
mediana	88,1356	88,1356	88,1356	88,1356	88,1356
mínima	79,6610	77,96610	79,6610	77,96610	79,6610
std	3,1073	3,2077	3,2292	3,3178	3,1871

Tabela II: Estatísticas de desempenho do GLVQ linear utilizando as diferentes técnicas de validação abordadas.

		Mínimo	Mediano	Máximo	Desvio padrão
N	Voto maj.	2	2	3	0,2564
	Silhouettes	2	2	4	0,3208
A	Voto maj.	2	3	4	0,4416
	Silhouettes	2	3	4	0,4462
B	Voto maj.	2	4	10	0,9860
	Silhouettes	3	4	4	0,4648

Tabela III: Distribuição dos valores de K_{opt} por classe através do voto majoritário e de *silhouettes*.

K_{opt}		2	3	4	5	...	8	9	10
Frequência	Voto maj.	3	29	66	0	...	0	1	1
	Silhouettes	0	31	69	0	...	0	0	0

Tabela IV: Frequência das sugestões de K_{opt} para baixa impedância.

²Média da soma dos erros quadráticos de quantização: $MSE_{QQ} = \frac{1}{n_k} \sum_{\mathbf{x} \in c_k} \|\mathbf{x} - \mathbf{m}_c^k\|^2$, em que n_k é o número de amostras da classe c_k e \mathbf{m}_c^k é o protótipo mais próximo pertencente à classe c_k .

balho, a solução de maior confiabilidade é a estratégia de voto majoritário em consequência das maiores estatísticas encontradas de média, mediana e mínima e em consequência do menor desvio padrão encontrado.

Já em cenários que apresentam recursos limitados, a melhor solução desejada é definida pela escolha de uma única métrica a ser considerada na análise dos dados. Logo, ao analisar os resultados obtidos entre os dois melhores índices, verifica-se que os critérios de desempate entre estes são a média e o desvio padrão onde, para estas estatísticas, *silhouettes* apresentou os resultados mais desejáveis.

Em uma última análise, avaliou-se a distribuição de quantidade de protótipos por classe sugerida pelo voto majoritário e por *silhouettes*. Estas distribuições são mostradas na Tabela III. Verifica-se, a partir destas sugestões de K_{opt} , que os dois métodos convergiram na mediana das sugestões, onde o vetor mediano de sugestões é $[K_N, K_A, K_B] = [2, 3, 4]$. Outro resultado relevante encontrado é dado pela quantidade máxima de protótipos reportada para representar a classe de baixa impedância ser mais sensível na modelagem por voto majoritário, em consequência de ter apresentado valores elevados ($K_B = 9$ e $K_B = 10$) que devem ser interpretados como exemplos excepcionais conforme mostrado na Tabela IV.

VI. CONCLUSÕES E TRABALHOS FUTUROS

Neste artigo, nós apresentamos uma metodologia baseada em validação de agrupamentos com o objetivo da obtenção de eficientes classificadores baseados em LVQ. Nossa motivação é originada pela solução da complexa tarefa de classificação de falhas que estamos investigando há alguns anos. O desafio almejado de detecção de falhas por curto-circuito entre espiras é desafiadora em decorrência da alta probabilidade de confusão entre exemplos de falhas de alta impedância como exemplos de normalidade.

Trabalhos anteriores a partir de soluções que utilizaram poderosas redes neurais supervisionadas, tais como redes MLP e RBF, originaram o desejo pela aplicação de classificadores baseados em protótipos em busca de soluções mais simples e que apresentem desempenhos aceitáveis na tarefa de classificação em interesse. Em comparação aos trabalhos anteriores do nosso grupo de pesquisa, nós conseguimos obter taxas de acurácia competitivas através de soluções que requerem menor custo computacional.

Atualmente, nós estamos investigando o desempenho desta abordagem a partir da utilização de outras técnicas de extração de características, entre elas as famílias de *wavelets* e MFCC (do inglês, *Mel-Frequency Cepstral Coefficients*).

AGRADECIMENTOS

Este estudo foi financiado pela Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES) - Código 001. Os autores também agradecem ao IFCE pela infraestrutura do Laboratório de Processamento de Energia.

REFERÊNCIAS

- [1] A. Webb, "Statistical pattern recognition, John Wiley & Sons," New York, USA, 2002.
- [2] V. N. Ghate and S. V. Dudul, "Optimal MLP neural network classifier for fault detection of three phase induction motor," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3468–3481, 2010.
- [3] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—a review," *IEEE transactions on energy conversion*, vol. 20, no. 4, pp. 719–729, 2005.
- [4] A. H. Bonnett, "Root cause failure analysis for ac induction motors in the petroleum and chemical industry," in *Petroleum and Chemical Industry Conference (PCIC), 2010 Record of Conference Papers Industry Applications Society 57th Annual. IEEE*, 2010, pp. 1–13.
- [5] G. Kliman, W. Premerlani, R. Koegl, and D. Hoeweler, "A new approach to on-line turn fault detection in ac motors," in *Industry Applications Conference, 1996. Thirty-First IAS Annual Meeting, IAS'96., Conference Record of the 1996 IEEE*, vol. 1. IEEE, 1996, pp. 687–693.
- [6] D. N. Coelho, G. A. Barreto, C. M. Medeiros, and J. D. A. Santos, "Performance comparison of classifiers in the detection of short circuit incipient fault in a three-phase induction motor," in *2014 IEEE Symposium on Computational Intelligence for Engineering Solutions (CIES)*. IEEE, 2014, pp. 42–48.
- [7] R. G. Vieira, C. M. Medeiros, and E. T. Silva, "Classification and sensitivity analysis to detect fault in induction motors using an mlp network," in *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 796–802.
- [8] D. P. Sousa, G. A. Barreto, and C. M. Medeiros, "Efficient selection of data samples for fault classification by the clustering of the som," in *Congresso Brasileiro de Inteligência Computacional. Niterói, RJ, 2017*.
- [9] D. N. Coelho, G. A. Barreto, and C. M. Medeiros, "Detection of short circuit faults in 3-phase converter-fed induction motors using kernel som," in *2017 12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM)*. IEEE, 2017, pp. 1–7.
- [10] D. P. Sousa, G. A. Barreto, C. C. Cavalcante, and C. M. Medeiros, "Lvq-type classifiers for condition monitoring of induction motors: A performance comparison," in *International Workshop on Self-Organizing Maps*. Springer, 2019, pp. 130–139.
- [11] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
- [12] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," 1973.
- [13] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "Clustering validity checking methods: part II," *ACM Sigmod Record*, vol. 31, no. 3, pp. 19–27, 2002.
- [14] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 224–227, 1979.
- [15] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern recognition*, vol. 37, no. 3, pp. 487–501, 2004.
- [16] A. K. Jain and R. C. Dubes, "Algorithms for clustering data," 1988.
- [17] 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.
- [18] T. Kohonen, "Improved versions of learning vector quantization," in *Neural Networks, 1990., 1990 IJCNN International Joint Conference on*. IEEE, 1990, pp. 545–550.
- [19] A. Sato and K. Yamada, "Generalized learning vector quantization," in *Advances in neural information processing systems*, 1996, pp. 423–429.
- [20] B. Hammer and T. Villmann, "Generalized relevance learning vector quantization," *Neural Networks*, vol. 15, no. 8-9, pp. 1059–1068, 2002.
- [21] B. Hammer, M. Strickert, and T. Villmann, "On the generalization ability of GRLVQ networks," *Neural Processing Letters*, vol. 21, no. 2, pp. 109–120, 2005.