

Diagnóstico de Curto-Circuito entre Espiras de Estator de Motores de Indução Usando Fluxo de Dispersão e Dados Estatísticos

Ailton de Oliveira Louzada* Wesley Angelino de Souza* Marcelo Favoretto Castoldi* Alessandro Goedtel*

* Universidade Tecnológica Federal do Paraná (UTFPR)
Av. Alberto Carazzai, 1640, CEP 86300-000, Cornélio Procopio, PR, Brasil
(e-mails: ailtonlouzada@alunos.utfpr.edu.br, wesleyangelino@utfpr.edu.br, marcastoldi@utfpr.edu.br, agoedtel@utfpr.edu.br)

Abstract: This work proposes the investigation of stator short circuit fault (SSCF) in Three Phase Induction Motors (TIM) related to stray flux. The level of defect is analyzed by means of an exploratory coil inserted in the motor housing, whose function is to sample the stray flux. Obtaining the voltage signals induced in the coil, it is possible to extract characteristics that can help in the automated identification of SSCF in TIM's. In this way, considering the need to identify these faults in an incipient, automatic way and considering the importance of finding relevant attributes for the adequate characterization of these faults, this article presents an attribute engineering method aimed at fault identification, presenting the process of data acquisition, extraction, selection, reduction and evaluation of statistical attributes to obtain an accurate diagnosis. The objective is to choose only potentially relevant attributes based on statistical information for the identification of SSCF, allowing the implementation in embedded systems with small computational capacity. The results show a reduction of 57 attributes initially extracted to 8, guaranteeing an accuracy of 86% in the identification of failures.

Resumo: Este trabalho propõe a investigação de falha de curto-circuito de estator (FCCE) em Motores de Indução Trifásico (MIT) relacionados ao fluxo de dispersão. Analisa-se o nível de defeito por meio de uma bobina exploratória inserida na carcaça do motor, cuja a função é amostrar o fluxo de dispersão. Obtendo os sinais de tensão induzida na bobina, tem-se a possibilidade de se extrair características que podem auxiliar na identificação automatizada de FCCE em MITs. Desta forma, considerando a necessidade de se identificar estas falhas de forma incipiente, automática e considerando a importância de se encontrar atributos relevantes para a adequada caracterização destas falhas, este artigo apresenta um método de engenharia de atributos visando a identificação de falhas, apresentando o processo de aquisição de dados, extração, seleção, redução e avaliação de atributos estatísticos para se obter um preciso diagnóstico. O objetivo é escolher somente os atributos potencialmente relevantes baseados em informações estatísticas para a identificação de FCCE, possibilitando a implementação em sistemas embarcados de pequena capacidade computacional. Os resultados apresentam uma redução de 57 atributos inicialmente extraídos para 8, garantindo a exatidão de 86% na identificação de falhas.

Keywords: Electric motor; stray flux; exploratory coil; feature extraction; statistical data; machine learning

Palavras-chaves: Motor elétrico; fluxo de dispersão; bobina exploratória; extração de características; dados estatísticos; aprendizado de máquina.

1. INTRODUÇÃO

Os Motores de Indução Trifásicos (MITs) são considerados excelentes opções quando se deseja conversão eletromecânica de energia. As vantagens desses motores são a sua robustez por praticamente não possuir elementos sensíveis, baixo custo, fácil manutenção, confiabilidade e simplicidade no controle (REIS, 2010).

Um aspecto importante é que os motores de indução são responsáveis pelo consumo de aproximadamente 40% de energia (mundial), sendo nos países desenvolvidos, esta porcentagem ficando em torno de 50% (FERREIRA, ALMEIDA, 2012). Apesar de sua robustez, os MITs estão suscetíveis a diversas falhas, entre as quais as de origem elétrica e mecânica. Como falhas mecânicas têm-se problemas

nos rolamentos, excentricidade do rotor, flexão (ou desalinhamento) do eixo, desalinhamentos causados por mau posicionamento de acoplamentos, contaminação interna, entre outras. As grandezas do motor, como corrente, tensão e torque, podem indicar características do estado de funcionamento, tendo elevado potencial para detectar possíveis falhas (REIS, 2010). As principais falhas elétricas podem ser falhas de espiras do estator (abertas ou curtos-circuitos); ou falhas de rotor (barras ou anéis terminais quebrados) (BELLINI *et al.*, 2008).

Estudos e pesquisas voltadas ao âmbito industrial revelam que uma grande parcela de falhas em MITs são originadas do estator. O isolamento elétrico do estator pode ser degradado rapidamente com uma pequena quantidade de espiras em curto-circuito, evoluindo para uma deterioração mais

acentuada, relacionada com o aumento de temperatura (TALLAM *et al.*, 2003). Pesquisas relacionadas às falhas de estator do MIT vêm sendo elaboradas, tanto com alimentação convencional da rede ou sendo alimentados por inversores de frequência (HENAO, CAPOLINO, 2003).

Na literatura, diversos trabalhos buscam meios para a identificação de falhas dos MITs. Em ERTAN e KEYSAN (2009), os harmônicos do entreferro foram investigados usando uma bobina externa instalada na carcaça de um MIT, focando na predição da velocidade do rotor, utilizando técnicas de análise espectral. Também foram explorados o melhor formato e posição de alocação desta bobina. Em SURYA *et al.* (2017), uma bobina observadora com 2 espiras foi inserida diretamente no fluxo radial, junto ao conjunto de bobinas principais de cada fase. Foram analisados o valor RMS, a fundamental e o terceiro harmônico, e tais grandezas foram diretamente interligadas com as falhas de curtos-circuitos de espiras de estator (FCCE) de sua fase correspondente.

FROSINI *et al.* (2017) utilizaram a Análise Discriminante Linear e Análise Espectral com três medições, sendo uma de corrente em uma das fases, e duas do fluxo de dispersão (uma bobina comercial e uma bobina confeccionada) para detectar falhas de curtos-circuitos nas espiras do estator, de MITs acionado por inversor. CAPOLINO *et al.*, (2019) investigam o uso do fluxo de dispersão utilizando-se de bobinas externas a carcaça da máquina para detectar falhas tais como barras, anéis e amortecedores quebrados em máquinas de corrente alternada (CA) e também, curtos-circuitos em estágios iniciais nos enrolamentos do estator.

Em boa parte dos estudos, uma ou mais técnicas são utilizadas e combinadas para a devida caracterização da falha. Na maioria dos casos, não é realizado um estudo de quais atributos poderiam ser utilizados com maior relevância, tendo maior potencial para caracterizar um defeito, sendo utilizado como dados de entrada para algoritmos de reconhecimento de padrões. Alguns atributos são escolhidos como relevantes na classificação, mas muitas vezes de forma aleatória ou por tentativa e erro.

Considerando a necessidade de se identificar falhas em motores e considerando a importância de se encontrar atributos relevantes para a adequada caracterização de falhas, este artigo apresenta um método de engenharia de atributos para a identificação de falhas de motores através de um processo de aquisição de dados, extração, seleção e avaliação de atributos estatísticos para um preciso diagnóstico de falha de motores. Tem-se assim o objetivo de escolher os atributos potencialmente relevantes baseados em informações estatísticas para a identificação de falhas de motores.

A Seção 2 apresenta o conceito de importância de atributos e algoritmos de classificação, a Seção 3 apresenta o método de engenharia de atributos para a identificação de falhas, a Seção 4 apresenta os resultados e discussões e a Seção 5 apresenta as considerações finais do trabalho e futuros trabalhos.

2. ATRIBUTOS E CLASSIFICADORES

2.1 Atributos Provenientes de Sinais

Nos tempos atuais, a complexidade e dispersão de sistemas conectados à internet levou a coleta de imensas quantidades de dados. Estes dados, por sua vez, possuem informações de suma importância para, dependendo da área de atuação, emular comportamentos de sistemas, tomar decisões assertivas, aumentar a eficiência entre outros. Desta maneira há a necessidade de transformar tais dados em informação e posterior conhecimento (BORELI, 2018). Assim, a extração de características se torna essencial, pois apresentam informações relevantes e únicas para cada padrão ou categorização, melhorando a separabilidade e posterior classificação. Além disto, é possível reduzir a quantidades de parâmetros para o classificador (NAGATA, 2018).

Uma das formas de se extrair características é através de recursos estatísticos, como as medidas de tendência central ou posição, bem como as de dispersão e as separatrizes. Considerando as informações estatísticas, foram selecionadas as seguintes grandezas:

- Média: uma medida de tendência central;
- Mediana: medida que divide os dados ordenados ao meio;
- Moda: valor que aparece com mais frequência;
- Desvio Absoluto Médio (DAM): média dos valores absolutos dos desvios em relação à média;
- Variância: medida de dispersão e é uma média dos quadrados dos desvios;
- Desvio Padrão: medida de dispersão em torno da média de uma variável. Extrai-se a raiz quadrada da variância, tornando a unidade de medida a mesma do conjunto de dados;
- Quartis: medidas separatrizes e dividem a distribuição dos dados ordenados em quatro partes: Q1 até 25% dos valores, Q2 que divide o conjunto ao meio (mediana) e Q3 o que separa os 75% dos valores;
- Curtose: representa o grau de achatamento ou alongamento da distribuição;
- RMS (*Root Mean Square*) é uma medida de estatística da magnitude de um valor;
- Histograma: semelhante ao diagrama de barras, porém refere-se a uma distribuição de frequências construída com intervalos de classes. Os intervalos de classes são colocados no eixo horizontal enquanto as frequências são colocadas no eixo vertical;
- Assimetria (*Skewness*) que define uma distribuição de frequência sendo simétrica se a metade esquerda de seu histograma é praticamente uma imagem

espelhada de sua metade direita, caso contrário, será assimétrica. (FERREIRA, 2015).

Convém observar que foi selecionado apenas um número pequeno de medidas estatísticas, na tentativa de avaliar a potencialidade das mesmas.

2.2 Algoritmos classificadores

Um algoritmo classificador, no contexto de aprendizado de máquinas, é um algoritmo que automaticamente ordena ou categoriza os dados em uma ou mais “classes”. Por exemplo, um algoritmo classificador pode ser treinado e utilizado para escanear e-mails e filtrá-los em duas classes: “Spam” ou “Não Spam”. Desta forma, os algoritmos de aprendizado de máquina são úteis para automatizar tarefas que antes tinham que ser feitas manualmente (SOUZA *et al.*, 2018).

Existem diversos algoritmos classificadores, cada um possuindo um modelo computacional característico, desempenhando um comportamento eficiente ou ineficiente em uma base de dados. Pode-se citar como exemplo alguns classificadores: *k*-NN (*k*-Nearest Neighbors), SVM (*Support Vector Machine*) Linear e com núcleo RBF (*Radial Basis Function*), *Gaussian Process*, *Naive Bayes*, DT (*Decision Tree*), RF (*Random-Forest*), MLP (*Multilayer Perceptron*), *AdaBoost* (*Adaptive Boosting*) e QDA (*Quadratic Discriminant Analysis*).

Considerando que os atributos e os algoritmos classificadores são relevantes no processo de identificação de padrões, a Seção 3 apresenta a técnica utilizada neste artigo para o processo de engenharia de atributos, que tem como objetivo extrair atributos adequados dos sinais aquisitados e apresentar os melhores atributos para a identificação de falhas em motores.

3. MÉTODO DE SELEÇÃO DE ATRIBUTOS

Esta seção apresenta o processo de engenharia de atributos, que tem como objetivo estabelecer os atributos que possuem melhor relevância a partir do sinal aquisitado. Desta forma, são apresentados os conceitos que foram utilizados: aquisição de dados, extração de atributos, construção da base de dados e o processo de seleção e redução de atributos.

3.1 Aquisição de dados

A aquisição de dados é um processo de conversão de sinais do mundo real para o domínio digital, considerando a coleta através de sensores ou dispositivos coletadores de sinais. Neste artigo, a aquisição de dados é feita através de um sinal periódico em alta frequência, sendo os dados brutos para o processo de extração de atributos. Foram coletadas 45 aquisições das formas de onda de tensão induzidas (chamadas de Aq01 até Aq45), em uma bobina de indução (sensor) a partir de um osciloscópio, em um motor de 1 HP (*Horse Power*). Variou-se o torque imposto pela carga de 0,5 N até 4,5 N, com passo de 0,5 N, e a condição do motor desde saudável e com falhas de porcentagem de curtos-circuitos de espiras de estator, variando desde 1%, 3%, 5% e 10% conforme apresentado na Tabela 1.

Tabela 1. Aquisição de dados

Aquisição - Motor de 1 HP					
Torque (N)	Saudável	1%	3%	5%	10%
0,5	Aq 01	Aq10	Aq19	Aq28	Aq37
1,0	Aq02	Aq11	Aq20	Aq29	Aq38
1,5	Aq 03	Aq12	Aq21	Aq30	Aq39
2,0	Aq 04	Aq13	Aq22	Aq31	Aq40
2,5	Aq 05	Aq14	Aq23	Aq32	Aq41
3,0	Aq 06	Aq15	Aq24	Aq33	Aq42
3,5	Aq 07	Aq16	Aq25	Aq34	Aq43
4,0	Aq 08	Aq17	Aq26	Aq35	Aq44
4,5	Aq 09	Aq18	Aq27	Aq36	Aq45

A Fig. 1 mostra seis formas de onda de seis ciclos de aquisição como exemplo.

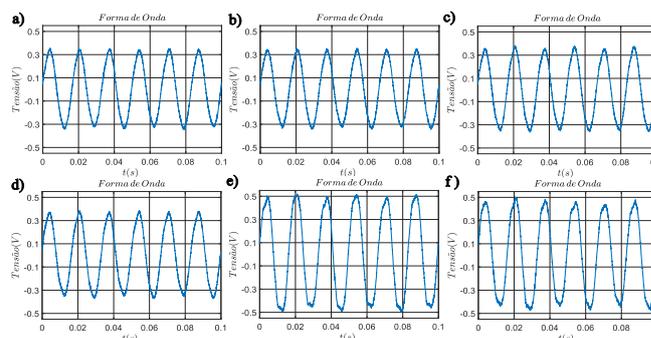


Fig. 1. Exemplos de formas de onda das aquisições (Matlab)

Observando-se as formas de onda da Fig. 1 tem-se: a) aquisição saudável com 0,5 N de carga, b) 1% de curto-circuito de espiras de estator com 0.5 N de carga, c) 3% de curto-circuito com 0.5 N de carga, d) 5% de curto-circuito com 0.5 N de carga, e) 10% de curto-circuito com 0.5 N de carga e f) 10% de curto-circuito com 1 N de carga. Na sequência é realizado uma normalização do sinal (mim, máx), deixando-as preparadas para o processo de extração de atributos.

Para a aquisição dos sinais foi utilizada uma bobina de indução externamente fixada à carcaça da máquina que captura o fluxo de dispersão do estator, que está presente na parte externa da carcaça. A bobina é constituída por um núcleo de material ferromagnético de grão não orientado, formato de U, com 291 espiras de fio de cobre esmaltado AWG23, sendo apresentada na Fig. 2.



Fig. 2 Bobina de indução (sensor).

Na Fig. 3 é possível observar os equipamentos que compõem a configuração experimental adotada: gerador CC, torquímetro Kistler, MIT e osciloscópio.

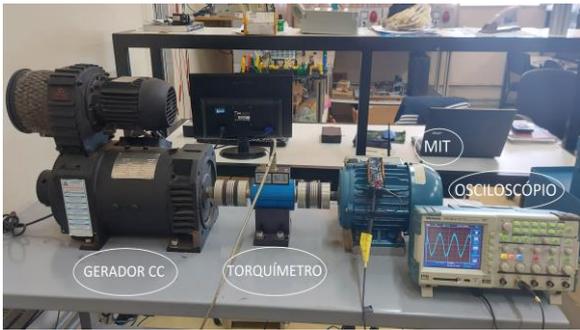


Fig. 3 Bancada experimental.

Por meio da bancada experimental, foi possível realizar ensaios com o motor operando desde 0.5 N à 4,5 N de carga. O elemento de carga do MIT é um gerador CC, ligado a resistores de potência. O MIT tem seus enrolamentos modificados para se inserir curtos-circuitos de espiras de estator da ordem de 1%, 3%, 5% e 10%. Os dados do MIT usados são: Marca: WEG, Modelo: W21 – 3~80, Rotor: Gaiola de esquilo, Potência: 1 HP, Número de Polos: 04, Tensão Nominal: 220/380V, Corrente Nominal: 3,02/1,75 A, Ip/In: 7,1, Rotação Nominal: 1720 RPM, Frequência: 60 Hz, Categoria: N, Fator de Serviço: 1,15, Fator de Potência: 0,82 e Rendimento: 79,5%.

O sinal da tensão induzida, proveniente da bobina, foi amostrado por um osciloscópio Tektronix modelo TPS2014, conforme Fig. 3. A tensão de acionamento do MIT possui um leve desbalanceamento, característica da própria rede. A taxa de amostragem do osciloscópio foi de 25 kHz, sendo feito 45 aquisições.

3.2 Extração de atributos

A partir do conjunto de formas de ondas aquisitadas, foram extraídos os atributos estatísticos: Média, Mediana, Moda, Desvio Padrão, Variância, 1º, 2º e 3º Quartis, Desvio Absoluto Médio (DAM), Curtose, *SkewNess*, RMS, (12 atributos somados) e histograma. Para o histograma, a divisão do intervalo de classes foi de 45 partes, de H1 até H45, totalizando 57 atributos extraídos (12 dos principais atributos estatísticos + 45 do histograma). A Tabela 2 apresenta a lista destes atributos.

Tabela 2. Atributos

Mediana	Média	Moda	Desv. Pad.	Variância	1-Quart	2-Quart	3-Quart	DAM	Curtose	SkewNess	RMS
Histogramas											
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24
H25	H26	H27	H28	H29	H30	H31	H32	H33	H34	H35	H36
H37	H38	H39	H40	H41	H42	H43	H44	H45			

3.3 Composição da base de dados

Devido à escassez de dados da Tabela 1 (45 formas de onda no total), foi feita uma expansão da base de dados. A expansão foi feita selecionando um trecho do sinal contendo 3 formas de onda, utilizando-se janela móvel. Dando um passo de uma forma de onda, seleciona-se as 3 subseqüentes, assim por diante, até serem extraídos 4 trechos do sinal, conforme Fig. 4.

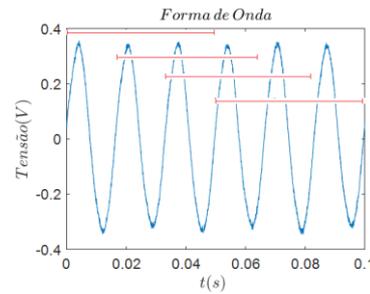


Fig. 4 - Expansão da base de dados

Assim, cada amostra original gerou mais 4 outras, totalizando 5, aumentando a base de dados de 45 amostras (5 classes de defeitos por 9 variações de torque) para 225 amostras. Os atributos foram extraídos desta nova base de 225 instâncias (maior quantidade) mas mantendo a relação 5 classes e 9 variações de torque.

3.4 Seleção de atributos

A seleção de atributos é uma etapa crucial no processo de aprendizado da máquina (SOUZA *et al.*, 2018). Atributos desnecessários aumentam o tempo computacional nas etapas de treinamento e classificação, também diminuem a interpretabilidade do modelo e podem diminuir a generalização dos dados (ABEDINIA *et al.*, 2017). Desta forma, tem-se a seleção de atributos conforme Fig. 5, tendo quatro etapas do método de escolha dos melhores atributos.

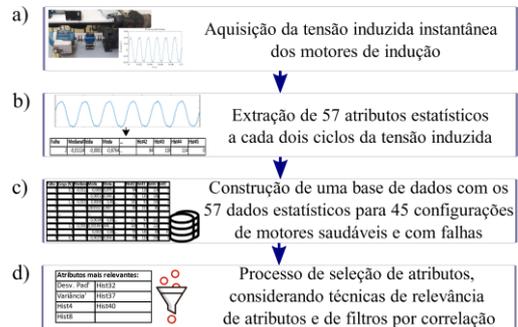


Fig. 5. Técnica de seleção de atributos

A técnica possui quatro etapas, que são:

a) Dados em alta frequência: Nesta etapa, tem-se os dados brutos, isto é, a forma de onda de tensão induzida de um motor, que foi coletada por um osciloscópio, com resolução de 8 bits e com 25kHz de frequência de amostragem. São necessários dados em alta frequência nesta etapa para que seja possível realizar a extração dos atributos estatísticos.

b) Cálculo dos 57 atributos estatísticos selecionados através de métodos estatísticos: Conforme Seção 3.3, extrai-se os 57 atributos estatísticos apresentados na Tabela 2.

c) Base de dados em baixa frequência com os 57 atributos e rótulo de falha do motor: A cada instância, tem-se 57 atributos e um atributo-alvo referente ao tipo de falha do motor medido. A base de dados para este trabalho é composta por 225 instâncias, conforme indicado na Seção 3.3.

d) Filtro dos atributos considerando a colinearidade: Com o conjunto de dados criado na etapa anterior, é analisada a relevância dos atributos no processo de classificação,

excluindo aqueles que não interferem fortemente nas métricas de desempenho dos algoritmos de aprendizado de máquina. O processo utilizado para a redução de atributos neste artigo é baseado na relevância dos atributos por floresta aleatória (RF - *Random Forest*) e posteriormente por colinearidade.

3.4.1 Floresta Aleatória como seletora de atributos

O RF é um algoritmo composto por um conjunto de árvores de decisão. O RF foi proposto por HO (1995), para promover o aprimoramento de árvore de decisão e resolver problemas de super ajuste dos modelos classificadores (*overfitting*). No processo de treinamento, são apresentados os dados rotulados e diversas árvores de decisão são criadas e, no processo de classificação, prevalece a classe indicada pela maioria das árvores de decisão que compõe a “floresta”. Cada árvore de decisão da RF é um conjunto de nós e folhas internas. No nó interno, a característica selecionada é usada para tomar a decisão de como dividir o conjunto de dados em dois conjuntos separados, o mais puro possível para fazer a separação de classes (HO, 1995). Neste presente trabalho, é utilizada a RF como indicadora de relevância de atributos com base na observação de quantas vezes um determinado atributo aparece nos nós das árvores de decisão e, conseqüentemente, o quanto o atributo contribui na classificação correta (MENZE, 2009). Os inconvenientes do método são a tendência de preferência de características numéricas e características categóricas com alta cardinalidade. Além disso, no caso de características correlatas, ele pode selecionar uma das características e negligenciar a importância da segunda (o que pode levar a conclusões erradas) (MENZE, 2009).

3.4.2 Colinearidade

Na estatística, a colinearidade é a análise do fenômeno no qual um atributo é altamente correlacionado a um outro (SOUZA, 2022). De forma analítica, se existe alta correlação entre pares de atributos, há conseqüentemente um comportamento aproximadamente linear entre eles, conforme é apresentado na Fig. 6. Nesta figura tem-se o par de atributos (Desvio Padrão, Variância) e (DAM, Desvio Padrão) com alta correlação e o gráfico do par de atributos apresenta um comportamento linear entre eles, apresentando então o conceito de que os atributos são altamente dependentes e que um deles pode ser removido.

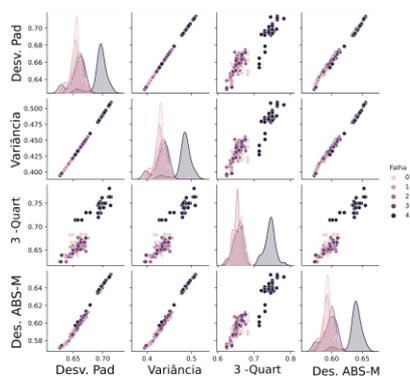


Fig. 6 Exemplo de alta correlação entre atributos.

Desta forma, se há alta correlação entre dois atributos, é possível dizer que existem informações semelhantes entre os

atributos e pode ser considerado que exista redundância de dados. Outro fenômeno que acontece na existência de atributos altamente correlacionados é o efeito de *overfitting*, que faz o ajuste super ótimo do modelo de classificação aos dados no processo de treinamento, mas, ao serem classificados novos dados, o modelo de classificação se torna ineficaz. Atributos com correlação maior do que 0,9 são considerados altamente dependentes e há um comportamento linear entre eles. Removendo um dos atributos, reduz-se o tempo de treinamento e classificação e, em alguns casos, aumenta-se os níveis de assertividade dos métodos.

3.5 Desempenho dos classificadores

Para explicar as métricas de desempenho dos classificadores, tais conceitos devem ser apresentados:

- Positivos verdadeiros (TP): Esses são os valores positivos preditos corretamente, o que significa que o valor da classe real é sim e o valor da classe prevista também é sim.
- Verdadeiros Negativos (TN): Estes são os valores negativos previstos corretamente, o que significa que o valor da classe real é não e o valor da classe prevista também é não.
- Falsos Positivos (FP): Quando a classe real é não e a classe prevista é sim, ou seja, houve um erro.
- Falsos Negativos (FN): Quando a classe real é sim, mas a classe prevista é não, ou seja, houve um erro.

Considerando estes conceitos, tem-se as métricas utilizadas na comparação de desempenho dos classificadores.

Exatidão: também chamado de acurácia, corresponde à proporção entre os pontos classificados corretamente em relação com a totalidade de dados, conforme Eq. (1).

$$\text{Exatidão} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Precisão: medida de desempenho intuitiva e é proporção da observação corretamente prevista para o total de observações. O valor da precisão é correspondente a Eq. (2).

$$\text{Precisão} = \frac{TP}{TP+FP} \quad (2)$$

Recall (Revocação): é a razão de observações positivas preditas corretamente para todas as observações na classe real verdadeira, conforme apresenta a Eq. (3).

$$\text{Revocação} = \frac{TP}{TP+FN} \quad (3)$$

F1-medida: a média ponderada de Precisão e Revocação. Portanto, essa pontuação leva em conta tanto os falsos positivos quanto os falsos negativos, conforme (4).

$$\text{F1-medida} = 2 * \frac{\text{Revocação} * \text{Precisão}}{\text{Revocação} + \text{Precisão}} \quad (4)$$

Em aspecto quantitativo, a exatidão é uma métrica mais generalista para conjunto de dados balanceados, porém apresenta desvantagens para conjuntos não balanceados. A precisão destaca a situação nas quais os falsos positivos são

mais prejudiciais, A revocação destaca os falsos negativos como mais prejudiciais. A F1-medida é uma métrica que balanceia tanto a revocação quanto precisão. Considerando tais aspectos, a Seção 4 apresenta os resultados e discussões sobre o processo de extração e seleção de atributos.

4. RESULTADOS E DISCUSSÕES

4.1 Técnica de seleção de atributos

Cada etapa apresentada da Fig. 3 é executada no procedimento experimental. Inicialmente, foram utilizados os dados de alta frequência, com os casos indicados na Tabela 1, tais dados possuem a oscilografia da tensão induzida, como mostrado na Fig. 5 (a). Posteriormente, foi realizada a extração de atributos, através de um *script* desenvolvido em Matlab. Os 57 atributos estatísticos apresentados na Tabela 2, foram calculados conforme apresentado na Seção 3.3. Tais atributos, juntamente ao atributo alvo, se tornam uma instância inserida no conjunto de dados, conforme mostrado na Fig. 5 (c). Para verificar se o conjunto de características oferece uma boa representação computacional para a identificação de falhas nos motores, foram utilizados três algoritmos de aprendizado de máquina: *k*-Vizinhos mais próximos (*k*-NN), Árvores Aleatórias como classificador (RF) e Perceptron Multicamadas (MLP).

Desta forma, os dados foram normalizados pelo método Z-score. No processo de avaliação dos resultados, foi utilizada a validação cruzada *k-fold* com 10 dobras (SHAO, 1993). As métricas exatidão, precisão e F1-medida foram utilizadas para avaliar o desempenho dos algoritmos de classificação. Além disso, o tempo médio de treinamento e testes foi coletado. Os testes foram realizados em um computador com sistema operacional Ubuntu Linux, processador i7-8750H, 16 GB de RAM. Os experimentos foram implementados usando o Python 3 com a biblioteca *sklearn*.

Para a melhor parametrização dos algoritmos de aprendizado de máquina, foi realizado o processo de busca em grade para o conjunto de dados, parametrizando os algoritmos *k*-NN ($k = \{1, 3, 5, \dots, 13, 15, 17\}$ e distância = {euclidiana, manhattan}), RF (árvores = {10, 20, 30, 40, ..., 170, 180} e profundidade da árvore = {5, 10, 15, 20, ..., 60, 70}) e MLP (taxa de aprendizado = {0,0001, 0,001, 0,01, 0,1, 1, 10, 100}), 1 a 4 camadas, 10:10:50 neurônios, função de ativação *Relu* e na saída função de ativação *Softmax*. Os melhores parâmetros para o *k*-NN foram $k = 1$ e distância manhattan, para o RF foram o número de árvores = 160 e profundidade da árvore = 20 e para o MLP foram 3 camadas, 40-30-40 neurônios e taxa de aprendizagem = 0,1.

Os algoritmos de classificação foram inicialmente utilizados para a base de dados composta pelos 57 atributos e, a cada etapa de seleção de atributos, os mesmos algoritmos são utilizados para a verificação do desempenho. Estas análises visam avaliar se a eliminação de características não interfere negativamente no desempenho dos algoritmos de classificação e, consequentemente, na capacidade de classificação das falhas de motores. As etapas são apresentadas a seguir.

a) Análise da base de dados com 57 atributos

Nesta etapa, foram considerados os 57 atributos estatísticos e a Tabela 3 apresenta o desempenho dos algoritmos de aprendizado de máquina, tendo o MLP com o melhor desempenho considerando tanto a exatidão quanto a macro F1-medida e precisão, porém há o maior tempo computacional. O método *k*-NN, por outro lado, apresentou o menor tempo computacional.

Tabela 3. Métricas de desempenho dos classificadores considerando os 57 atributos estatísticos

Métrica	<i>k</i> -NN	RF	MLP
Exatidão	0,88	0,85	0,90
Macro F1-medida	0,88	0,82	0,90
Precisão	0,88	0,84	0,90
Tempo de treinamento (s)	-	0,26	0,33
Tempo de testes (s)	0,04	0,06	0,14

b) Análise da base de dados com os atributos pelo RF como seletor de atributos

Posteriormente, tem-se a técnica de coletar os atributos mais relevantes considerando o algoritmo RF como seletor de atributos. Neste caso, dos 57 atributos apresentados na Fig. 7, foram selecionados os 13 com relevância maior que 2%. Os atributos selecionados são: Desvio Padrão, Variância, 1º Quartil, 3º Quartil, DAM, Curtose, SkewNess, Histogramas H4, H8, H32, H37 e H40 e RMS.

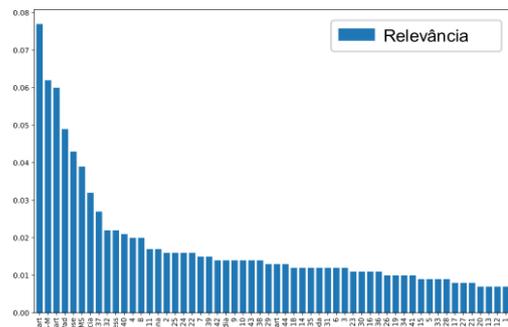


Fig. 7 Gráfico de ordenação dos atributos mais relevantes utilizando o algoritmo RF.

A Tabela 4 apresenta os índices de desempenho dos três algoritmos de aprendizado de máquina. Neste caso, o algoritmo RF como classificador apresentou o melhor desempenho considerando as métricas de exatidão, precisão e f-medida. Em tempo computacional, o *k*-NN e o RF tiveram o tempo computacional de 0,01s no processo de testes, mas o *k*-NN, por não possuir estágio de treinamento, ainda se torna um algoritmo com o melhor desempenho para este estágio de atributos.

Tabela 4. Métricas de desempenho considerando os 13 atributos que restaram após o uso do algoritmo RF

Métrica	<i>k</i> -NN	RF	MLP
Exatidão	0,83	0,85	0,65
Macro F1-medida	0,83	0,84	0,69
Precisão	0,83	0,84	0,65
Tempo de treinamento (s)	-	0,15	0,19
Tempo de testes (s)	0,01	0,01	0,13

Após utilizar o subconjunto de dados composto pelos 13 atributos, foi feito o estudo de correlação, conforme apresentado na Fig. 8, para verificar a possibilidade de remoção manual de atributos considerando o critério da colinearidade.

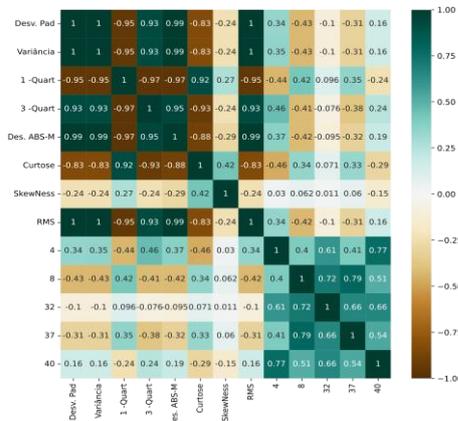


Fig. 8 Correlação entre os 13 atributos selecionados.

Desta forma, considerando a remoção um dos atributos com colinearidade maior que 0,9, tem-se os pares (Desvio Padrão, Variância), (Desvio Padrão, DAM), (Desvio Padrão, RMS), (Desvio Padrão, 1º Quartil), (Desvio Padrão, 3º Quartil) e (1º Quartil, Curtose) com alta correlação, sendo possível remover um dos atributos, pois possuem informações dependentes. Desta forma, dos 13 atributos, foram removidos a Variância, o valor RMS, 1º Quartil, 3º Quartil e o DAM, restando 8 atributos. A correlação dos 8 atributos restantes é apresentada na Fig. 9.



Fig. 9 Correlação entre os 8 atributos selecionados após a remoção por colinearidade > 0,9.

Para verificar se os atributos restantes ainda possuem um bom desempenho no processo de identificação de falhas, os três algoritmos de aprendizado de máquina foram executados, tendo os resultados apresentados na Tabela 5.

Tabela 5. Métricas de desempenho dos classificadores considerando os 8 atributos que restaram após o uso do algoritmo RF e de colinearidade maior que 0,9

Métrica	k-NN	RF	MLP
Exatidão	0,83	0,86	0,71
Macro F1-medida	0,82	0,85	0,69
Precisão	0,83	0,85	0,69
Tempo de treinamento (s)	-	0,09	0,17
Tempo de testes (s)	0,01	0,01	0,11

Tem-se o RF como melhor classificador, apresentando um melhor desempenho ao comparar com o processo de 13 atributos, passando para 86% de exatidão. O tempo computacional também foi reduzido nos processos de treinamento do algoritmo RF e da MLP, tendo o k-NN e RF como os que apresentaram melhor desempenho.

4.3. Análise de Resultados e Discussões

Os processos utilizados na seleção de atributos permitiram uma significativa redução da base de dados em relação a atributos, passando de 57 para 13 utilizando o algoritmo RF como relevância de atributos e de 13 para 8 utilizando a remoção por colinearidade acima de 0,9. Esta significativa redução fez com que o tempo de execução dos algoritmos k-NN e RF reduzissem a ponto de serem executados em 10 milissegundos, tornando-os interessantes em aplicações como em dispositivos com processamento limitado, como microcontroladores. Porém, a qualidade dos atributos restantes fez com que a exatidão do melhor algoritmo caísse de 0,9 para 0,86.

A mudança do melhor algoritmo classificador, que com 57 atributos tem-se o MLP e com 8 atributos tem-se o RF classificador, mostra que a MLP e o k-NN captam pequenas variações nos dados, sendo interessante na existência de atributos com poucas variações e podendo possuir altas correlações, enquanto o RF tem um desempenho estável em todas as etapas de seleção de atributos.

Outro ponto a ser observado é que o algoritmo k-NN, por não ter o estágio de treinamento, faz a carga de processamento no estágio de classificação. Na base de dados deste artigo, considerada pequena por ser composta por 225 instâncias de diferentes cargas do motor e condições de falhas, conforme apresenta a Fig. 10 (curva de aprendizado dos três classificadores utilizados neste trabalho). De acordo com a Fig. 10, a acurácia aumenta gradativamente quando a quantidade de instancias de treinamento aumenta, e com 225 instâncias, os modelos classificadores não convergiram ao nível estável de acurácia.

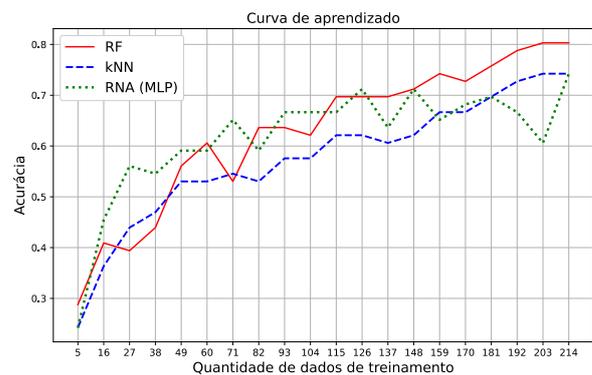


Fig. 10 - Curva de aprendizado.

Mesmo considerando a não estabilidade da acurácia devido a poucas instâncias, há um bom desempenho. Desta forma, pode-se concluir que o algoritmo RF como classificador se torna um interessante algoritmo para ser utilizado na base de dados para a identificação de falhas de estator em MITs.

5. CONCLUSÕES

Este artigo apresenta uma metodologia usada para a verificação tanto quantitativa como qualitativa de atributos estatísticos básicos originários de sinais senoidais periódicos, coletados de uma bobina externamente acoplada em um motor de indução (aquisitadas a partir de um simples osciloscópio) que poderiam ser usados como dados de entrada para algoritmos de aprendizado de máquina.

Neste método, é feita a redução de atributos baseado na floresta aleatória (RF) e colinearidade. Na sequência é investigada se os 57 atributos iniciais como as reduções satisfazem a relação custo computacional x precisão, utilizando como parâmetros comparativos os algoritmos: k -Vizinhos mais Próximos (k -NN), Árvores Aleatórias (RF) como classificador e Perceptron Multicamadas (MLP). Mesmo após a redução de 57 para 8 atributos, a exatidão apresentou valor de 86%. Desta forma, os dados estatísticos se mostraram promissores em se tratando de simplicidade computacional, uma vez que a definição de cada um deles é de cálculos de fácil execução.

Deve-se destacar que a aquisição foi feita com um osciloscópio com resolução horizontal de 2500 pontos, resolução vertical de 8 bits o que definitivamente, é extremamente baixa se comparada a dispositivos microcontroladores atuais (de 10 a 24 bits). Estes equipamentos não são dedicados a coleta, mas sim visualização do sinal. Vendo desta forma esta técnica se torna muito promissora em se tratando de sistemas embarcados de baixo custo, onde a capacidade de aquisição e processamento dos dados é bem limitada.

REFERÊNCIAS

ABEDINIA, O., AMJADY, N., & ZAREIPOUR, H. (2016). A new feature selection technique for load and price forecast of electrical power systems. *IEEE Transactions on Power Systems*, 32(1), 62-74.

BELLINI, A., FILIPPETTI, F., TASSONI, C., CAPOLINO, G. A. (2008) Advances in Diagnostic Techniques for Induction Machines. *IEEE Transactions On Industrial Electronics*, 55(12), 4109 – 4126.

BORELI, A. F. Extração de Características em Sinais Biológicos Utilizando Estatística de Alta Ordem. 2018. 75f. Dissertação (Mestrado em Eng. Eletrônica) – Laboratório de Modelagem, Análise e Controle de Sistemas Não-Lineares, Universidade Federal de Minas Gerais, Belo Horizonte.

CAPOLINO, G. A., ROMARY, R., HÉNAO, H., PUSCA, R. (2019) State of the Art on Stray Flux Analysis in Faulted Electrical Machines *IEEE Workshop on Electrical Machines Design, Control and Diagnosis* 2019.

ERTAN, H. B., KEYSAN, O. (2009) External Search Coil as a Means of Measuring Rotor Speed of an Induction Motor. *International Symposium on Advanced Electromechanical Motion Systems & Electric Drives Joint Symposium*.

FERREIRA, F. J. T. E., ALMEIDA, A. T., (2012). Induction motor downsizing as a low-cost strategy to save energy. *Journal of Cleaner Production*, 24, 117-131.

FERREIRA, V. A. Estatística Básica, 1ª edição, Rio de Janeiro: Estácio, 2015.

FROSINI, L., ZANAZZO, S., BECCARISI, F. (2017) Linear Discriminant Analysis for an Automatic Detection of Stator Faults in Induction Motor Drives. *IEEE International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives* 2017.

HENAO, H., DEMIAN, C., CAPOLINO, G.A. (2003) A Frequency-Domain Detection of Stator Winding Faults in Induction Machines Using an External Flux Sensor. *IEEE Transactions on Industry Applications*, 39(5), 1272-1279.

HO, T. K. (1995). Random decision forests. In *Proceedings of international conference on document analysis and recognition* (Vol. 1, pp. 278-282). IEEE.

MENZE, B. H., KELM, B. M., MASUCH, R., HIMMELREICH, U., BACHERT, P., PETRICH, W., HAMPRECHT, F. A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC bioinformatics*, 10(1), 1-16.

NAGATA, E. A. Detecção, Segmentação E Classificação De Afundamentos De Tensão Em Sistemas Elétricos De Potência. 2018. 92f. Dissertação (Mestrado em Eng. de Sistemas e Automação) Universidade Federal de Lavras, Minas Gerais.

REIS, A. J. S. Reconhecimento de Padrões de Falhas em Motores Trifásicos Utilizando Redes Neurais. 2010. 73f. Dissertação (Mestrado em Eng. Elétrica) – Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal.

SHAO, J. (1993) Linear Model Selection by Cross-validation, *Journal of the American Statistical Association*, 88:422, 486-494.

SURYA, G., KHAN, Z. J., BALLAL, M. S., SURYAWANSHI, H. M. (2017) A Simplified Frequency Domain Detection of Stator Turn Fault in Squirrel Cage Induction Motors Using Observer Coil Technique. *IEEE Transactions on Industrial Electronics*. 64(2), 1495-1506.

SOUZA, W. A., ALONSO, A. M., BOSCO, T. B., GARCIA, F. D., GONÇALVES, F. A., MARAFÃO, F. P. (2022). Selection of features from power theories to compose NILM datasets. *Advanced Engineering Informatics*, 52, 101556.

SOUZA, W. A., MARAFÃO, F. P., LIBERADO, E. V., SIMÕES, M. G., DA SILVA, L. C. (2018). A nilm dataset for cognitive meters based on conservative power theory and pattern recognition techniques. *Journal of Control, Automation and Electrical Systems*, 29(6), 742-755.

TALLAM, R.M., HABETLER, T. G., HARLEY R.G. (2003). Experimental Testing of a Neural-Network-Based Turn-Fault Detection Scheme for Induction Machines under Accelerated Insulation Failure Conditions. *IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives* 2003.