Identificação de Defeitos em Rolamentos de Motores de Indução Trifásicos Utilizando Medidas de Informação e Ferramentas Meta-Heurísticas^{*}

Gustavo H. Bazan^{*,**} Alessandro Goedtel^{**}
Fábio R. Durand^{**} Wagner F. Godoy^{**}

^{*} Instituto Federal do Paraná (IFPR), Av. Dr Tito s/n - Jacarezinho, Paraná, Brasil (e-mail: gustavo.bazan@ifpr.edu.br)

^{**} Universidade Tecnológica Federal do Paraná (UTFPR-CP), Av. Alberto Carazzai 1640, Cornélio Procópio, Paraná, Brasil, (e-mail: agoedtel@utfpr.edu.br, fabiodurand@utfpr.edu.br, wagnergodoy@utfpr.edu.br).

Abstract: Three-phase induction motors are widely employed in industrial production processes due to their characteristics. Detecting a fault when these motors are in operation has received special attention from the industry, because can lead to a reduction of losses in the processes and reliable operation of industrial systems. Thus, this paper presents a study concerning the use of meta-heuristics tools in the classification of bearing fault diagnosis in induction motors. The extraction of the fault characteristics performed based on information measurements between the current signals. After, it is used the Artificial Bee Colony algorithm to select the relevant attributes and to optimize the pattern classifier input matrix. To evaluate the classification accuracy under various fault severity levels, the performance of two different pattern recognition techniques were compared: decision tree C4.5 and artificial neural network type multilayer perceptron. Experimental results confirm the effectiveness of the proposed methodology.

Resumo: Os motores de indução trifásicos são amplamente utilizados nos processos de produção industrial devido às suas características. Detectar uma falha quando esses motores estão em operação recebeu atenção especial da indústria, pois pode levar a redução de perdas nos processos e operação confiável de sistemas industriais. Assim, este artigo apresenta um estudo sobre o uso de ferramentas meta-heurísticas na classificação do diagnóstico de falhas de rolamentos em motores de indução. A extração das características são realizadas com base em medidas de informação entre os sinais das correntes. Depois, emprega-se o algoritmo de colônia artificial de abelhas para selecionar os atributos relevantes e otimizar a entrada do classificador de padrões. Para avaliar a precisão da classificação em vários níveis de severidade de falhas, comparou-se o desempenho de duas técnicas de reconhecimento de padrões: árvore de decisão C4.5 e rede neural artificial multilayer perceptron. Resultados experimentais confirmam a eficácia da metodologia proposta.

Keywords: Bearing Fault Detection; Mutual Information; Artificial Colony of Bees; Pattern Recognition.

Palavras-chaves: Detecção de Falhas de Rolamentos; Informação Mútua; Colônia Artificial de Abelhas; Reconhecimento de Padrões.

1. INTRODUÇÃO

As máquinas de indução são amplamente utilizadas nas aplicações industriais, devido a sua simples estrutura, robustez em ambientes agressivos e adaptabilidade a diversos tipos de cargas. Além disso, apresentam um baixo custo de aquisição e manutenção (Frosini et al., 2015). Entretanto, essas máquinas estão sujeitas a vários tipos de falhas elétricas e mecânicas. Dentre as falhas mecânicas, recentes trabalhos destacam as falhas de rolamentos, que representam de 40% a 52% de todas as falhas encontradas no motor de indução trifásico (MIT) (Frosini et al., 2015;

^{*} Este trabalho foi apoiado em parte pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) no âmbito dos processos 474290/2008-5, 473576/2011-2, 552269/2011-5, 201902/2015-0 e 405228/2016-3, pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Estado do Paraná no âmbito dos processos 338/2012 e 06/56093-3, pela Universidade Federal de Tecnológica do Paraná (UTFPR) e pelo Instituto Federal do Paraná (IFPR).

Leite et al., 2015). É possível que essas falhas influenciem na operação das máquinas, ocasionando paradas indesejadas. Geralmente, as falhas nos rolamentos estão relacionadas a contaminação, corrosão, lubrificação inapropriada e problemas de instalação (Leite et al., 2015).

Tendo em vista a grande possibilidade de ocorrência de falhas, atualmente, vários autores estão realizando estudos sobre o monitoramento das condições desses motores (Frosini et al., 2015; Leite et al., 2015). Foram propostas metodologias que identificam os defeitos incipientes relativos a esses motores, visto que o monitoramento apropriado oferece vantagens significativas aos processos industriais, como a prevenção de desligamentos inoportunos, redução nos custos de tempo de interrupção e reparos, e fornecimento de uma operação mais confiável dos motores. Assim, verifica-se que a disponibilidade de uma metodologia capaz de identificar defeitos incipientes é de grande interesse para a comunidade científica (Garcia-Escudero et al., 2017).

As metodologias de diagnóstico são classificadas pela variável de leitura analisada, pelas ferramentas de processamento de sinais e classificação de padrões. Em relação as variáveis empregadas, essas são divididas nas seguintes categorias: monitoramento de vibração, de temperatura, de emissão acústica, de corrente, de tensão, de potência ou de sinais magnéticos, e análise de óleo ou detritos (Trachi et al., 2016). No entanto, algumas dessas metodologias exigem sensores invasivos que dificultam a sua implementação, necessitando o acesso ao interior das máquinas, sendo esta uma tarefa difícil durante a operação. Assim, recentemente, as abordagens baseadas na medida da corrente do estator tem recebido uma grande atenção, por ser considerado um método não invasivo, além de possuírem sensores de relativo baixo custo e com tecnologia confiável que aproveite integralmente os benefícios do monitoramento das condições das máquinas de indução. Ainda, a utilização da corrente apresenta vantagens como o simples acesso, a facilidade de implementação e a não necessidade de sensores adicionais (Trachi et al., 2016).

As ferramentas de processamento de sinais colaboram no monitoramento das falhas, pois possibilitam a extração e seleção das características mais relevantes dos sinais. Na literatura, são encontrados metodologias tradicionais no domínio da frequência, as quais empregam a decomposição espectral dos sinais, como por exemplo a Transformada Fourier (TF) (Leite et al., 2015; Frosini et al., 2015; Trachi et al., 2016), a Transformada Wavelet (TW) (Trachi et al., 2016) e a Transformada Hilbert (TH) (Batista et al., 2016). Recentemente, alguns pesquisadores estão utilizando ferramentas mais avançadas como o algoritmo de classificação de sinais múltiplos (MUSIC - *MUltiple SIgnal Classification*) (Garcia-Calva et al., 2017) e as técnicas de invariância rotacional (ESPRIT - *Estimation of Signal Parameters via Rotational Invariance Techniques*) (Kim et al., 2013). Em oposição as metodologias no domínio da frequência, alguns trabalhos estão propondo abordagens alternativas, que utilizam sensores de baixo custo e realizam as análises dos sinais no domínio do tempo, como em Godoy et al. (2016) e Palácios et al. (2016).

A Teoria da Informação é uma outra metodologia que tem mostrado grande utilidade na área de monitoramento e detecção de falhas, pois tem a capacidade de determinar

a redundância entre as variáveis envolvidas no problema (Romero-Troncoso et al., 2011). Uma das ferramentas é a Informação Mútua (MI - *Mutual Information*) que é capaz de identificar as dependências lineares e não-lineares entre as séries temporais (Melia et al., 2015). Haja vista que o MIT é um sistema não-linear complexo, neste estudo é utilizado a IM como medida de similaridade entre as correntes, com o intuito de extrair as informações significantes dos sinais, possibilitando assim um eficiente diagnóstico de defeitos em rolamentos de MITs.

Semelhante as ferramentas de processamento de sinais, os Sistemas Inteligentes (SI) tornaram-se importantes no diagnóstico de falhas em máquinas elétricas, devido a sua eficiência em identificar o defeito sem a necessidade de modelos matemáticos complexos, tornando mais simples a implementação computacional, além de possibilitar a detecção da origem das falhas, contribuindo para uma operação mais confiável das máquinas. Dentre as ferramentas, pode-se citar as Árvores de Decisão (DT - *Decision Trees*), as Redes Neurais Artificiais (RNA), o *Support Vector Machine* (SVM), o *k*-Nearest Neighbor (*k*-NN), dentre outras, conforme Palácios et al. (2016), Godoy et al. (2016) e Bazan et al. (2017).

Ainda, recentemente, alguns pesquisadores de várias áreas têm desenvolvido em seus trabalhos ferramentas de inteligência de enxames na busca de soluções próximas do ideal para problemas de otimização complexos. Huang and Ai (2015) já empregaram a metodologia de otimização por Colônia Artificial de Abelhas (ABC - *Artificial Bee Colony*) na localização de falhas em *softwares* de computadores. Os resultados demonstraram que o método de localização de falhas proposto é promissor, tendo em vista que o algoritmo ABC pode eficientemente evitar o ótimo local e garantir a validade da localização da falha em todo o *software*. No âmbito do presente trabalho, essas ferramentas de otimização podem colaborar na seleção dos atributos mais relevantes da matriz de entrada dos classificadores de padrões, a fim de reduzir essa matriz.

Assim, a contribuição deste trabalho consiste na apresentação de uma metodologia alternativa para o diagnóstico de falhas de rolamentos em motores de indução trifásicos baseada na otimização da matriz de entrada dos classificadores de padrões por meio do algoritmo ABC. Serão selecionados os atributos mais relevantes das medidas da informação mútua entre os sinais de correntes obtidas na etapa de processamento de sinais. Por fim, realiza-se uma avaliação dos classificadores de padrões, Árvore de Decisão C4.5 e Rede Neural Artificial Multilayer Perceptron (MLP), que melhor identificam os espectros de defeitos extraídos na etapa de processamento de dados. A validação do modelo proposto é realizada por meio ensaios experimentais, onde são coletadas 438 amostras do motor operando em regime permanente senoidal e alimentados por meio de inversores de frequência. São emuladas as condições de operação que podem ser encontradas na indústria, como variação na frequência de alimentação do inversor (20% a 100% da frequência de alimentação nominal) e variação no conjugado de carga ($10\% \leq T_n \leq 120\%$, onde T_n é o torque nominal) para a obtenção dos resultados experimentais. Os níveis de severidade analisados, em ambos os casos, são de

15 a 90 minutos de desgaste dos rolamentos do motor de indução trifásico de 0.74 kW.

Este artigo está dividido em cinco seções. Na Seção 2 são abordados os principais conceitos das ferramentas utilizadas. Já na Seção 3, apresenta-se os materiais e métodos do trabalho. Os resultados experimentais são demonstrados na Seção 4 e, finalmente, na Seção 5 uma conclusão e discussões finais são realizadas.

2. EMBASAMENTO TEÓRICO

Nesta seção são abordadas as principais ferramentas utilizadas no sistema proposto de reconhecimento de padrões aplicado a desgaste abrasivo em rolamentos de MITs com alimentação não-senoidal. São descritos os aspectos gerais relacionados às etapas de extração de características, bem como de classificação de padrões de acordo com a metodologia proposta neste trabalho.

2.1 Informação Mútua

A MI é definida como a medida de dependência estocástica entre as variáveis aleatórias, possibilitando mensurar a informação comum entre elas. Ainda, aponta a redução no nível de incerteza associado a uma variável aleatória pela informação trazida por uma outra (Cover and Thomas, 2006). Tendo isso em vista, emprega-se esta ferramenta em sinais de máquinas elétricas, com o intuito de quantificar a associação entre as variáveis e possibilitar o monitoramento das condições de operação da máquina (Bazan et al., 2017). Neste trabalho, os sinais temporais analisados são as correntes das fases A e B do motor, as quais estão defasadas entre si. Assim, faz-se necessário utilizar uma outra abordagem conhecida como informação mútua atrasada (DMI - *Delayed Mutual Information*), que é uma grandeza que vem sendo utilizada em análises de estruturas dinâmicas de sistemas complexos, pois possibilita estimar a dependência entre as variáveis aleatórias, buscando obter a medida de similaridade em função de um deslocamento de tempo τ (Melia et al., 2015).

Após a abordagem dos conceitos relacionados à etapa de extração de características, são descritos os aspectos gerais associados ao otimizador heurístico e aos classificadores de padrões. Neste trabalho utiliza-se a árvore de decisão C4.5 e a rede neural artificial MLP, pois são técnicas amplamente aplicadas em pesquisa recentes, na área de mineração de dados, reconhecimento de padrões e identificação de falhas em máquinas elétricas (Palácios et al., 2016; Godoy et al., 2016).

2.2 Algoritmo por Colônia Artificial de Abelhas

O algoritmo ABC é baseado na meta-heurística de enxame e foi introduzido por Karaboga em 2005 para resolver problemas de otimização, baseando-se no comportamento inteligente das abelhas na natureza (Basturk and Karaboga, 2006). No ABC, a colônia de abelhas artificiais contém três grupos de abelhas: as operárias, associadas a fontes de alimentos específicos; as seguidoras, assistindo à dança das operárias dentro da colmeia para escolher uma fonte de alimento; e as campeiras, em busca de fontes de alimentos de forma aleatória. Tanto as seguidoras quanto as campeiras também são chamadas de abelhas desempregadas. O

algoritmo ABC trabalha da seguinte forma: inicialmente, todas as posições de fonte de alimentos são descobertas pelas campeiras. Posteriormente, o néctar de fontes de alimentos é explorado pelas operárias e seguidoras, e esta exploração contínua acabará levando ao esgotamento da fonte. Em seguida, as operárias que exploravam a fonte de alimento esgotada se tornam campeiras em busca, mais uma vez, de fontes alimentares. Analogamente no ABC, a posição de uma fonte de alimento representa uma possível solução para o problema e a quantidade de néctar de uma fonte de alimento corresponde à qualidade da solução associada. Assim, o número de abelhas operárias é igual ao número de soluções, uma vez que cada operária está associada com apenas uma fonte de alimento.

2.3 Árvore de Decisão C4.5

Segundo Fürnkranz et al. (2012), a árvore de decisão é um modelo de classificador cuja estrutura é constituída de um determinado número de nós e arcos, os quais representam as expressões ou regras que irão auxiliar na classificação das amostras nunca vistas. A classificação dessas amostras se dá ao percorrer toda a árvore de decisão, de cima para baixo (*Top-Down*). Uma das topologias idealizadas em Quinlan (1993) foi a árvore de decisão C4.5. Este algoritmo trabalha dividindo o conjunto total de dados de treinamento em subconjuntos, de maneira recursiva, utilizando-se de dados estatísticos como critérios de seleção. Assim, são calculados a Entropia, o Ganho de Informação e a Taxa de Ganho de Informação dos atributos de cada amostra presente no subconjunto de treinamento. De forma subsequente, o referido algoritmo busca minimizar a quantidade de informação necessária para a classificação de uma determinada amostra e garantir que uma árvore simplificada seja obtida.

2.4 Rede Neural Artificial MLP

A RNA *Multilayer Perceptron* (MLP) é composta ao menos de uma camada neural escondida, encontrada entre a entrada e a saída. Nesta topologia de rede, o fluxo de informações na estrutura da rede segue estritamente um caminho, o qual se inicia na camada de entrada e percorre, em seguida, as camadas intermediárias a de saída, respectivamente (Haykin, 2008; Silva et al., 2010). Esta topologia utiliza o algoritmo *backpropagation*, o qual possui dois estágios: *forward* e *backward*. O primeiro é um estágio supervisionado onde os vetores de entrada de uma amostra do conjunto de treinamento são propagados camada a camada até a produção das suas respectivas saídas, objetivando assim obter as respostas das redes. Nesta etapa, são considerados apenas os valores atuais dos pesos sinápticos e dos limiares dos neurônios. Como trata-se de um processo de aprendizagem supervisionado, as respostas produzidas pelas saídas da rede são comparadas com as suas respectivas respostas desejadas. Assim, em função do desvio entre o valor atual e o desejado, no segundo estágio do algoritmo *backpropagation* são ajustados os valores dos pesos sinápticos e dos limiares dos neurônios. Aplicando estes dois estágios sucessivamente, estes valores são alterados automaticamente em cada iteração do processo de treinamento, acarretando assim em uma diminuição gradual da soma dos erros produzidos pelas respostas da

rede frente aos seus respectivos valores desejados (Haykin, 2008; Silva et al., 2010).

3. ARQUITETURA PARA O DIAGNÓSTICO DE FALHAS

O MIT quando inserido em um ambiente industrial está sujeito a diversas condições de operação, como variação de carga, desequilíbrio de tensão, instalação em ambiente hostil, defeitos elétricos e/ou mecânicos, dentre outras (Palácios et al., 2016; Godoy et al., 2016; Bazan et al., 2017). O monitoramento adequado com o intuito de detectar os defeitos incipientes relativos a esses motores, traz vantagens importantes a um processo produtivo industrial, como a prevenção de desligamentos dispendiosos, que proporciona tempo suficiente para o desligamento controlado do processo, reduzindo assim os custos de tempo de interrupção e reparos, além de diagnosticar as necessidades atuais de manutenção, agendando a manutenção preventiva futura, reduzindo os custos com manutenção e fornecendo uma operação mais segura dos motores.

A fim de simular essas condições agressivas ao MIT, uma base de dados experimental foi criada por meio da aquisição dos sinais de corrente de um MIT operando sem falhas e com falhas de rolamento. Os níveis de severidade de falhas de rolamentos aplicados nos motores foi de 15, 30, 60 e 90 minutos. Ainda, com o intuito de emular os rolamentos em desgaste excessivo, foram reproduzidos defeitos de desgaste excessivos para avaliar a similaridade do processo de degradação do rolamento por meio do seu uso excessivo, falta de lubrificação e carga excessiva no eixo. Após a limpeza do rolamento, a lubrificação foi substituída por uma pasta abrasiva usada para desgastar as peças e o motor entrou em funcionamento pelos tempos previamente estabelecidos de 15, 30, 60 e 90 minutos para a degradação dos rolamentos. Após este processo, os rolamentos foram limpos e lubrificados novamente para a aquisição adequada dos dados.

Neste contexto, a base de dados experimentais consiste de 438 testes realizados no motor de 0,74 kW operando em regime permanente, sob a alimentação por meio de inversor de frequência, além diversas condições de operação, todas estas comuns a um ambiente industrial. Para os ensaios realizados, variou-se o nível de conjugado de carga de 0,5 N.m a 5,0 N.m, em intervalos de 0,5 N.m, e a frequência de alimentação, de 20% a 100% do valor nominal, com intervalos de 10%. As condições descritas anteriormente, foram coletadas para cada nível de desgaste excessivo nos rolamentos.

Para o presente estudo, os dados experimentais foram adquiridos utilizando a bancada experimental de ensaios apresentada na Figura 1, que foi projetada para o monitoramento de medidas elétricas e mecânicas, tais como tensão, corrente, vibração, torque e velocidade, além de permitir a simulação de diversas condições de operação como variação na alimentação e variação no conjugado de carga.

A bancada também permite a variação na alimentação do motor por meio de variadores de tensão independentes por fase e por meio da alteração na frequência de alimentação do inversor de frequência. Ainda, possibilita a imposição

Figura 1. Bancada Experimental de Ensaios

de conjugado no eixo do MIT, por este estar acoplado a um gerador de corrente contínua, além de possuir um torquímetro de dupla faixa de atuação com sensor de velocidade integrado da marca Kistler, os quais são responsáveis pela leitura dos sinais de torque e velocidade. Os sinais de correntes utilizados neste trabalho são coletados de forma individual por meio de sensores Hall, encarregado de condicionar os sinais das correntes que são transferidas às entradas analógicas da placa de aquisição de dados, DAQ USB-6221 da National Instruments, tal qual está conectada a um microcomputador.

A partir do algoritmo apropriado de importação de dados, os conjuntos descritos anteriormente são manipulados e avaliados no software Matlab®. O objetivo desta etapa consiste em extrair as informações presentes nos dados, buscando assim simplificar o número de entradas do sistema de classificação. A Figura 2 apresenta o procedimento de aquisição, processamento de sinais e classificação utilizados neste trabalho.

Figura 2. Fluxograma do sistema de identificação de defeitos de rolamentos proposto

Como os sinais das correntes analisadas estão defasados entre si, para conseguir verificar a medida de similaridade entre eles, utiliza-se a versão atrasada da informação mútua, previamente descrita na Seção 2.1. Tendo isso em vista, um algoritmo para o cálculo da DMI foi desenvolvido no Matlab®. Esse algoritmo trabalha com os sinais das correntes das fases, A e B, do motor, bem como com o deslocamento τ . A corrente i_b deve ser deslocada de 1 em 1 ponto até que se atinja o deslocamento em τ pontos. Para cada iteração, deve-se calcular as *pdfs* marginais das

correntes i_a e i_b , a *pdf* conjunta entre estas duas correntes, além do valor da MI. Esse valor da MI calculado é, então, armazenado em uma matriz de dados. Estes passos são realizados recursivamente até que se atinja o deslocamento τ . Por fim, é encontrado a DMI do sinal em análise. Utiliza-se um deslocamento de 150 pontos, pois observou-se que, com esta quantidade de pontos, as correntes estão em fase em pelo menos um ciclo. Assim, atenua-se o nível de incerteza associado a corrente da fase A devido à informação trazida pelo sinal de corrente da fase B. A Figura 3 exibe os sinais de correntes das fases A e B do motor para uma determinada condição de operação. Já a Figura 4, ilustra a sua respectiva MI encontrada.

Figura 3. Sinais de Correntes das Fases A e B - Motor operando sem falhas, com 5 N.m de carga em seu eixo e com as tensões equilibradas

Figura 4. Análise da variação no nível de desequilíbrio - Motor operando sem falhas, com 5 N.m de carga em seu eixo e com as tensões equilibradas

Após finalizar o processamento dos dados de todas as amostras, deve-se então normalizar os valores entre 0 e 1, em função do valor máximo da informação mútua (MI_{max}) encontrado para todas as situações de operação. Realizado a normalização dos dados, seleciona-se então os atributos mais relevantes da informação mútua por meio do algoritmo ABC, objetivando reduzir a matriz de entrada dos classificadores de padrões C4.5 e MLP. Nessa seleção, emprega-se o algoritmo Seletor de Características baseada na Correlação (*Correlation-based on Feature Selector* - CFS), uma ferramenta que correlaciona os atributos do conjunto de dados baseado em uma função heurística. O limiar de ativação desta função tem como particularidade atuar em subconjuntos altamente correlacionados com uma classe, contudo não relacionados entre si. Assim, tanto os atributos insignificantes quanto os redundantes devem ser descartados, por apresentarem baixa correlação

com a classe em estudo e por serem, provavelmente, altamente relacionados com um ou mais atributos restantes. Logo, a seleção de um atributo depende apenas de como ele irá predizer as classes ainda não exploradas por outros atributos. A Equação 1 define o cálculo de avaliação de subconjuntos da ferramenta CFS.

$$M_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \quad (1)$$

onde M_s é definido como sendo o *merit* de um subconjunto S contendo k características, r_{cf} é a correlação principal entre as características de classe, enquanto que r_{ff} é a média entre as correlações entre as características.

Por fim, maximizando a função descrita em 1, obtém-se a matriz de entrada do sistema de classificação de falhas no rolamento de MIT. Os dados dessa matriz são rotulados de acordo com sua classe, as amostras de sinais do motor saudável recebem uma *tag* de saída desejada igual a 0 e, os sinais com falhas de rolamentos, uma *tag* de saída desejada igual a 1.

4. RESULTADOS EXPERIMENTAIS

O conjunto das amostras da matriz de entrada são aleatoriamente separadas, de forma a permitir que o conjunto de dados de treinamento seja representativo o suficiente em relação à base de dados na qual se deseja prever. Tendo isso em vista, neste estudo utiliza-se o método de validação cruzada de k -partições, com $k = 10$. O processo é realizado por k vezes, sendo que em cada iteração, 1 subconjunto é designado como conjunto de teste e os outros $(k-1)$ subconjuntos restantes, designados como conjunto de treinamento. A cada iteração, a taxa de classificação incorreta do conjunto de teste é calculada, sendo que ao final, o desempenho é estimado pelo erro médio do classificador nos k testes realizados.

Para a realização dos ensaios experimentais, utilizou-se o *software Waikato Environment for Knowledge Analysis* (WEKA), que permite a resolução de problemas concretos de mineração de dados e classificação de padrões por meio do conjunto de algoritmos que nele estão disponíveis, disponibilizando os classificadores apresentados na Seção 2 (Hall et al., 2009).

Na seleção dos atributos de maior relevância, empregou-se o algoritmo ABC. Optou-se por esses seguintes parâmetros: população inicial de 30 abelhas, número máximo de 20 iterações, tipo de mutação bit-flip e probabilidade de mutação de 0,01. Já na etapa de classificação das amostras saudáveis e defeituosas, utilizou-se a rede neural MLP e a árvore de decisão C4.5, descritas na Seção 2. Para a Rede Artificial MLP, a taxa de aprendizagem é definida em 0,3, o termo momentum recebe 0,2 e o número máximo de épocas para o treinamento da rede é 500. O número de neurônios na camada intermediária é de 10 neurônios (empiricamente definido). Para a camada intermediária e de saída são utilizados a função de ativação logística e linear, respectivamente. Na construção do modelo de árvore de decisão C4.5, não são consideradas divisões binárias nos atributos nominais. Uma semente com valor 1 foi utilizada para randomizar as amostras. O número mínimo de atributos por folha recebe um valor de 2. De modo a construir uma

árvore de decisão mais simplificada e generalista, utiliza-se o método de poda. O fator de confiança da poda é ajustado em 0,25, e a quantidade de dados utilizados para a redução de erros na poda é fixado como 3.

4.1 Resultados Experimentais sem o ABC

Nesta Seção são apresentados os resultados experimentais dos testes para o motor de 0,74 kW conectado por meio de inversor de frequência sem a utilização do algoritmo ABC para a seleção dos atributos mais relevantes. Assim, utilizam-se como dados de entrada para os classificadores: Árvore de Decisão C4.5 e Rede Neural MLP, todos os valores da curva de DMI (150 atributos) obtidos na etapa de extração de características. Conforme descrito anteriormente, são emuladas diversas condições de operação encontradas em um ambiente industrial, como variação na frequência de alimentação, variação no conjugado de carga e desgaste abrasivo nos rolamentos de 15, 30, 60 e 90 minutos. A Tabela 1 apresenta os resultados de classificação para esse primeiro caso.

Tabela 1. Resultados experimentais utilizando toda a matriz de entrada da DMI - Motor 0,74 kW

Dados	Motor 0,74 kW	
Classificador	C4.5	MLP
Acurácia Global (%)	94,3	99,8
Acurácia Normal (%)	85,1	100
Acurácia Falha (%)	96,6	99,7
Índice Kappa	0,82	0,99
Construção do Modelo (s)	0,40	4,99

A partir da Tabela 1, analisa-se que ao ser utilizado toda a matriz de entrada da DMI, atinge-se o melhor resultado utilizando o classificador MLP. Observa-se que são classificadas corretamente 99,8% do conjunto global de amostras, 100% do conjunto de amostras saudáveis e 99,7% do conjunto de amostras defeituosas. Ainda, destaca-se o índice Kappa de 0,99 que confirma os resultados descritos anteriormente. Isso demonstra que utilizando a informação mútua combinada com a rede MLP é possível fazer uma boa separação entre as classes sem defeitos e com defeitos. Já quando utiliza-se a árvore C4.5, verifica-se que são atingidos resultados inferiores, no entanto, identificam-se ainda corretamente 94,3% de todas as amostras. É importante destacar que o classificador C4.5 despende de um tempo de 0,40 segundos para a construção do modelo, o qual é menor quando comparado com a rede MLP que despende de 4,99 segundos. Assim, com os resultados satisfatórios de classificação, conforme apresentado na Tabela 1, a árvore de decisão C4.5 juntamente com a DMI também apresenta-se como uma promissora proposta para o diagnóstico de falhas de rolamentos de MIT alimentado por inversor de frequência.

4.2 Resultados Experimentais Utilizando o ABC

Nesta Seção, apresentam-se os resultados experimentais dos testes para o motor de 0,74 kW quando para a identificação das falhas são utilizados classificadores meta-heurísticos. Semelhante aos ensaios realizados na Seção 4.1, efetua-se o teste utilizando todo o conjunto de dados do motor, ou seja, com variação na frequência de alimentação de 20% à 100% do da frequência nominal e uma

variação no conjugado de carga de 10% a 120% de seu valor nominal.

Conforme descrito na Seção 3, ao empregar o algoritmo CFS presente no ABC, que utiliza a meta-heurística *merit* para a predição das classes, os 150 atributos da DMI são avaliados individualmente. Os atributos que apresentam os maiores *merits* são identificados como os mais relevantes para a separação entre as classes saudáveis e defeituosas, assim, são utilizados como dados para a construção do modelo classificador. Foram identificados 23 atributos: 29, 51, 55, 60, 66, 68, 82, 84, 86, 91, 92, 96, 106, 108, 110, 113, 114, 116, 119, 120, 121, 142 e 144, os quais compõem a matriz de entrada de 150 atributos da curva da DMI exemplificada na Figura 4. A Tabela 2 apresenta os resultados de classificação utilizando os classificadores meta-heurísticos, para a condição em que o motor está sujeito ao desgaste abrasivo nos rolamentos de 15, 30, 60 e 90 minutos, além de ser alimentado por uma fonte não-senoidal.

Tabela 2. Resultados experimentais utilizando o ABC na seleção dos atributos da DMI - Motor 0,74 kW

Dados	Motor 0,74 kW	
Classificador	C4.5	MLP
Acurácia Global (%)	93,6	98,6
Acurácia Normal (%)	81,6	96,6
Acurácia Falha (%)	96,6	99,1
Índice Kappa	0,80	0,96
Construção do Modelo (s)	0,22	1,31

Ao utilizar o ABC na seleção dos atributos, verifica-se que houve uma diminuição de 0,70% e 1,20% de acurácias para os classificadores C4.5 e MLP, respectivamente, em relação aos testes apresentados na Seção 4.1. No entanto, ainda são encontradas satisfatórias acurácias de classificação, como pode ser visualizado na Tabela 2. Empregando o classificador meta-heurístico MLP são identificadas corretamente 98,6%, 96,6% e 99,1% das amostras globais, saudáveis e defeituosas, respectivamente. Nessa situação, o coeficiente Kappa de 0,96 ratifica os resultados atingidos. Já quando utiliza-se o classificador meta-heurístico C4.5, identificam-se 93,6% de todas as amostras, 81,6% das amostras sem falhas e 96,6% das amostras com falhas. O índice Kappa de 0,80 confirma a total concordância com os resultados, haja visto que valores iguais ou superiores a 0,80 indicam essa relação. Novamente, como na Seção 4.1, o classificador MLP apresenta uma melhor identificação das amostras. Destaca-se que utilizando o ABC na seleção dos atributos, obteve-se acurácias similares e com um gasto de tempo computacional menor quando comparado aos ensaios anteriores. Nestes novos ensaios, os tempos para as construções dos modelos foram de 0,22 e 1,31 segundos para os classificadores meta-heurísticos C4.5 e MLP, respectivamente. Assim, pensando no âmbito computacional, essas ferramentas apresentam ser mais adequadas para o diagnóstico de falhas de rolamentos em MIT, uma vez que encontram-se resultados satisfatórios com um custo computacional mais baixo.

Pelos resultados apresentados nas Seções 4.1 e 4.2, demonstra-se que o sistema proposto apresenta promissor desempenho na identificação de defeitos de rolamentos em MITs alimentados por fonte de alimentação não-senoidal

independente da faixa de velocidade e de carga em que atua.

5. CONSIDERAÇÕES FINAIS

Este estudo apresentou uma abordagem alternativa para o diagnóstico de defeitos de desgaste abrasivo em rolamentos de MITs, a qual é baseada em medidas de informação mútua, algoritmo por colônia artificial de abelhas e sistemas inteligentes. Os ensaios experimentais foram realizados utilizando um conjunto de dados composto por vários sinais do MIT sujeito a diferentes condições de operação, como variação no conjugado de carga e na frequência de alimentação e também desgaste abrasivo de rolamentos.

Assim, na etapa de extração de características, utilizou-se a DMI para calcular as medidas de similaridade entre os sinais de corrente das fases A e B do motor, a fim de extrair as características de falhas. Após, realizou-se a avaliação dos classificadores de padrões (C4.5 e MLP) que é mais adequado a esse tipo de dados. Os resultados confirmaram que a rede neural artificial MLP é a mais compatível com a abordagem de extração de características, pois propicia resultados promissores de classificação das amostras, independente das circunstâncias de operação do motor. No entanto, é importante destacar que com a árvore de decisão C4.5, a metodologia também apresenta resultados satisfatórios, atingindo uma taxa de classificação correta de 94,3% de todas as amostras.

Quando utiliza-se o algoritmo ABC na seleção dos atributos mais relevantes da curva da DMI, verifica-se novamente que os classificadores meta-heurísticos C4.5 e MLP apresentam resultados satisfatórios, superiores a 93,6%. No entanto, observa-se uma melhor separação entre as classes sem defeitos e com defeitos. Ao utilizar esse classificador, são selecionados os 23 atributos mais relevantes da curva da DMI de 150 atributos, atingindo resultados próximos ao dos testes sem o algoritmo ABC. Isso demonstra que ao empregar o algoritmo ABC para a seleção dos atributos, pode-se obter resultados satisfatórios com um custo computacional inferior, uma vez que para a construção do modelo classificador é gasto um tempo menor.

REFERÊNCIAS

- Basturk, B. and Karaboga, D. (2006). An artificial bee colony (abc) algorithm for numeric function optimization. In *IEEE Swarm Intelligence Symposium 2006*. Indianapolis, Indiana, USA.
- Batista, F.B. et al. (2016). An empirical demodulation for electrical fault detection in induction motors. *IEEE Transactions on Instrumentation and Measurement*, 65, 559–569.
- Bazan, G.H., Scalassara, P.R., Endo, W., Goedel, A., Godoy, W.F., and Palácios, R.H.C. (2017). Stator fault analysis of three-phase induction motors using information measures and artificial neural networks. *Electric Power Systems Research*, 143, 347 – 356.
- Cover, T. and Thomas, J. (2006). *Elements of Information Theory*. Wiley-Interscience., New York, 2nd edition.
- Frosini, L., Harliska, C., and Szabo, L. (2015). Induction machine bearing fault detection by means of statistical processing of the stray flux measurement. *IEEE Transactions on Industrial Electronics*, 62(3), 1846–1854.
- Fürnkranz, J., Gamberger, D., and Lavrac, N. (2012). *Foundations of Rule Learning*. Springer-Verlag Berlin, Heidelberg.
- Garcia-Calva, T.A., Morinigo-Sotelo, D., and de Jesus Romero-Troncoso, R. (2017). Non-uniform time resampling for diagnosing broken rotor bars in inverter-fed induction motors. *IEEE Transactions on Industrial Electronics*, 64(3), 2306–2315.
- Garcia-Escudero, L.A., Duque-Perez, O., Fernandez-Temprano, M., and Morinigo-Sotelo, D. (2017). Robust detection of incipient faults in vsi-fed induction motors using quality control charts. *IEEE Transactions on Industry Applications*, 53(3), 3076–3085.
- Godoy, W.F., da Silva, I.N., Goedel, A., Palácios, R.H.C., and Lopes, T.D. (2016). Application of intelligent tools to detect and classify broken rotor bars in three-phase induction motors fed by an inverter. *IET Electric Power Applications*, 10(5), 430–439.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I.H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explor. Newsl.*, 11(1), 10–18.
- Haykin, S. (2008). *Neural Networks and Learning Machines*. Pearson, New Jersey, 3rd edition.
- Huang, L. and Ai, J. (2015). Automatic software fault localization based on artificial bee colony. *Journal of Systems Engineering and Electronics*, 26(6), 1325 – 1332.
- Kim, Y.H., Youn, Y.W., Hwang, D.H., Sun, J.H., and Kang, D.S. (2013). High-resolution parameter estimation method to identify broken rotor bar faults in induction motors. *IEEE Transactions on Industrial Electronics*, 60(9), 4103–4117.
- Leite, V.C.M.N., da Silva, J.G.B., Veloso, G.F.C., da Silva, L.E.B., Lambert-Torres, G., Bonaldi, E.L., and d. L. de Oliveira, L.E. (2015). Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. *IEEE Transactions on Industrial Electronics*, 62(3), 1855–1865.
- Melia, U. et al. (2015). Mutual information measures applied to EEG signals for sleepiness characterization. *Medical Engineering and Physics*, 37(3), 297–308.
- Palácios, R.H.C., da Silva, I.N., Goedel, A., and Godoy, W.F. (2016). A novel multi-agent approach to identify faults in line connected three-phase induction motors. *Applied Soft Computing*, 45, 1–10.
- Quinlan, J. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Francisco.
- Romero-Troncoso, R.J. et al. (2011). FPGA-based online detection of multiple combined faults in induction motors through information entropy and fuzzy inference. *IEEE Transactions on Industrial Electronics*, 58, 5263–5270.
- Silva, I.N., Spatti, D.H., and Flauzino, R.A. (2010). *Redes Neurais Artificiais: para engenharia e ciências aplicadas*. Artliber, São Paulo, SP, Brasil.
- Trachi, Y., Elbouchikhi, E., Choqueuse, V., and Benbouzid, M.E.H. (2016). Induction machines fault detection based on subspace spectral estimation. *IEEE Transactions on Industrial Electronics*, 63(9), 5641–5651.