

APLICAÇÃO DO ALGORITMO RANDOM FOREST COMO CLASSIFICADOR DE PADRÕES DE FALHAS EM ROLAMENTOS DE MOTORES DE INDUÇÃO

TIAGO DRUMMOND LOPES*, ALESSANDRO GOEDEL†, RODRIGO H. CUNHA PALÁCIOS†, WAGNER FONTES GODOY†

**Instituto Federal de Santa Catarina, Departamento de Eletroeletrônica, Itajaí, Santa Catarina, Brasil*

†*Universidade Tecnológica Federal do Paraná, Departamento de Engenharia Elétrica e Computação, Cornélio Procopio, Paraná, Brasil*

Emails: tiago.drummond@ifsc.edu.br, agoedel@utfpr.edu.br,
rodrigopalacios@utfpr.edu.br, wagnergodoy@utfpr.edu.br

Abstract— Due to its robustness and its convenient relation power-volume, three-phase induction motors have dominated the field of electromechanical energy conversion. However, considering industrial application, these machines are subject to stresses and operating regimes, which can lead to faults in its different parts. Many of these defects can occur in their bearings parts. Several patterns classifiers methods have been widely discussed in the literature and they are able to predict such defects. The purpose of this study is to apply the random forest algorithm for identifying bearings defects in induction motors. The experimental results were collect from tests carried out in a workbench employing motors operating with faulty bearings, under varying conditions of load torque and voltage unbalance, considering the analysis of two currents signals in the time domain.

Keywords— Bearing Fault Detection, Current Sensor, Random Forests, Three-phase Induction Motor.

Resumo— Graças à sua robustez e a sua conveniente relação potência-volume, os motores de indução trifásicos têm dominado o campo de conversão de energia eletromecânica. No entanto, na aplicação industrial, essas máquinas são sujeitas a esforços e regimes de operação que podem proporcionar falhas nas suas diferentes partes. Grande parte desses defeitos nos motores ocorrem em seus rolamentos. Diversos métodos classificadores de padrões são capazes de prever os defeitos dos motores elétricos e têm sido amplamente abordados na literatura. A proposta desse trabalho consiste em utilizar o algoritmo random forest para identificar defeitos em rolamentos de motores de indução. Os resultados experimentais são apresentados a partir de ensaios realizados em bancada com motores operando com rolamentos defeituosos, em condições variadas de conjugado de carga e desequilíbrio de tensão, possibilitando a análise do sinal de duas das três correntes no domínio do tempo.

Palavras-chave— Detecção de Falhas em Rolamentos, Florestas Aleatórias, Motor de Indução Trifásico, Sensor de Corrente.

1 Introdução

O Motor de Indução Trifásico (MIT) é, dentre as máquinas elétricas girantes, a mais utilizada no setor industrial. Este tipo de motor possui como suas principais características, a elevada eficiência, robustez, baixa manutenção e uma conveniente relação potência-volume, quando comparado aos demais motores. Tais características são responsáveis por tornar o MIT tão atraente para a conversão de energia elétrica em energia mecânica motriz (Gonzalez et al., 2014; Popescu et al., 2013; Keskes et al., 2013).

A constatação da vasta utilização do MIT pode ser realizada por meio de dados apresentados no balanço energético nacional apresentado pela ELETROBRÁS-PROCEL (2014). Esse relatório apresenta uma estimativa de que por ano são destinados 210,10 TWh (Tera Watt-hora) às indústrias, representando 34,4% da energia elétrica produzida no Brasil. Tendo em vista a pouca aplicação do MIT em ambientes não industriais, pode-se inferir que a parcela de consumo energético do mesmo é superior a 24% da produção nacional de energia elétrica.

Na aplicação industrial, o motor de indução é submetido a um esforço inevitável e embora habi-

tualmente seja bem construído, pode estar sujeito a falhas geradas, por exemplo, a partir da umidade e poeira do ambiente, problemas de qualidade de energia elétrica, sobreaquecimento, etc. Estas falhas podem ser classificadas em dois grupos: falhas elétricas e falhas mecânicas. Problemas relativos ao enrolamento de estator, o enrolamento de rotor, as barras quebradas no rotor, os anéis quebrados no rotor e conexões elétricas são perturbações que caracterizam as falhas elétricas. Sob outra perspectiva, as falhas mecânicas destacam-se por problemas de rolamentos, excentricidade, desgaste de acoplamento e desalinhamento (Bellini et al., 2008).

Outros problemas relacionados aos motores, como os providos pelos componentes de acionamento e de potência ligados à máquina ou a má qualidade da energia elétrica, também podem ocasionar problemas no funcionamento deste equipamento. Entretanto, estes são considerados externos aos motores e não são tratados como falha do mesmo. Estudos detalhados apresentados por Zarei et al. (2014) e Bellini et al. (2008) mostram que entre os principais elementos deteriorados nos MIT estão os rolamentos com mais de 40% dos casos de falhas.

Tais defeitos causam prejuízos imponderáveis

no ambiente industrial, pois essas falhas podem acarretar a interrupção inesperada da operação do motor e conseqüentemente a paralisação de processos fabris. Dessa forma, a confiabilidade e a continuidade de serviço das máquinas são os fatores que precisam ser levados em conta. Para alcançar estes objetivos, um sistema eficaz para a detecção e diagnóstico de falhas do motor se torna essencial (Seera et al., 2014). No entanto, o processo de identificação precoce de defeitos, embora muito importante, não é uma tarefa simples. Tal fato, torna atrativo o desenvolvimento de pesquisas com o intuito de detectar avarias nas máquinas antes que as mesmas entrem em colapso.

A verificação direta ou indireta do estado da máquina pode ser realizada mediante leitura de uma das fases na corrente elétrica de alimentação ou pela tensão elétrica, monitoramento do campo eletromagnético, medidas de temperatura ou por outros parâmetros (Ghate and Dudul, 2010). Deste modo, a interpretação dessas grandezas mostra as alterações de comportamento do equipamento contribuindo na tarefa de detecção de defeitos ainda em fase de evolução. Assim, uma operação corretiva pode ser planejada de forma a atuar no processo reduzindo os impactos e evitando uma parada inesperada (Araújo et al., 2011).

Nos últimos anos, a academia e a indústria vêm mostrando grande interesse nos estudos e pesquisas relacionadas às mais variadas técnicas de detecção de defeitos em máquinas elétricas. Nesse sentido, os sistemas inteligentes têm sido desenvolvidos como uma abordagem flexível que pode ser aplicada a vários tipos de problemas relacionados a máquinas elétricas, em particular aqueles associados com a identificação de padrões de falha. Dentre esses sistemas, Random Forest (RF) é um método de aprendizado amplamente utilizado para classificar padrões.

Conforme relatado na literatura (Seera et al., 2014), ao monitorar e comparar os sinais de motores com defeito com aqueles a partir de um sinal sem perturbações, falhas incipientes podem ser detectadas e localizadas. Pode-se verificar tal fato nos estudos de Leite et al. (2015) e Harlisca et al. (2013), onde o primeiro demonstra a detecção de falha de rolamento de um MIT analisando o espectro da corrente do estator. Algoritmos baseados em Spectral Kurtosis, Kurtogram rápido e o Kurtogram Wavelet, também são aplicados para melhorar a análise. Testes experimentais são realizados considerando as falhas em diferentes fases de evolução e os resultados são promissores. Já no segundo estudo, a detecção de três tipos de falhas de rolamentos por meio de processamento estatístico das medições de fluxo é detalhada. O método não invasivo desenvolvido requer apenas duas sondas e alguns cálculos.

A pesquisa de Son et al. (2009) apresenta um

sistema de sensores inteligentes que adquire sinais de vibração, corrente e fluxo do motor de indução. Os sinais alimentam classificadores de padrões com o intuito de diagnosticar falhas de rolamento, rotor e estator. O desempenho dos classificadores Random Forest, Support Vector Machine e Linear Discriminant é comparado. Segundo os autores, os resultados experimentais mostram que o sistema de sensores inteligentes tem o desempenho similar a técnicas tradicionais de diagnóstico de falhas em máquinas, porém com o custo reduzido.

É proposta no presente trabalho a verificação do desempenho do método Random Forest, aplicado na detecção de defeitos no rolamento de MIT. O classificador é alimentado com sinais reais de corrente do estator da máquina oriundos de experimentos realizados em ambiente laboratorial. Desequilíbrios de tensão na alimentação da máquina, ampla variação do perfil de carga no eixo do motor e diversos defeitos nos rolamentos (defeitos nas pistas externa e interna, desgaste excessivo e descarga elétrica por meio de curto circuito) são considerados nos ensaios.

O artigo está dividido nas seguintes seções: o método classificador utilizado na pesquisa é descrito na Seção II; na Seção III é exibido o sistema proposto neste trabalho e na Seção IV são apresentados os resultados experimentais. Finalmente, na Seção V expõem-se as conclusões deste artigo. a Seção II descreve

2 Random Forest

Na área de aprendizado de máquinas espera-se que a combinação dos resultados de vários classificadores apresente melhor desempenho e maior convicção na tomada de decisão do que apenas um classificador. Ainda, segundo Sirikulviriyaya and Sinthupinyo (2011), dada uma mesma quantidade de informação para treinamento, é amplamente aceito que o desempenho e a confiabilidade de muitos classificadores é geralmente melhor do que um único classificador. Dessa forma, manifestam-se um abundante interesse na pesquisa e exploração de métodos *ensemble*, os quais caracterizam-se pela geração de muitos classificadores e combinação de seus resultados (Dietterich, 2000). Como exemplos clássicos de métodos *ensemble* pode-se citar: *boosting*, *bagging* e recentemente *random forest*.

O algoritmo de Random Forest (RF) introduzido por Breiman (2001) é um termo geral para métodos de *ensemble* utilizando classificadores do tipo árvore. A RF constrói uma grande quantidade de árvores de decisão para fora do sub-conjunto de dados a partir de um treinamento único definido. Tal treinamento é realizado usando *bagging* (um meta-algoritmo para melhorar a classificação e a regressão de modelos de

acordo com a estabilidade e a precisão da classificação).

A utilização de *bagging* no treinamento além de reduzir a variância ajuda a evitar o *over-fitting*. Este procedimento extrai casos aleatoriamente a partir de conjuntos de dados de treinamento originais e os conjuntos são usados para construir cada uma das árvores de decisão que compõe a RF. Cada árvore classificadora é apontada como um componente preditor. A RF constrói sua decisão por meio da contagem dos votos dos componentes preditores em cada classe e, em seguida, seleciona a classe vencedora em termos de número de votos acumulados.

Assim, todo o algoritmo inclui duas fases importantes: o período de formação de cada árvore e o período de votação. A primeira fase consiste em treinar cada árvore de decisão, e os sub-conjuntos de dados são selecionados a partir do conjunto de dados de treinamento e definidos usando a estratégia *bagging* aleatória. Em seguida, os dados de teste são classificados pela votação majoritária. Cerca de um terço dos casos são deixados de fora das amostras de *bootstrap* e não são usados na construção de uma árvore em particular. As amostras deixadas de fora da árvore de ordem k são executadas através da k -ésima árvore para obter uma classificação. Desta forma, uma classificação é obtida para cada caso do conjunto de testes. Cerca de um terço das árvores podem ser utilizadas para avaliar a precisão do classificador (Son et al., 2009).

A abordagem de Yang et al. (2015) utiliza a técnica Multiple Class Feature Selection (MCFS) para selecionar as características que são usadas para treinar e testar o classificador Random Forest. A fim de verificar a eficácia do algoritmo MCFS para diagnóstico de falhas de motores de indução, foram realizados experimentos em um emulador de falha da máquina. Com este sistema, os sinais de corrente do estator e de vibração podem ser adquiridos quando o motor funciona sob seis condições diferentes (motor normal, curto circuito no estator, rotor desbalanceado, defeito de pista interna e defeito de pista externa de rolamento, quebra de barra do rotor e rotor descentralizado).

O trabalho de Patel and Giri (2016) explora experimentalmente o desenvolvimento de um classificador Random Forest, uma técnica de aprendizado de máquina que surgiu recentemente, para o diagnóstico de falha mecânica multi-classe de um motor de indução. Em primeiro lugar, os sinais de vibração são colhidos a partir do rolamento utilizando um acelerômetro. Parâmetros do sinal de vibração são extraídos sob a forma de características estatísticas e usado como entrada para o classificador. Esses padrões de rolamentos defeituosos são classificados pelo algoritmo RF em quatro classes (motor sem falhas, defeito na esfera, pista interna e externa do rolamento).

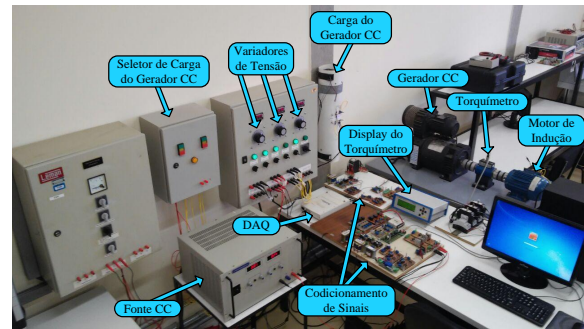


Figura 1: Estrutura de ensaios

Neste sentido, este trabalho apresenta um estudo do desempenho do classificador de padrões Random Forest frente ao problema de detecção de falhas em rolamentos de motores de indução trifásico.

3 Aquisição e Tratamento de Dados

A metodologia proposta nesta pesquisa consiste em criar um banco de dados com sinais de correntes de motores de indução operando com rolamentos novos e com rolamentos defeituosos. Posteriormente usar os sinais amostrados de duas das três correntes da máquina para alimentar o classificador RF, o qual indica se a máquina possui ou não defeitos no rolamento. Para isso, dois motores do fabricante WEG, ambos de 1 cv e alto rendimento foram utilizados para realização dos ensaios, sendo o Motor 1 e o Motor 2 com respectivos rendimentos de 80,5% e 82,6%. Essas máquinas operaram com rolamentos com condições operacionais normais e defeituosos, em condições de desequilíbrio de tensão e variação de carga aplicada no eixo do motor. Tal procedimento foi realizado em uma bancada, a qual visa emular o comportamento de um MIT frente a diferentes pontos de operação, e assim reproduzir as mesmas condições encontradas em ambiente industrial.

A estrutura de ensaios apresentada na Fig. 1 baseia-se em uma máquina de corrente contínua, um torquímetro girante, um quadro de acionamento de máquinas elétricas com variadores de tensão e sensores de efeito Hall de corrente.

O funcionamento da bancada resume-se em acoplar a Máquina de Corrente Contínua (MCC) ao MIT, através do mesmo eixo. Assim, com a MCC ligado para operar como um Gerador de Corrente Contínua (GCC) é possível atuar na variação da tensão de campo, por meio de uma fonte de corrente contínua, e, conseqüentemente, obter variação no torque de carga do MIT. A energia gerada pelo GCC no processo é dissipada em uma carga resistiva ligada à armadura da máquina. O torquímetro é responsável por fornecer a velocidade e torque imposto ao MIT e os sensores de efeito Hall são utilizados para mensurar a cor-

rente. O motor é submetido a diferentes tipos de alimentação através de variadores de tensão (Variacs) conectados diretamente à rede, para que sejam obtidos os diferentes pontos de operação da máquina.

Tal processo de aquisição e tratamento de dados inicia-se com os sensores de efeito Hall aferindo os sinais senoidais de corrente da máquina. Para que os sinais possam ser armazenados em um microcomputador (PC), utiliza-se a placa de aquisição de dados (DAQ) da fabricante National Instruments modelo NI USB6221, que através de uma porta USB faz a comunicação entre os sensores de corrente e o PC. O software Matlab é encarregado de registrar os dados, organizá-los em matrizes e processá-los.

De modo a permitir que os sinais sejam observados pelo algoritmo RF no domínio do tempo prevê-se o tratamento dos dados baseado na técnica que encontra-se respaldada na proposta de (do Nascimento et al., 2011), a qual utiliza este método para estimação dos valores de harmônicos presentes nos sinais senoidais de corrente elétrica de cargas não lineares. Recentemente, os trabalhos Palácios et al. (2014), Godoy et al. (2014), e Broniera et al. (2013) também basearam-se na mesma técnica de pré-processamento apresentando resultados promissores.

Inicialmente, os sinais referentes a duas das três fases de correntes do motor são somados vetorialmente e em seguida o algoritmo de tratamento de dados é aplicado. Esse processo consiste em discretizar o sinal com o objetivo de reduzir o número de amostras por semi ciclo. No caso deste trabalho foi sub-amostrado em 25 pontos por semiciclo. Para simplificar a aplicação da random forest, usa-se o módulo do sinal, evitando valores negativos. Na sequência do processamento, ocorre a normalização pelo valor de pico de cada sinal amostrado, ou seja, os valores ficam no intervalo entre zero e um. Por fim, as matrizes a serem apresentadas como entradas dos classificadores são construídas.

A fim de obter sinais de corrente de MIT com rolamentos defeituosos, pares de rolamentos novos foram submetidos a condições que os tornassem falhos. Quatro tipos de defeitos foram recriados de forma controlada, tais como: defeito na pista externa, falha na pista interna, curto circuito e desgaste excessivo.

4 Resultados Experimentais

Os resultados experimentais desse estudo são apresentados nessa seção, a qual visa mostrar o desempenho da utilização do algoritmo Random Forest para classificar falhas de rolamento de motores de indução.

Os ensaios experimentais visam reproduzir em laboratório as condições de operação da máquina

no ambiente industrial. Assim, neste trabalho emulam-se problemas de qualidade de energia por meio dos desequilíbrios de tensão e combinadamente varia-se o conjugado de carga aplicado ao eixo da máquina, a fim de amostrar os sinais de corrente do MIT nessas condições de operação.

Para obter a classificação quanto ao funcionamento adequado do rolamento, foram amostrados os sinais das correntes trifásicas dos MIT em análise. Para tanto, foram separados os dados para as construções dos modelos inteligentes, utilizando os sinais resultantes da soma das correntes I_A , I_B e I_C duas a duas, ou seja, $I_A + I_B$, $I_B + I_C$ e $I_A + I_C$. Salienta-se que as amostras são sinais dos motores em regime permanente, adquiridas em ensaios na bancada experimental, com conjugado de carga variando de 10% a 150% do torque nominal e com 16 combinações de desequilíbrio de tensão entre as fases, considerando o limite máximo de $\pm 10\%$. A partir dessas características construiu-se um banco de dados de 275 ensaios, sendo que em 102 experimentos o motor operou com rolamentos sem defeitos e em 173 defeituosos, a saber que considerou-se quatro tipos de rolamentos defeituosos, quais sejam: pista externa, pista interna, desgaste excessivo e curto-circuito.

Com o intuito de obter resultados expressivos usou-se a seguinte metodologia para definir os parâmetros do classificador RF. Para determinar o número de árvores para a RF, testes foram realizados considerando 5, 10, 25, 50, 100, 200, 300, 400 e 500 árvores. No processo de construção das estruturas classificadoras utilizou-se validação cruzada com 10 partições e levou-se, também, em conta o erro *out-of-bag*. Os melhores resultados foram obtidos considerando 200 árvores. Para os testes realizados neste trabalho, variou-se o número de sementes de 1 a 5 e o número de características aleatórias foi variado de 1 a 15. Com base nos resultados obtidos, o número de sementes foi definido como 1 e número de características aleatórias definido como 6.

Os classificadores treinados e validados possuem 25 entradas obtidas da soma de dois sinais de corrente ($I_A + I_B$, $I_B + I_C$ e $I_A + I_C$). Ainda, realizou-se testes dividindo os dados por faixa de desequilíbrio de tensão em 5 grupos distintos, quais sejam: $\leq 0,5\%$, $\leq 2\%$, $\leq 4\%$, $\leq 8\%$ e $\leq 10\%$ de desequilíbrio. Estes testes se aplicam a fim de avaliar a robustez e capacidade de generalização do algoritmo RF frente à problemas de qualidade da energia elétrica. São apresentados na Tabela 1 os parâmetros, as acurácias e as estatísticas para o algoritmo RF separados pelas diversas condições de entradas em análise neste estudo. Na tabela, as classes "0" e "1" representam experimentos considerando motor sob condições saudáveis e com defeito, respectivamente.

Observando os dados presentes na Tabela 1, nota-se que frente a variação do desequilíbrio de

Tabela 1: Parâmetros, acurácia e estatísticas do classificador Random Forest com 25 entradas, separados por faixa de desequilíbrio de tensão

25 entradas		Situação 1				Situação 2						
Desequilíbrio de tensão		≤ 0,5%				≤ 2%						
Correntes	$I_A + I_B$		$I_B + I_C$		$I_A + I_C$		$I_A + I_B$		$I_B + I_C$		$I_A + I_C$	
Classe	0	1	0	1	0	1	0	1	0	1	0	1
True Positive Rate	0,583	0,972	0,583	0,972	0,5	0,931	0,789	0,937	0,711	0,956	0,6	0,918
False Positive Rate	0,028	0,417	0,028	0,417	0,069	0,5	0,063	0,211	0,044	0,289	0,082	0,4
Acurácia	0,913	0,824	0,913	0,824	0,783	0,788	0,877	0,887	0,901	0,854	0,806	0,802
Recall	0,583	0,972	0,583	0,972	0,5	0,931	0,789	0,937	0,711	0,956	0,6	0,918
F-Measure	0,712	0,892	0,712	0,892	0,61	0,854	0,83	0,911	0,795	0,902	0,688	0,856
Área ROC	0,915	0,915	0,894	0,894	0,826	0,826	0,949	0,949	0,942	0,942	0,85	0,85
Acurácia Global (%)	84,25		84,25		78,70		88,35		86,74		80,32	
Erro Out-of-bag	0,1019		0,1574		0,1759		0,1245		0,1124		0,2088	
Tempo de Construção (s)	0,18		0,21		0,22		0,54		0,64		0,78	
Estatística Kappa	0,6107		0,6107		0,4733		0,7421		0,6991		0,5487	
Erro Absoluto Médio	0,2292		0,2672		0,3104		0,2197		0,2312		0,3279	
Erro Absoluto Relativo (%)	51,39		59,93		69,62		47,56		50,05		70,97	
Amostras	108		108		108		249		249		249	
25 entradas		Situação 3				Situação 4						
Desequilíbrio de tensão		≤ 4%				≤ 8%						
Correntes	$I_A + I_B$		$I_B + I_C$		$I_A + I_C$		$I_A + I_B$		$I_B + I_C$		$I_A + I_C$	
Classe	0	1	0	1	0	1	0	1	0	1	0	1
True Positive Rate	0,799	0,928	0,701	0,952	0,5	0,892	0,794	0,921	0,706	0,949	0,524	0,909
False Positive Rate	0,072	0,201	0,048	0,299	0,108	0,5	0,079	0,206	0,051	0,294	0,091	0,476
Acurácia	0,865	0,888	0,894	0,846	0,727	0,755	0,855	0,884	0,89	0,846	0,772	0,765
Recall	0,799	0,928	0,701	0,952	0,5	0,892	0,794	0,921	0,706	0,949	0,524	0,909
F-Measure	0,83	0,908	0,786	0,896	0,593	0,818	0,823	0,902	0,788	0,895	0,624	0,831
ROC Area	0,928	0,928	0,942	0,942	0,81	0,81	0,927	0,927	0,94	0,94	0,821	0,821
Acurácia Global (%)	88,04		86,00		74,80		87,37		85,90		76,65	
Erro Out-of-bag	0,112		0,1374		0,201		0,1248		0,1351		0,2129	
Tempo de Construção (s)	1,04		1,27		1,53		2,37		2,33		3,83	
Estatística Kappa	0,7382		0,6843		0,4192		0,7251		0,6842		0,4636	
Erro Absoluto Médio	0,2497		0,2417		0,3572		0,2658		0,2393		0,3538	
Erro Absoluto Relativo (%)	53,76		52,02		76,88		56,99		51,32		83,91	
Amostras	393		393		393		681		681		681	
25 entradas		Situação 5										
Desequilíbrio de tensão		≤ 10%										
Correntes	$I_A + I_B$		$I_B + I_C$		$I_A + I_C$							
Classe	0	1	0	1	0	1						
True Positive Rate	0,752	0,921	0,752	0,94	0,565	0,913						
False Positive Rate	0,079	0,248	0,06	0,248	0,087	0,435						
Acurácia	0,849	0,863	0,881	0,865	0,794	0,781						
Recall	0,752	0,921	0,752	0,94	0,565	0,913						
F-Measure	0,797	0,891	0,811	0,901	0,66	0,842						
Área ROC	0,907	0,907	0,943	0,943	0,827	0,827						
Acurácia Global (%)	85,81		87,03		78,42							
Erro Out-of-bag	0,1345		0,1297		0,2012							
Tempo de Construção (s)	2,88		3		4,43							
Estatística Kappa	0,6888		0,7134		0,5087							
Erro Absoluto Médio	0,2752		0,2366		0,3516							
Erro Absoluto Relativo (%)	58,95		50,68		75,31							
Amostras	825		825		825							

tensão, mesmo nas condições mais críticas, o classificador RF manifestou taxa de acerto percentual sempre superior a 84,25% para os sinais $I_A + I_B$ e $I_B + I_C$, denotando robustez e capacidade de generalização, ainda que na presença de problemas de qualidade da energia elétrica. No entanto, percebe-se que quando o método é alimentado com sinal $I_A + I_C$, apresenta acurácia inferior. Tal desigualdade é consequência do desequilíbrio de tensão de até +10% presente nos sinais de I_A e I_C , que quando somados atingem um desequilíbrio de aproximadamente 20%, caracterizando uma condição atípica e fora das recomendações conforme (Bollen, 2000). Ainda, acrescentando-se nível DC na soma das amplitudes dos sinais pode alterar o padrão da entrada, dificultando a correta classificação dos padrões. Porém, mesmo frente a tais condições críticas, o classificador mostrou-se robusto e generalista atingindo uma acurácia mínima de 77,80%.

5 Conclusões

O presente estudo expôs uma aplicação do algoritmo Random Forest, onde o método foi utilizado para identificação de falhas em rolamentos de motores de indução.

Dois motores de 1 cv operando com rolamentos saudáveis e defeituosos tiveram seus sinais de

corrente coletados em experimentos realizados em laboratório. Os sinais de duas das três correntes da máquina são somados, discretizados em 25 pontos por semiciclo e normalizados em função do valor máximo. Assim sendo, são criadas as matrizes que alimentam a entrada dos classificadores RF, e, suas saídas são usadas para auxiliar o diagnóstico de defeitos nos rolamentos dos MIT.

Por meio dos resultados experimentais apresentados, pôde-se analisar o desempenho do classificador Random Forest na identificação de defeitos em rolamentos e constar que a metodologia empregada pode ser usada como ferramenta complementar no processo de detecção de falhas.

Agradecimentos

Os autores agradecem o apoio da Fundação Araucária (Processo N° 06/56093-3) e do CNPq (Processo N°474290/2008-5, 473576/2011-2, 552269/2011-5).

Referências

Araújo, R., Rodrigues, R., de Paula, H. and Bacarini, L. (2011). Premature wear and recurring bearing failures in three-phase induction motors: A case study, *IEEE Latin America Transactions* 9(4): 528–533.

- Bellini, A., Filippetti, F., Tassoni, C. and Capolino, G. A. (2008). Advances in diagnostic techniques for induction machines, *IEEE Transactions on Industrial Electronics* **55**(12): 4109–4126.
- Bollen, M. (2000). *Understanding Power Quality Problems: Voltage Sags and Interruptions*, IEEE - TP 139-0, Wiley.
- Breiman, L. (2001). Random forests, *Machine Learning* **45**(1): 5–32.
- Broniera, P., Gongora, W., Goedtel, A. and Godoy, W. (2013). Diagnosis of stator winding inter-turn short circuit in three-phase induction motors by using artificial neural networks, *2013 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED)*, pp. 281–287.
- Dietterich, T. G. (2000). Ensemble methods in machine learning, *Proceedings of the First International Workshop on Multiple Classifier Systems*, MCS '00, Springer-Verlag, London, UK, UK, pp. 1–15.
- do Nascimento, C. F., de Oliveira Jr, A. A., Goedtel, A. and Serni, P. J. A. (2011). Harmonic identification using parallel neural networks in single-phase systems, *Applied Soft Computing* **11**(2): 2178–2185.
- ELETROBRÁS-PROCEL (2014). Balanço energético nacional.
- Ghate, V. N. and Dudul, S. V. (2010). Optimal mlp neural network classifier for fault detection of three phase induction motor, *Expert Systems with Applications* **37**(1): 3468–3481.
- Godoy, W., da Silva, I., Goedtel, A., Palácios, R. and Gongora, W. (2014). Neural approach for bearing fault classification in induction motors by using motor current and voltage, *2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 2087–2092.
- Gonzalez, A., Hernandez, C. and Arjona, M. (2014). A novel high-efficiency parallel-winding connection for a three-phase induction motor fed by a single-phase power supply, *IEEE Transactions on Energy Conversion* **29**(2): 269–277.
- Harlisca, C., Szabo, L., Frosini, L. and Albini, A. (2013). Bearing faults detection in induction machines based on statistical processing of the stray fluxes measurements, *2013 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED)*, pp. 371–376.
- Keskes, H., Braham, A. and Lachiri, Z. (2013). Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet {SVM}, *Electric Power Systems Research* **97**(0): 151 – 157.
- Leite, V., Borges da Silva, J., Cintra Veloso, G., Borges da Silva, L., Lambert-Torres, G., Bonaldi, E. and De Lacerda de Oliveira, L. (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.
- Palácios, R. H. C., Silva, I. N., Goedtel, A., Godoy, W. F. and Oleskovicz, M. (2014). A robust neural method to estimate torque in three-phase induction motor, *Journal of Control, Automation and Electrical Systems* **25**(4): 493–502.
- Patel, R. K. and Giri, V. (2016). Feature selection and classification of mechanical fault of an induction motor using random forest classifier, *Perspectives in Science* pp. –.
- Popescu, M., Dorrell, D., Alberti, L., Bianchi, N., Staton, D. and Hawkins, D. (2013). Thermal analysis of duplex three-phase induction motor under fault operating conditions, *IEEE Transactions on Industry Applications* **49**(4): 1523–1530.
- Seera, M., Lim, C. P. and Loo, C. K. (2014). Motor fault detection and diagnosis using a hybrid fmm-cart model with online learning, *Journal of Intelligent Manufacturing* pp. 1–13.
- Sirikulviriyaya, N. and Sinthupinyo, S. (2011). Integration of rules from a random forest, *2011 International Conference on Information and Electronics Engineering*, Vol. 6, pp. 194–198.
- Son, J.-D., Niu, G., Yang, B.-S., Hwang, D.-H. and Kang, D.-S. (2009). Development of smart sensors system for machine fault diagnosis, *Expert Systems with Applications* **36**(9): 11981 – 11991.
- Yang, X., Yan, R. and Gao, R. X. (2015). Induction motor fault diagnosis using multiple class feature selection, *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 256–260.
- Zarei, J., Tajeddini, M. A. and Karimi, H. R. (2014). Vibration analysis for bearing fault detection and classification using an intelligent filter, *Mechatronics* **24**(2): 151 – 157.