# Performance Comparison of classifiers in the detection of Short Circuit Incipient Fault in a Three-Phase Induction Motor

COELHO, D. N; BARRETO, G. A.
Departamento de Engenharia de Teleinformática – DETI
Universidade Federal do Ceará - UFC
Fortaleza, Ceará, Brasil
Email: davidcoelho89@gmail.com
Email: guialenbar@gmail.com

MEDEIROS, C. M. S.; SANTOS, J. D. A.
IFCE – Instituto Federal de Educação Ciência e Tecnologia do Ceará
Fortaleza, Ceará, Brasil
Email: claudiosa@ifce.edu.br
Email: jdalencars@gmail.com

***Abstract*—This paper aims at the detection of short-circuit incipient fault condition in a three-phase squirrel-cage induction motor fed by a sinusoidal PWM converter. In order to detect this fault, different operation conditions are applied to an induction motor, and each sample of the real data set is taken from the line currents of the PWM converter aforementioned. For feature extraction, the Motor Current Signature Analysis (MCSA) is used. The detection of this fault is treated as a classification problem, therefore different supervised algorithms of machine learning are used so as to solve it: Multi-layer Perceptron (MLP), Extreme Learning Machine (ELM), Support-Vector Machine (SVM), Least-Squares Support-Vector Machine (LSSVM), and the Minimal Learning Machine (MLM). These classifiers are tested and the results are compared with other works with the same data set. In near future, an embedded system can be equipped with these algorithms.**

***Keywords—SVM, LSSVM, MLP, ELM, MLM, Fault Detection, Three-phase Induction Motor***

## I. Introduction

The induction motors, due to their robustness, efficiency and simplicity, are the workhorses of the present industry [1-3]. According to [2], they can typically consume between 40 to 50% of all the generated capacity of an industrialized nation.

Even with its robustness, due to machine aging, installation environment conditions, inadequate applications or lack of preventive maintenance, the induction motor is subject of various faults [1] [4]. According to [5], the “motor failure rate is conservatively estimated as 3-5% per year”. The most common are bearing faults, stator or rotor isolation faults, open bars or crack of the rings, and eccentricity [4] [6].

Among all these faults, the insulation breakdown in the stator winding corresponds to nearly 40% of the total motor failures [4] [7]. This fault initiates as a high impedance fault [8]. Then, the fault current can cause a local heating, making the failure spread quickly in the winding [9]. If this fault is detected at the beginning of its occurrence, maintenance team can actuate and save production costs or the induction machine can be reused after the motor rewinding [10].

The cost of an unscheduled production downtime is too high [11]. As the stator winding inter-turn short-circuit takes just few minutes to evolve [2], a constant online monitoring of this fault is an important tool to reduce costs and save the machines.

Therefore, many studies in this fault detection have been made recently. As examples, in [12], Neural Networks and Wavelet transform are used in order to detect temporary short circuit in induction motor winding, and in [13], a wavelet-based probabilistic neural network is used for interturn fault detection. In other hand, as example of another kind of fault detection, [14] uses wavelet packet decomposition and artificial neural networks for bearing fault classification.

There are many ways to detect faults in induction motors, and among these, the Motor Current Signature Analysis (MCSA) is highly used. In inter-turn short-circuit detection, it is common to use some components of the current frequency spectrum to compose the input vectors. But these components can exist previously in the electric system or be affected by more than one kind of fault, and the current signals may be embedded in strong background [1]. These conditions can create severe difficulties for fault detection, and make the problem non linear.

Hence, together with the MCSA, different non-linear machine learning algorithms (that will be hereafter detailed) were tested and their accuracy was compared. All of these algorithms don´t need any motor mathematical model [7], but it is necessary a consistent and significant amount of data that can represent properly the specific problem. So, a real data set with different levels of fault was used.

The use of a frequency converter with a (machine learning based) detection system previously embedded in it can be a competitive advantage, because there is no need to

978-1-4799-4509-2/14/$31.00 ©2014 IEEE

considerable hardware alterations, and this machine is widely used in order to match all the applications that need speed variation. As examples of applications, one can mention fans, blowers, conveyors, crushers, compressors, cranes and pumps [5]. Moreover, the frequency converters can reduce maintenance and improve reliability to the motor [15].

The remainder of the paper is divided as follows: in section 2, the experimental data acquisition is explained; in section 3, the classifiers and the parameter choice are described; in section 4 the results are shown and discussed; finally, in section 5, the conclusions are made.

## II. Experimental Data Acquisiton

A standard WEG three-phase squirrel-cage induction motor has been used as base to this data set. Its main characteristics are 0.75 kW, 220/380 V, 3.02/1.75 A, 79.5% efficiency, 1720 rpm, Ip/In = 7.2, and 0.82 power factor. The data set used at this work was generated with the aforementioned motor operating in different conditions.

Each part of the data acquisition scheme is shown in figure 1, and is hereafter explained.

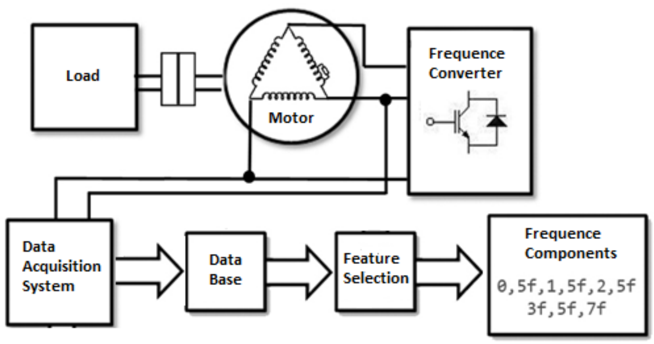


Figure 1. Data Acquisition Scheme

Firstly, a Foucault´s Break is used in order to apply three different levels of load: 0% (without load), 50% of nominal load and 100% of nominal load. The test bench made in order to apply these load levels can be seen in Figure 2.

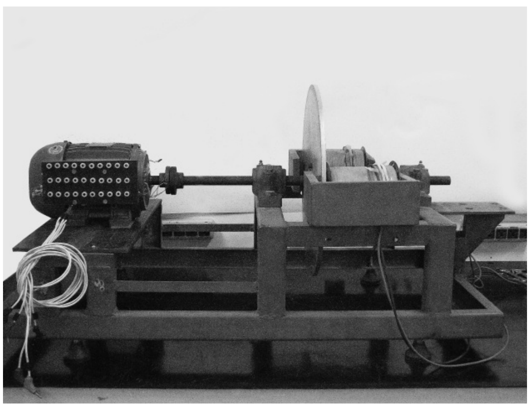

Figure 2. Test Bench

In [16] each part of the previously shown test bench is thoroughly described.

Secondly, in order to vary the speed of the motor, a frequency converter (also known as inverter drive) WEG CFW-09 was utilized with seven different frequencies: 30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, 55 Hz and 60 Hz.

This Frequency converter can be vector controlled (have a closed loop operation), but, at this work, only open loop operation is used.

Furthermore, three Hall's sensors were used to measure the line currents of each phase of this frequency converter.

Moreover, the motor was rewound so that some extra terminals are available exposing bobbins derivations of each phase. So, it is possible to emulate different inter-turn short-circuit levels. Hence, this is the only fault presented here.

At this work, three different levels of fault are used. In the lowest level (level 1), 5 turns were short-circuited, totaling 1.41% of the turns of one phase. In the medium level (level 2), 17 turns (4.8%) were short-circuited. Finally, in the highest level (level 3), 32 turns (9.26%) were short-circuited.

Finally, an auxiliary command system was built to execute two kinds of short-circuit schemes: the high impedance (that imitates the initial condition of short-circuit process) and the low impedance. With these two short-circuit schemes and three levels of faults, there are six different fault conditions of the motor.

All these operation conditions of the motor are shown in Table 1. The load level applied to the motor, the phase identification, the frequency of the voltage applied by the frequency converter and the "fault extent" are respectively presented.

In this last operation condition, "H" represents High impedance fault, "L" represents low impedance fault, and the numbers "1", "2" and "3" represent the level of the fault. All these conditions totalize 441 (3 x 3 x 7 x 7) time domain sample vectors.

TABLE 1. Operation conditions of the motor

| | Vector's Characteristics | | | | | | |
|---|---|---|---|---|---|---|---|
| Load Level | 0% | 50% | 100% | | | | |
| Inversor Phase | Ph 1 | Ph 2 | Ph 3 | | | | |
| Inversor Frequence | 30 Hz | 35 Hz | 40 Hz | 45 Hz | 50 Hz | 55 Hz | 60 Hz |
| Fault Extent | Normal | HI1 | HI2 | HI3 | LI1 | LI2 | LI3 |

It is important to mention that, as shown in figure 1, the motor was delta connected. With this configuration, two line currents of the frequency converter were directly connected to the faulty phase of the motor. As the authors want to make a system that can detect the fault just by one phase of the converter, just one of these previously mentioned phases was used in order to avoid redundancy of information.

So, 294 samples are used: 147 from phase 1 (directly connected to the fault current) and 147 from phase 3 (indirectly connected to the fault current). These samples are represented in figure 3.

As can be seen in this figure, the problem here explored can be treated as a multi-class problem, if one considers each fault extent as a class (normal, H1, H2, H3, L1, L2, L3). With this configuration, each class has 42 samples.

In other hand, if one want to classify the data between normal condition, high impedance fault or low impedance fault, the problem remains as a multi-class problem, but there are just three classes: one with 42 samples (normal), the other with 126 samples ( HI1, HI2 and HI3) and the last one with 126 samples (LI1, LI2 and LI3)..

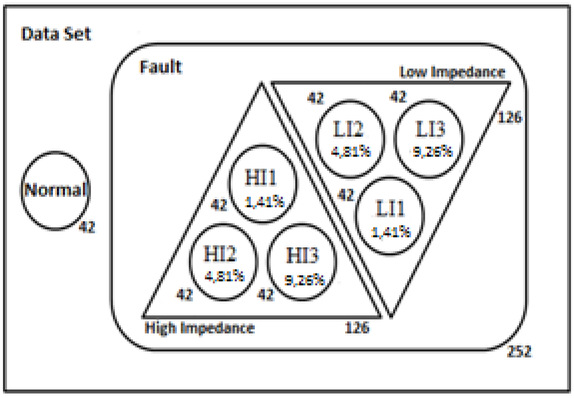


Figure 3. Data Set Distribution

But, it is important to mention that, at this work, it is not a considerable problem if the classifiers confound the level or the kind of the fault. Hence, the problem was treated as a binary problem: normal or faulty motor.

With this last configuration, there are 252 samples of faulty conditions (H1, H2, H3, L1, L2 and L3) and 42 samples of "normal" motor.

The data was acquired by a data acquisition system Agilent U2352 with 16 bits of resolution, passing through a 1 kHz analog filter and a signal amplifier. Each sample of the data set is the result of 10 seconds of acquisition with 10 kHz of sampling rate. So, each sample has 100,000 points.

After acquiring the time domain data, the Fast Fourier Transform (FFT) is used to obtain the spectrum in frequency domain. As there is an 1 kHz analog filter, the authors could choose until 500 Hz frequency as a characteristic of this problem.

According to [17], there are no novel components in stator motor current frequency spectrum due to the isolation fault, but there is just the increasing of some existing components.

In [18], the authors proposed a method to detect the short-circuit by monitoring some harmonic components in the axial leakage flux waveform. [16] done a statistic analysis of the variance of each harmonic generated by this method, noticed that some harmonics of the fundamental frequency inverter voltage varied the most with the different conditions of faults, and used these harmonics as attributes to this problem. In other hand, in [11], after taking into account other methods in addition to [18], and testing each component relevancy to an specific Artificial Neural Network (ANN), the authors chose the frequencies $0{,}5fs$, $1{,}5fs$, $2{,}5fs$, $3fs$, $5fs$ e $7fs$ as attributes to this problem, where "$fs$" is the source fundamental frequency.

At this paper, the previously cited harmonic frequencies were also used as attributes, because of all the methods that were taken into account in [11].

## III. CLASSIFIERS AND PARAMETER CHOICE

As the problem here is treated as a binary problem, each sample of the data set has just one label: +1 for normal operation condition or -1 for failure condition. With this configuration, 86% of the samples (252) are from faulty motors and 14% of the samples (42) are from normal motors.

As the principal issue of this work is to define if the motor is with or without fault (for, in near future, embed an algorithm), this problem can be treated as a classification or a pattern recognition task.

According to [19], "The central problem in pattern recognition is to define the shape and placement of the boundary so that the class assignment errors are minimized".

To solve this problem, state-of-the-art supervised classifiers with different learning machine paradigm were chosen. The MLP and the ELM are neural networks, the SVM and LSSVM classifiers have their theory grounded on the Statistical Learning Theory. Finally Learning in MLM consists in reconstructing the mapping existing between input and output distance matrices and then estimating the response from the geometrical configuration of the output points. Hereafter, each algorithm is briefly explained.

In order to choose the hyperparameters of each classifier, a grid search with 10-fold cross-validation was used.

### *A. MLP and ELM*

Single hidden layer Multi-Layer Perceptrons (MLPs) are known as universal approximators, because they can approximate any continuous function with a certain precision since it has neurons enough [20].

Moreover, vast research activities in neural classification have established that neural networks are a promising alternative to various conventional classification methods. Among these, the MLPs are the most used neural networks classifiers [21]. Hence, this algorithm was the first chosen to be tested.

There are many ways to train this algorithm, but "the MLP trained by BP algorithm (MLP/BP) is probably the most studied and classical neural model, especially in classification applications" [11]. At this work, what is called MLP is a MLP trained with the backpropagation algorithm.

The Extreme Learning Machine [22] has the same "structure" of the MLP trained by the backpropagation, but the training part is fast and one just has few parameters to adjust.

There are no standard formulas in order to determine the optimum parameters of these classifiers. In order to do this, attempt and error, or heuristic rules can be used.

In backpropagation training, one must adjust the learning rate, the momentum term, and the number of neurons in the hidden layer. On the other hand, in ELM training, one must adjust only the number of neurons in the hidden layer.

As previously mentioned, in order to choose the number of hidden neurons of MLP and ELM classifier, a grid search with 10-fold cross validation was used.

In [11], after a lot of tests with the same data set used here, the best results were achieved with 5 hidden neurons for the MLP and 20 hidden neurons for the ELM. So, the grid search for this parameter, in MLP was from 2 to 10 hidden neurons and the grid search for this parameter in ELM was from 15 to 25.

Finally, for both neural networks, the hyperbolic tangent function was used as the activation function, and, for MLP, the learning rate and the momentum term were chosen by attempt and error.

### B. *SVM and LSSVM*

The Support-Vector Machine (SVM) [23] [24] was introduced to solve binary pattern recognition problems but it is also capable of solving function estimation problems.

This algorithm tries to find the optimal hyperplane that separates the data from two classes, from some samples of these classes.

The learning step of the SVM classifiers aims at the minimization of both the Empirical Risk and the Structural Risk. The main part of this step is to calculate the samples that will contribute to create the optimal hyperplane (Support Vectors). This calculation involves solving a quadratic programming problem with inequality restrictions.

To treat the nonlinearity, a kernel function is used in order to map the data into a higher space, and constructing an optimal hyperplane in this space. At this work, a Gaussian (Radial Basis Function) Kernel, which is described in equation 1, is used.

$$K(x, x_i) = exp\{-\|x - x_i\|^2/\sigma^2\} \qquad (1)$$

The Least-Squares Support-Vector Machine [25] is also capable of solving classification and function estimation problems. But the principal difference is that this algorithm considers equality type constraints in order to solve a set of linear equations, instead of quadratic programming.

This characteristic makes this classifier's training faster, but, in the validation part, all the training samples are used in order to create the optimal hyperplane.

In order to select the optimal hyper-parameters of these 2 algorithms, a 10-fold cross-validation is used.

### C. *MLM*

The MLM [26] is a recent algorithm, firstly designed for multi-response regression, but it can also be used for classification problems adding one more step in the validation part.

"The basic idea behind the Minimal Learning Machine (MLM) is the existence of a mapping between the geometric configurations of points in the input and output spaces" [26].

Primarily, the data set must be divided in two parts: the training data set and the test data set. Then, "K" reference points have to be randomly chosen. These reference points correspond to some samples of the training data set. Each sample has its input (the attributes) and the output (the label).

The number of reference points is the only hyperparameter of this algorithm. And, like with the other classifiers, a 10-fold cross validation was used in other to choose this parameter, and the grid search was between 02 and 20 reference points.

The first step of the training part, is to calculate $D_x \in \mathcal{R}^{N\,x\,K}$, where each element $d_{nk}$ represents the distance from the "nth" (n = 1, … , N) input sample of the training data set to the "kth" (j = 1, … , K) input of the reference points (represented by $\boldsymbol{m_k}$). At this work the Euclidean distance is used.

The second step is to calculate $\Delta_y \in \mathcal{R}^{N\,x\,K}$ where each element of this matrix represents the distance from each output sample of the training data set to each output of the reference points (represented by $\boldsymbol{t_k}$).

The last step is to calculate a linear regression model that relates matrix $D_x$ and matrix $\Delta_y$. Equation 2 represents the linear model, matrix E corresponds to the residuals and, in equation 3, the Moore-Penrose method is used in order to calculate matrix $B \in \mathcal{R}^{K\,x\,K}$.

$$\Delta_y = D_x . B + E \qquad (2)$$

$$\hat{B} = (D_x' D_x)^{-1} D_x' \Delta_y \qquad (3)$$

The validation part consists of, for a testing sample **x,** calculating a vector $\in \mathcal{R}^{1\,x\,K}$ **,** which elements represent the distance between the input reference points and the sample **x**. Then, the distance between the outputs of the sample x and the outputs of the reference points are estimated by the equation 4:

$$\hat{\delta}(\boldsymbol{y}, \boldsymbol{t_k}) = d.\hat{B} \qquad (4)$$

where each element of $\hat{\delta}(\boldsymbol{y}, \boldsymbol{t_k}) \in \mathcal{R}^{1\,x\,K}$ represents the distance between the output of the reference points and the output from the sample x.

Finally, the output **y** is estimated by solving the optimization problem described in equation 5:

$$\hat{y} = \text{argmin}_y \sum_{k=1}^{K} \left( (\boldsymbol{y} - \boldsymbol{t_k})' (\boldsymbol{y} - \boldsymbol{t_k}) - \hat{\delta}^2 (\boldsymbol{y}, \boldsymbol{t_k}) \right)^2 \quad (5)$$

At this work, the Levenberg-Marquardt Algorithm is used to solve the aforementioned problem.

## IV. Results and Discussion

With each classifier, fifty (50) independent turns of training and validation are made, and, for each turn, four steps are performed: the hold out (division of the data set between training data set and validation data set), the hyperparameters choice (grid search and cross validation), the training part and the validation part.

At the end of each turn, the accuracy rate of each classifier is calculated and saved.

In order to separate the entire data set between training and validation data sets (hold out), three methods were used. The results will be hereafter discussed according to each method.

At the first method, the data are randomly divided as follows: 80% for training and 20% for validation. As can be seen in the boxplot of Figure 4, all classifiers achieve a high maximum accuracy rate (more than 90%). But, as there are 252 data from faulty motor and just 42 for normal motor, the algorithms tend to classify most of the data as faulty motors. So, many data from normal motor are classified as failure, or, in other words, a false positive happens.

With the MLP for example, only 21.5% of the data from normal condition was correctly classified. This result can be verified with the confusion matrix of table 2.

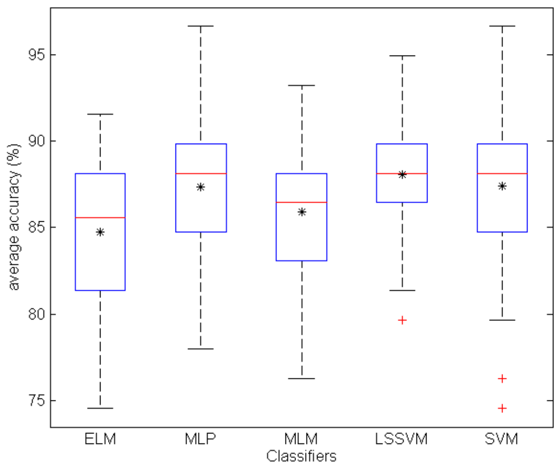


Figure 4. Accuracy of the first method

The same problem occurred in [16], where the accuracy was 87.5%, but the accuracy specifically in the normal condition data set is just 52%.

Table 2 – Confusion Matrix of MLP in the first method

| MLP | Confusion Matrix | |
|---|---|---|
| | ***Faulty*** | ***Normal*** |
| Faulty | 2487 | 45 |
| Normal | 328 | 90 |

These are not desirable conditions for the problem of incipient short circuit detection, because false positives can bring unexpected downtime in the production and, as previously mentioned, reduce industry profits.

In order to solve this problem, at the second hold out method, the authors tried to improve the training part of all algorithms.

As there was more samples of faulty conditions than normal conditions in the training part, in the second hold out method, the authors balance the faulty samples and the normal samples: 80% of randomly chosen normal condition data (33 samples) and 33 samples of failure condition (totaling 66 samples, or 22% of the entire data set) are used for the training part. The algorithms are validated with the remaining samples. The results of this method are shown at the box plot of figure 5.

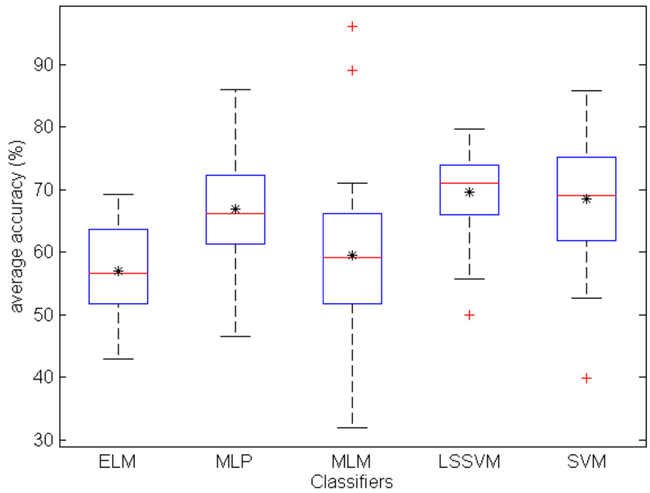


Figure 5. Accuracy of the second method

The false positives were highly reduced. For example, with the MLP, in the first method, there was a false positive mean rate of 78.5%, and, in the second method, this mean rate decayed for 28.4%, as can be seen in table 3.

Table 3 – Confusion Matrix of MLP in the second method

| MLP | Confusion Matrix | |
|---|---|---|
| | ***Faulty*** | ***Normal*** |
| Faulty | 7314 | 3636 |
| Normal | 322 | 128 |

In [11], the authors treated the problem differently by just taking into account the tests in which all the normal condition motor data were classified correctly.

But the second method brought another problem: just few data to train the algorithms. The false positives were reduced, but the classification rate was also reduced. Moreover, ELM and MLM had the worst results.

So, in order to have a high accuracy rate and a low false positive mean rate, the authors tried a third method.

At this method, the number of normal data and faulty data was equalized. This was made by generating 210 artificial normal data. Each of these artificial data consists of a real normal sample corrupted by a normally distributed noise.

The new data set has 252 samples of data from normal conditions (42 real data plus 210 artificial data) and 252 samples of data from faulty condition, totaling 504 data. With this data set, as the first method, the data are randomly divided as follows: 80% for training and 20% for validation.

As can be seen in the boxplot of figure 6, the SVM and LSSVM classifiers achieved almost 97.5% and 99.5% of mean accuracy rate respectively. With the other classifiers, the accuracy rate and the false positive rate were balanced.

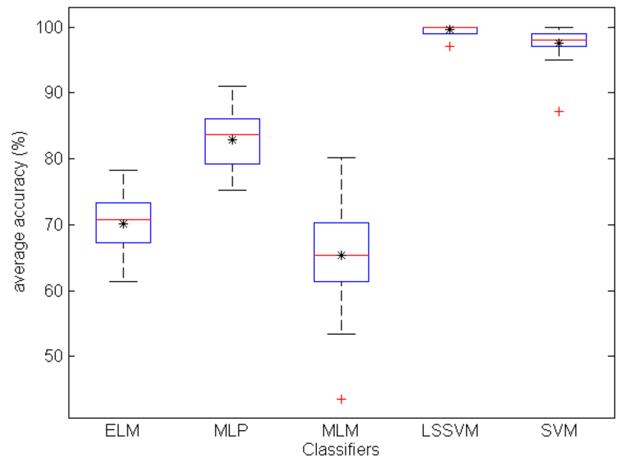


Figure 6. Accuracy of the third method

With MLP, for example, there was a mean false positive rate of 22.9% (5.9% less than in the second method), and a mean accuracy rate of 82.8% (just 4.5% less than in the second method). This result can be verified by the confusion matrix of table 4.

Table 4 – Confusion Matrix of MLP in the third method

| **MLP** | **Confusion Matrix** | |
|---|---|---|
| | ***Faulty*** | ***Normal*** |
| Faulty | 2237 | 290 |
| Normal | 578 | 1945 |

Just taking into account the SVM and the LSSVM classifiers (which got the best results at this method), the initial difficulty, with the other methods, to classify the data with a low false positive rate was due to the poor representation of the normal condition class.

Moreover, as the authors intend to embed, in near future, an algorithm in a frequency converter, the SVM will probably be the best choice, because it uses less support vectors from the training data set than the LSSVM. This characteristic makes the validation part of the SVM faster than the validation part of LSSVM and this can enhance the speed for the classification of the fault.

In the last "hold out" method, for example, the average number of support vectors with SVM was 60% of the entire data set. In other hand, the LSSVM uses all the samples as support vectors, which corresponds to 80% of the entire data set.

## V. CONCLUSION

At this work, different "state-of-the-art" classifiers were used and compared in the detection of short circuit incipient fault, and a mean accuracy rate of 97.5% and 99.5% were achieved with the SVM and the LSSVM classifier respectively.

The results are promising. But, as the main objective of this work is not only classify offline the faults, but also build an embedded system for motor fault detection, the accuracy rate is not the only detail that should be studied.

Other characteristics such as false positives (which can lead unexpected downtime in the production), and processing time (the time it takes for each classifier to detect the fault) will be taken into account.

Moreover, other techniques of feature extraction and selection, and other classifiers will be, in near future tested. Therewith, the best match between a feature extraction technique and a classifier will be embedded in a frequency inverter.

Besides this, the authors intend to, in near future, apply a closed loop operation in the motor and see if, with this condition, the short circuit incipient fault can still be detected.

On the other hand, the problem discussed here, can also be treated as a multiclass problem, if one wants to know the level of the fault. So, a supervisory system can be made if a classifier can distinguish the levels of the faults.

## *Acknowledgment*


The Authors thank CAPES/CNPQ for the financial support, the Federal Institute of Ceará (IFCE) that provided the means for this work.


## *References*


[1] Ghate, V. N.; Dudul, S. V. Optimal MLP neural network classifier for fault detection of three phase induction motor. Expert Systems With Applications 37. 2010

[2] Thomson, W.T.; Fenger, M.; , "Current signature analysis to detect induction motor faults," Industry Applications Magazine, IEEE , vol.7, no.4, pp.26-34, Jul/Aug 2001

[3] Bacha, K. et al. Induction machine fault detection using stray flux EMF measurement and neural network-based decision. Electric Power Systems Research 78. 2008

[4] NandiI, S.; Toliyat, H. A.; LI, X. Condition monitoring and fault diagnosis of electrical machines - a review. IEEE Transactions on Energy Conversion, v. 20, n. 4, p. 719_729, 2005.

[5] Vico, J.; Hunt, R. Projection Principle for Electrical Motors in the Cement Industry. Cement Industry Technical Conference, 2010 IEEE-IAS/PCA 52nd

[6] Bonnet, A. H. Root Cause Failure Analysis for AC Induction Motors in the Petroleum and Chemical Industry. Petroleum and Chemical Industry Conference (PCIC). 2010.

[7] Martins, J. F. ; Pires, V. F. ; Pires, A. J. Unsupervised Neural-Network-Based Algorithm for an On-Line Diagnosis of Three-Phase Induction Motor Stator Fault. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, (2007).

[8] Nataranjan, R. Failure identification of induction motors by sensing unbalanced stator currents. IEEE Transactions on Energy Conversion, v. 4, n. 4, p. 585_590, 1989.

[9] Tallam, R. M. et al. A survey of methods for detection of stator related faults in induction machines. In: Proceedings of the IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED'03). [S.l.: s.n.], 2003. p. 35_46.

[10] Thomson, W. T. On-line mcsa to diagnose shorted turns in low voltage stator windings of 3-phase induction motors prior to failure. In: Proceedings of the IEEE International Conference on Electric Machines and Drives (IEMDC'01). [S.l.: s.n.], 2001. p. 891_898.

[11] Oliveira, A. G.; Medeiros, C. M. S.;Pontes, R. S. T. Stator Winding Interturns Short Circuit Fault Detection in a Three Phase Induction Motor Driven by Frequency Converter Using Neural Networks. Energy Efficiency in Motor Driven Systems 2013 Conference, EEMODS 2013. Rio de Janeiro – RJ. 2013

[12] Asfani, D. A. et al. Temporary short circuit detection in induction motor winding using combination of wavelet transform and neural network. Expert Systems with Applications 39. 2012

[13] Seshadrinath J.; Panigrahi, B. K. Incipient Interturn Fault Diagnosis in Induction Machines Using an Analytic Wavelet-Based Optimized Bayesian Inference. IEEE Transactions on Neural Networks and Learning Systems, VOL. 25, NO. 5, MAY 2014

[14] Yadav, M.; Wadhwani, S. Automatic Fault Classification of Rolling Element Bearing using Wavelet Packet Decomposition and Artificial Neural Network. International Journal of Engineering and Technology Vol.3 (4), 2011, 270-276

[15] Sawa, T.; Kume, T. Motor Drive Technology – History and Visions for the Future. In: ANNUAL IEEE POWER ELECTRONICS SPECIALISTS CONFERENCE, 35, 2004, Auchen, Alemanha.

[16] Coelho, D.; Medeiros, C. Short Circuit Incipient Fault Detection and Supervision in a Three-Phase Induction Motor with a SOM-Based Algorithm. *Book of Advances in Self-Organizing Maps*. vol. 198, p. 315-323. Jan. 2013. ISBN 978-3-642-35229-4.

[17] Joksimovic, G. M.; Penman, J. The detection of inter-turn short circuits in the stator windings of operating motors. IEEE Transactions on Industrial Electronics, v. 47, n. 5, p. 1078_1084, 2000.

[18] Penman, J.; Sedding, H.G.; Lloyd, B.A.; Fink, W. T., "Detection and location of interturn short circuits in the stator windings of operating motors," Energy Conversion, IEEE Transactions on , vol.9, no.4, pp.652,658, Dec 1994

[19] Principe, J. C.; Euliano, N. R.; Lefebvre W. C. Neural and Adaptive Systems: Fundamentals through Simulations. Ed. Willey. 2000. 672 pgs.

[20] Hornik, K.; Stinchcombe, M.; White, H. "Multilayer feedforward networks are universal approximators". Neural Networks, v. 2, p 359-366, 1989

[21] Zang, G. P. Neural Network for Classification: A Survey. IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews).Vol. 30, No. 4. 2000. pp. 451-462.

[22] Guang-Bin Huang; Qin-Yu Zhu; Chee-Kheong Siew, "Extreme learningmachine: a new learning scheme of feedforward neural networks," Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on , vol.2, no., pp.985,990 vol.2, 25-29 July 2004

[23] Vapnik, V. N. The nature of Statistical learning theory. Springer-Verlag, New York, 1995

[24] Vapnik, V. N. Statistical Learning Theory. John Wiley and Sons, 1998.

[25] Suykens, J. A. K.; Vandewalle, J. Least squares support vector machine classifiers. Neural Processing Letters, v. 9, n. 3, p. 293–300, 1999

[26] Souza Junior, A. H. et al. Minimal Learning Machine: A New Distance-Based Method for Supervised Learning. 12th International Work-Conference on Artificial Neural Networks, IWANN 2013, Puerto de la Cruz, Tenerife, Spain, June 12-14, 2013, Proceeding.