Deep learning based optimum fault diagnosis of electrical and mechanical faults in induction motor

Vikas Singh¹, Purushottam Gangsar¹, Ashok Atulkar¹ and Rajkumar Porwal¹

¹Mechanical Engineering Department, Shri G S Institute of Technology and Science (SGSITS), Indore, India

E-mail: vikassingh14101992@gmail.com

Abstract. Among all the motors Induction motor (IM) plays a vital role in industry and the demand for their reliability and safe operation is increasing day by day. They are reliable but they do wear out if not maintained timely which in turn will lead to excessive loss of revenue and also man power and this motivates us to develop an intelligent methodology for diagnostic of incipient faults in Induction Motor. This paper focuses on the development of optimum Deep Learning based diagnostic technique to detect the mechanical and electrical faults in Induction Motor. Here, mechanical and electrical faults of different severity level are being tested on machine fault simulator using acquired current and vibrational signals. In this work, optimum model of Deep Neural Network (DNN) based on critical vibration and current features is developed and finally used to timely and effectively detect that which kind of faults the given IM is dealing with. The results are added and discussed in the result and discussions.

Keywords: Induction motor, electrical faults, mechanical faults, deep neural network, overall accuracy, loss percentage, overfitting, relu and softmax.

1. Introduction

World is approaching towards fourth industrial revolution also coined as industry 4.0. The main objective of this era is to incorporate the sensor information and use it for condition monitoring and fault diagnosis of machines. Manufacturing machines often encountered with the failures which may result in some huge accidents, losses and even causalities. With growing demand for high quality of products has increased our reliability over these machines and also the pressure on these machines have also been increased significantly. So, effective diagnosis of manufacturing machine i.e IM has become our most important and prime task which drags our interest towards its progress in the field of research related with its diagnosis [1]. Because of this experimentation on fault diagnosis of induction motor picks up speed in its data collection with the help of advanced sensors and the decisions are concluded on the basis of these sensor data which has been seen success in determining the fault in Induction motor [2,3,4]. The skilled manual labour is required only to analyze and handle the data which is coming out during the working of induction motor which helps in its early identification of failures. Recently, various methods have been acknowledged and introduced in the fault diagnosis of induction motor continuously [1, 5].

For each diagnosis method engineers need some basic knowledge and some experience and because of lack of such people the traditional methods become more tedious and tricky. Even if somehow we arrange such experts the knowledge and training given by these people are not hand in hand and for small scale industries it is sometime becomes very expensive task to avail such kind of experts [6]. Overpast 30 years several Artificial Intelligence techniques have been evolved and put in the monitoring of Induction motor like Support Vector Machines (SVM), Artificial Neural Network (ANN), Fuzzy Logic(FL), Neuro-Fuzzy etc [6, 7, 8, 9]. All these A.I techniques differ from one to one by their accuracy, loss function and prediction performance. These machine learning algorithms require a lots of labelled data and exhaustive feature analysis for effectively training of the model [10].
Since DNNs acquire valiant representation skills due to its heaped hidden layers, they can extract the useful features automatically. They do not require exhaustive feature analysis and therefore they have been getting popularity as a classifier in many machine fault diagnosis [11, 12, 13]. In this work, DNN was considered to develop the optimum fault diagnosis model for IM. Ten mechanical and electrical faults were simulated on machine fault simulator and used to generate motor current and vibration signals. These signals were further used to extract three critical fault features i.e., standard deviation, kurtosis and variance then these features are fed to DNN for IM fault diagnosis. A significant amount of accuracy and considerable amount of loss function has been observed which is best in terms of Industry benchmark.

2. Introduction to DNN
Deep learning is a subdivision of ML, one can say it is refined version of Machine Learning which is based on ANN based on Reorientation Learning. DNN is one of the architecture of the Deep learning which is used for the classification of simple data frame. Basically DNN is nothing but an ANN in which multiple hidden layers are arranged between input and output layer. Composition of Neural Network includes mainly neurons, synapses, weights, biases, and Activation functions. These components are clubbed altogether and can be arranged in a format like human nervous system and function equivalent to human brain. It is generally feed forward neural network in which information run between input and output layer without bending back. At earliest a map of virtual neurons is created by the DNN and assigns weight (numerical values) to them then weights and inputs are multiplied adding with biases in it which is further shoted to output layer with the help of activation function which are sitting on each hidden neurons. Lastly the output layer shows the result between 0 to 1. Sometimes network do not accurately predict the pattern so algorithm automatically adjust its weight that the way giving importance to some other input variable which is more dominant, until it finds the right mathematical computation to process the data completely[14].

The deep in Deep Learning refers to the large no. of layers of neurons that help to learn various representations of data. This can be understood by Figure 1. Figure 1 shows the working of DNN i.e information (features) given to the input layer has to travel through three hidden layers with some extra information loaded in the form of weights and biases and shoted to output layer for Prediction. There is still a basic question that comes in most of the A.I engineers that why most of the people demands deep learning instead of Machine Learning, although they hold the better understanding of Machine Learning. Now as explained earlier that Deep learning is a subdivision of ML and is used to extract useful pattern from data computationally. In simple word one can say that Machine Learning models needs human intervention for the final outcome but Deep Learning models make predictions independent of human Intervention [15].

Here below are some hyperparameters that has to be optimised for better prediction of the model. Hyperparameters are variables inside the structure of the DNN that can be altered for the betterment of our model. The optimization of these hyperparameters is itself a very lengthy and time taking process because it is based on hit and trial methodology, but we can proceed our steps with some rules that are given by the researchers [14, 15].

![Figure 1. Functioning of Deep Neural Network](image-url)
2.1. Selection of Layers and Neurons

There is no such rule in selecting the optimum no. of hidden layers and neurons but we can restrict our trial and error method of selecting neurons and layers by getting some knowledge from different theories given by researchers in the past like:

1. Amount of hidden neurons have to be around the magnitude of the input and output neurons.
2. Amount of hidden neurons have to be 2/3 of the magnitude of the input neurons in addition to the size of the output neurons.
3. Amount of hidden neurons have not be more than the two times of the input neurons.

2.2. Selection of epoch

As per the no. of epochs are considered one can say that it is not the significant hyperparameter that is to be optimized. More important is the training and validation error, as long as these error are decreasing one should continue the training. We should keep track on our validation losses that is if the loss starts increasing after certain no. of epochs then that will lead to over fitting of the model which is further addressed by different regularization technique.

2.3. Selection of Batch Size

Batch size is another type of hyperparameter that tells us the no.of specimen to work through before renovating the internal model parameters. Generally there is no specific rule for the selection of the proper batch size. Larger the batch size result in faster progress in training but converges very slow and vice versa. So for our convenience we take batch size in order of 32, 64, 128 …so on.

2.4. Selection of Activation Function:

Activation function is a node that has to be added in between the two hidden layers and the output layer, sometimes also termed as transfer function. Actually there are two types activation function like-

1. Linear
2. Non-linear

Now we use only non linear activation function for our classification problem because our output label is between 0 to 1 whereas linear function may bring any number between –infinity to infinity. Five most commonly used Activation Functions are:-

1. Sigmoid- This kind of Activation Function we use generally for binary classification on output layer where result is provided as 0 or 1. We cannot use sigmoidal function in between the hidden layers as it kills and saturate the gradients.

2. Tanh- It is widely used in between the hidden layers as its value lie down around -1 to 1 therefore the mean avlue comes out to be 0 for hidden layers or close to it, hence assisting the data to be centralized and also makes learning easy for the next layer.

3. Relu- This kind of transfer function is relatively easier as compare to tanh and sigmoid as it includes very simple mathematical computation. For an instance only some neurons are mobilized in connecting the network ambivalent and productive.

4. Leaky relu- It is just used as the apparent to Relu as it overcome the dilemma of vanishing of the gradient which we saw in Relu.

5. Softmax-It is called when we have multiclass problem in front of us as in we have to categorize the output layer of the dataframe. It can only be used in output layer not in between the hidden layers.

Choosing the right kind of activation function for our model is a very complex and lengthy task as it involves no. of iteration because Good or bad- there is no thumb rule for the right selection of the activation function but we can converges this issue by some of our prior knowledge about these activation function and also by knowing the properties of our problem statement.

3. Experimental Study

The experiment is carried out on n test rig (as shown in figure 2) comprised of a Machine Fault Simulator (MFS) which is available in mechanical engineering department, IIT Indore, MP. This MFS as shown in Figure 2 comprises of 3-φ Squirrel cage Induction motor, shaft, coupling, and pulley with
V-type belt, Speed controller, Magnetic brake and a photovoltaic sensor. Vibration signals are measured with the help of tri axial accelerometer in all the three orthogonal directions and current signals in all the three phases are to be measured by three AC current probes. Both were reinforced on Induction Motor. The photovoltaic sensor which measures the rotational speed of the shaft requires DC power source and was mounted near the coupling. In order to vary load on IM Magnetic brakes was used.

![Induction motor experimental test-rig](image)

**Figure 2.** Induction motor experimental test-rig

In this work following ten different faults of the IM were considered including healthy motor. Faults were arranged as five electrical related faults i.e. Stator Winding fault with severity level 1 and 2 (SWF1 and SWF2), Broken Rotor Bar fault (BRB), Phase Unbalance with severity level 1 and 2 (PU1 and PU2) and four Mechanical Related i.e. Unbalanced Rotor (UR), Bowed Rotor fault (BR), Bearing Fault (BF), Misaligned Rotor fault (MR). The motor conditions are shown in figure 3. These faults were simulated on MFS one after the other to set up data. A Data Acquisition System was used to obtain vibration and current data. At 1000 Hz sampling rate the data is obtained and 10,000 no. of sample points. Datasets were collected for 40 Hz frequency and high load for each fault conditions. Overall 25 raw data sets (25*10000 sample points) were gathered for each fault condition. These data now would be used for DNN based fault diagnosis.

![Induction motor with various seeded fault](image)

**Figure 3.** Induction motor with various seeded fault
4. Result and Discussion

The signals acquired from the experiments are further used to extract three critical fault features i.e., standard deviation, kurtosis and variance. As per previous study, these features are found to be critical for IM fault diagnosis [16, 17]. These features of vibration and current signals extracted firstly and then fed to DNN for present work. The performance of DNN depends on its Hyperparameters. For developing an effective DNN model, two simulations are considered in this study one with number of batch size and second with number of Epochs. For this study no. of layers is arbitrarily selected as 4, no. of neurons is selected as 64 in first, 45 in second, 32 in third and 16 in fourth hidden layer. From the research it is found that the Relu is the finest activation function for hidden Layers and Softmax is best for output layer in case multiple class output [18]. Therefore in this study, the Relu and Softmax functions are used in hidden and output layer, respectively. The effect of no. of batch size and epochs are checked on individual and overall accuracy of detection.

4.1 Selection of Optimum Number of Batch size

This simulation focuses on the selecting optimum number of batch size. For this five simulations are considered starting from batch size 8 to 100. Table 1 shows the effect of no. of batch size on the classification accuracy and percentage loss. The experiment starts with 100 epochs for all the cases.

Results from table 1 show that the minimum and maximum overall accuracy are found to be 96 % (with 16 batch size) and 100% (with 8 batch size), respectively. Based on the result in Table 1 one can easily observed that, on increasing the no. of batch size the testing accuracy dropped significantly in few steps but hits some peaks in between. So accuracy cannot be alone factor that decides the model to be optimized but loss percentage also plays a vital role in this picture. From table 1 the minimum percentage losses are found to be 0.69 and 0.12 (with 8 batch size) for training and testing, respectively. In this table it can be easily seen that with increase in batch size the loss percentage keeps on increasing. So both loss percentage and testing accuracy should be our main concern in optimization of DNN model.

| Batch Size | Individual Accuracy, in % | Overall Accuracy (%) | Loss, % |
|------------|---------------------------|----------------------|---------|
|            | BF | BRB | BR | PU2 | SWF | ND | PU1 | RM | SWF | UR |
|            | Tr. | Tst. | Tr. | Tst. | Tr. | Tst. | Tr. | Tst. | Tr. | Tst. |
| 8          | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 0.69 | 0.12 |
| 16         | 100 | 100 | 100 | 100 | 100 | 96.1 | 100 | 100 | 100 | 96.2 | 1.3 | 8.7 |
| 32         | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 99.5 | 100 | 2.6 | 2.4 |
| 64         | 100 | 96.2 | 100 | 100 | 100 | 100 | 100 | 100 | 96.2 | 97.5 | 96.1 | 9.6 | 9.3 |
| 100        | 100 | 100 | 100 | 100 | 100 | 96.2 | 100 | 100 | 96.2 | 100 | 95.1 | 96.1 | 19.8 | 20.3 |

Figure 4 clearly signifies the acceptance of the batch size 8 with no. of epoch 100 as throughout the training and validation of the data both the lines follows each other approximately. For batch size 16 figure 5 shows same trend as in figure 4 for batch size 8. However the overall testing accuracy is less in case of batch size of 16 as compared to case of batch size of 8. Therefore it can conclude that the batch size of 8 is optimum value for the present diagnosis.
4.2 Selection of Optimum Number of Epoch

This simulation focuses on the selecting optimum number of epoch. For this five simulations are considered starting from 20 epoch to 200 epoch, table 2 shows the effect of no. of epochs on the classification accuracy and percentage loss. Here optimum batch size is taken as 8 from the results shown in table 1.

Based on the result in table 2, it can be observed that with increase in number of epoch which is ultimately the no. of iterations on training algorithm, the testing accuracy is increasing with little ups and down. Testing accuracy depends upon the no. of samples it is being tested so every time we cannot conclude alone on the basis of our testing accuracy. The loss percentage as one of the factor which is decreasing upon increasing the no. of iterations. So it can also be one of other criteria for selection of optimized model.

| Epoch | Individual Accuracy, in % | Overall Accuracy (%) | Loss, % |
|-------|---------------------------|----------------------|--------|
|       | Tr. | Tst. | Tr. | Tst. | Tr. | Tst. | Tr. | Tst. |
| 20    | 100 | 96.1 | 100 | 92.3 | 100 | 96.1 | 100 | 97.5 | 100 | 88.4 | 19.1 | 33.5 |
| 40    | 100 | 96.1 | 100 | 92.3 | 100 | 92.3 | 100 | 99.1 | 96.1 | 96.1 | 10.5 | 27.1 |
| 60    | 100 | 96.1 | 100 | 96.1 | 100 | 96.1 | 100 | 98.5 | 96.1 | 96.1 | 7.3 | 16.9 |
| 80    | 100 | 96.1 | 100 | 96.1 | 100 | 96.1 | 100 | 96.1 | 96.1 | 96.1 | 3.1 | 6.4 |
| 20    | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 0.1 | 0.91 |

Results from table 2 show that the minimum and maximum overall accuracy are found to be 88.4% (with 8 batch size and 20 epoch) and 100% (with 8 batch size and 200 Epoch), respectively. The minimum percentage losses are found to be 0.1 and 0.91 (with 8 batch size and 200 epochs) for training
and testing, respectively. But batch size 8 and epoch 200 cannot be considered as optimized hyperparameters for the present model because it is the case of over-fitting (both lines are going away from each other). The problem of overfitting in this case can be noticed from Figure 6. However the case with batch size 8 and epoch size 80 is performing well in terms of accuracy and loss percentage and also there is no problem of overfitting as shown in figure 7.

Choosing the optimized batch size and epoch is a kind of hit and trial method, the more we train the model the closer we come to our optimized model. So we cannot guess the exact size of batch and no. of epoch. We can only come to a range of values for both the hyper parameters. So for this problem, optimum batch size is found to be in between 8 to 16 and the optimum no. of epoch is found to be in between 60 to 100 for our optimized model. The final overall testing accuracy with these optimized parameters (batch size 8 and epoch 100) are found to be 100 %. It means the optimized DNN model performs perfectly for fault diagnosis of IM based on three critical features (standard deviation, kurtosis and variance) of vibration and current signals. All the mechanical as well as electrical faults with different severity levels are diagnosed perfectly using proposed methodology.

![Figure 6. Model Accuracy and Model Loss with Epochs for batch size 8 and Epoch 200](image)

![Figure 7. Model Accuracy and Model Loss with Epochs for batch size-8 and Epoch-80](image)

5. Conclusion
In this study, DNN based fault diagnosis are developed for IM to predict up to ten electrical faults with their severity levels and mechanical faults respectively. The vibration and current signals from all the ten working state of the IM acquired from MFS were used. The intent of this work is to develop optimum DNN based diagnosis with three critical features (standard deviation, kurtosis and variance) to give us the best classification accuracy along with least loss percentage. There is always a limitation of batch size and no. of epoch i.e for small dataset, increasing both hyper parameters cannot be a better proposal, chance of over-fitting will always be there. Therefore, in this work, the optimal no. of batch size and no. of epochs are selected and then optimal IM fault diagnosis is performed using these optimized parameters. The maximum prediction accuracy obtained was 100 % which shows the effectiveness of
the developed methodology. The study can be applied with other optimization algorithms and operating conditions of the motor to check the robustness of the developed diagnostics.

Acknowledgement
Authors would like to thank NPIU (TEQIP III) for providing financial support through Collaborative Research Scheme (CRS). We are also thankful to IIT Indore for providing experimental facilities for this collaborative work

References
[1] Gangsar P and Tiwari R 2020 Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review Mechanical Systems and Signal Processing 144 106908.
[2] M Riera Guasp, J A Antonino Daviu and G A Capolino 2014 Advances in electrical machine, power electronic, and drive condition monitoring and fault detection: state of the art IEEE Transactions on Industrial Electronics 62(3) 1746–1759.
[3] He Q and Du D M 2007 Fault diagnosis of induction motor using neural networks. International Conference on Machine Learning and Cybernetics 2 1090- 1095 IEEE.
[4] Li X and Zhang W 2020 Deep learning-based partial domain adaptation method on intelligent machinery fault diagnostics IEEE Transactions on Industrial Electronics.
[5] Chouhan A, Gangsar P, Porwal R and Mechefske C K 2020 Artificial neural network based fault diagnostics for three phase induction motors under similar operating conditions. Vibroengineering Procedia 30 55-60.
[6] Gangsar P and Tiwari R 2018 Effect of noise on support vector machine based fault diagnosis of IM using vibration and current signatures In MATEC Web of Conferences 211 03009.
[7] Lei Y, Yang B, Jiang X, Jia F, Li N and Nandi A K 2020 Applications of machine learning to machine fault diagnosis: A review and roadmap Mechanical Systems and Signal Processing 138 106587.
[8] ALShorman O, Irfan M, Saad N, Zhen D, Haider N, Glowacz A, and Al Shorman A 2020 A Review of Artificial Intelligence Methods for Condition Monitoring and Fault Diagnosis of Rolling Element Bearings for Induction Motor Shock and Vibration.
[9] Gangsar P and Tiwari R 2019 Performance Analysis of Support Vector Machine and Wavelet Packet Transform Based Fault Diagnostics of Induction Motor at Various Operating Conditions Mechanisms and Machine Science 61.
[10] Zhao X and Jia M 2020 A novel unsupervised deep learning network for intelligent fault diagnosis of rotating machinery Structural Health Monitoring 19(6), 1745-1763.
[11] Zhang W, Li X, Jia X D, Ma H, Luo Z and Li X 2020 Machinery fault diagnosis with imbalanced data using deep generative adversarial networks Measurement 152 107377.
[12] Sun W, Shao S, Zhao R, Yan R, Zhang X and Chen X 2016 A sparse auto-encoder-based deep neural network approach for induction motor faults classification Measurement 89 171–178.
[13] Sun W, Zhao R, Yan R, Shao S and Chen X 2017 Convolutional discriminative feature learning for induction motor fault diagnosis IEEE Transactions on Industrial Informatics 13(3) 1350-1359.
[14] Cipollini F, Oneto L, Coradu A and Savio S 2019 Supervised learning for a novel deep learning network for feature learning in bearing condition monitoring Data-Enabled Discovery and Applications 3(1) 1.
[15] Zhao X, Jia M, Ding P, Yang C, She D and Liu Z 2020 Intelligent fault diagnosis of multichannel motor-rotor system based on multimanoifold deep learning machine IEEE/ASME Transactions on Mechatronics 25(5) 2177-2187.
[16] Gangsar P and Tiwari R 2018 On-line diagnostics of mechanical and electrical faults in induction motor using multiclass support vector machine algorithms based on frequency domain vibration and current signals ASME J. Risk Uncertainty Part B 5(3) 031001.
[17] Gangsar P and Tiwari R 2016 Taxonomy of induction-motor mechanical-fault based on time-domain vibration signals by multiclass SVM classifiers Intelligent Industrial Systems 2(3) 269-281.
[18] Agarap A F 2018 Deep learning using rectified linear units (relu). arXiv preprint arXiv:1803.08375.