International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 06 | June 2019 www.irjet.net p-ISSN: 2395-0072

Fault Detection and Maintenance Prediction for Gear of an Industrial Gearbox using Machine Learning Approaches

Er. Ashish Kumar Srivastava[1], Er. Abhishek Singh Gangwar[2]

[1]M.Tech. Scholar, Dept. of Mechanical Engineering, Goel Institute of Technology & Management, Lucknow, U.P., India
[2]Assistant Professor, Dept. of Mechanical Engineering, Goel Institute of Technology & Management, Lucknow, U.P., India

---

Abstract - *Gear can be termed as a mechanical component which is used to transfer amount of force and torque generated by an engine (as in automobile) or from an electric motor (as in industrial machine) to the desired output. It can be termed as an integral part for any gearbox. Gear needs to be monitored very carefully for any kind of signs which can result in failure. For this many a times machine is needed to be under shutdown so that proper inspection can be done, this increases time of unproductive work by machine. To avoid this time of unproductive work machine learning algorithms can be used to provide a prediction based for a best suited maintenance type according to the condition of gear so that it can be done to avoid failure of the component. All experiments have been performed on the gear dataset. This work explores and analyze the different machine learning algorithms named K-nearest neighbor algorithm, decision tree, random forest algorithm, support vector machine algorithm and multilayer perceptron algorithm. Waikato Environment for Knowledge Analysis (WEKA) tool is used to perform these experiments. According to our experiments, random forest algorithm is performing the best with 89.15% accuracy and with 0.172 root mean square error.*

Key Words: Fault Detection, Gear, Machine Learning, WEKA, Decision tree, random forest algorithm, support vector machine algorithm.

1. INTRODUCTION

In a gearbox there is always a surface contact between gears so that the torque can be transmitted by them. This results in many factors such as friction, vibration, heat generation due to friction, etc. that may lead to failure of gear like bending, pitting, scoring, abrasive wear, corrosive wear. For this material to be selected must have wear resistance, good surface hardness, etc. and the best suited material for any industrial gear is steel. But inspite of all the measures (such as use of lubrication, proper ventilation, etc.) failure may creep into gear, for this a proper and systematic inspection is required.

This may be quite a hectic process for both man and machine as no work can be taken while inspection thus, reducing productivity for a certain time. To reduce such loss Machine Learning concept can be used for monitoring condition of gear while it is still in service and a suitable maintenance type can be predicted according to the condition of gear.

Many researchers have used machine learning techniques for gear fault detection. Praveen Kumar et al. [2] used vibration signals as features and for fault detection support vector machine is used. Biswanath Samanta [3] also did experiments using artificial neural network and SVM. Sreenath et al. [4] have used Navie Bayes and decision tree algorithm-s for classification. Many research articles used tree-based methods like decision tree and SVM algorithm.

In this work, we are aiming to approach the machine learning algorithm of different type. K-NN algorithm is selected form instance-based learning algorithms. Two algorithms are selected from tree-based algorithms named decision tree and random forest algorithm. SVM algorithm selected form kernel-based algorithm and from deep learning multilayer perceptron algorithm is selected.

The main objective of this paper is to predict the fault detection with higher accuracy. To achieve this, we have done experiment with different machine learning algorithms and one deep learning algorithm. This paper also includes the comparative study of these algorithms.

This paper is organised as follow: Section 2 describes the machine learning algorithms that used in experiments. Section 3 discussed the experimental results. Section 4 concludes this work and last section 5 explains the future work.

2. MACHINE LEARNING ALGORITHMS

This section describes the different machine learning algorithms.

2.1 K-Nearest Neighbor Algorithm

K-NN is an Instance based machine learning algorithm and also know as lazy learning algorithm. There is no training phase, at the time of inference an instance search for a K nearest neighbors based on the distance metric like Euclidean, Manhattan, city block etc. The value of k is positive and odd number [7].

2.2 Decision Tree

Decision tree is tree-based algorithm and also known as J48 algorithm. This tree is consisting of node and edges where node is feature and edge represents the features value. This tree can be build using information gain method. Information gain method selects the best feature and this feature is selected as root node. After that search for a feature that is with root feature performing best using information gain method [6].

2.3 Random Forest Algorithm

Leo Breiman has presented this algorithm and this is improved version of decision tree. Decision tree algorithm uses information gain method to find the best feature whereas random forest algorithm gives equal priority to all features. So, this algorithm will create as many trees as many features a dataset has. At the time of inferencing, a new instance go through all the trees and every tree will predict the class label and majority voting method is used to find the final class label. This algorithm outperforms to as compare to decision tree algorithm.

2.4 Support Vector Machine

Support vector machine is one of the famous algorithms. This algorithm separates out two classes with maximum margin. This algorithm work on the principle of linear separable dataset [8]. In this work, we have used optimized SVM algorithm named SMO [9]. Based on this SVM algorithm can be of two types:

SVM linear classifier: If dataset is linearly separable and SVM algorithms predict the linear hyper plane.

SVM nonlinear classifier: If dataset is not linearly separable and SVM predicts the nonlinear hyper plane. Nonlinear hyperplane can be polynomial kernel, radial basis kernel.

2.5 Multi-Layer perceptron

Multilayer perceptron is deep learning algorithm uses the concept of biological nervous system. Simple neural network algorithm or perceptron algorithm work on the principle of the linear separability [10].

If dataset is not linear separable then multiplayer perceptron is used. In multilayer perceptron algorithm, hidden layers introduced in between input layer and output layer.

3. EXPERIMENTAL RESULTS

This section discusses experimental results for gear fault detection. This section is subdivided into different subsections named as Subsection 3.1 Dataset description. Subsection 3.2 explains the performance metric used for evaluation of the machine learning algorithms. Subsection 3.3 explains experimental setup of this work. Subsection 3.4 discuss the results of different machine learning algorithms.

3.1 Dataset Description

This work uses the gear fault detection dataset provided by [5]. This dataset has ten features named as:

Table 1: Feature of Gear dataset

1: Motor Vibration	2: Vibration of planetary gearbox in x direction
3: Vibration of planetary gearbox in y direction	4: Vibration of planetary gearbox in z direction
5: Motor torque	6: Vibration of parallel gear box in x direction
7: Vibration of parallel gear box in y direction	8: Vibration of parallel gear box in z direction
9: Rotating speed of motor spindle	10: Applied load

This dataset has five class labels. Table 2 has shown the names as:

Table 2: Class labels of Gear Dataset

Class label 1: Chipped (crack in gearbox)
Class label 2: Miss (missing feet in gear)
Class label 3: Root (crack in root of gear)
Class label 4: surface (crack in surface of gear)
Class label 5: Heathy gear

3.2 Performance parameters

Evaluating the machine leaning algorithm is an machine learning algorithms, we used accuracy and root mean square error.

Accuracy can be calculated with the number of instances that classified correctly divided by number of total instances. Confusion matrix is used to show how well instances are classified in each class labels.

3.3 Experimental Setup

All experiments have been performed using WEKA tool [1]. WEKA is open source machine learning used for analysis. Arff file format is used in this tool as input of dataset. We have preprocessed the dataset and converted into arff format for experiments.

3.4 Results and Discussion

This section explains and discuss the results. Experiments have included four machine learning and one deep learning algorithm. Machine learning algorithms includes K-nearest

neighbor, decision tree algorithm, random forest algorithm and support vector machine algorithm and one deep learning algorithm name is multilayer perceptron algorithm. These results are obtained using 10-fold cross validation on gear dataset.

K-NN algorithms used K is equal to five mean it considers five neighbors to make final decision. Decision tree algorithm uses information gain method to build tree. MLP uses 20 hidden layers.

Figure 1 shows the accuracy of different machine learning algorithms where K-NN algorithm achieves least accuracy of 80.26 %. Decision tree algorithm achieves 87.57%. SVM and MLP algorithm achieves almost similar accuracy 88.79% and 88.80 % respectively. Random forest algorithm achieves best actuary of 89.15%.

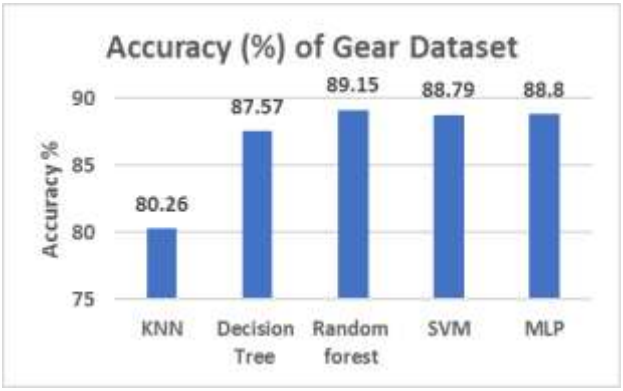

Figure 1: Accuarcy of Gear Dataset

Measuring accuarcy of machine learning algorithms is not apropriate metric, it is also important to find the error of each algorithm to make final deciosn for best algorithm. Referring to Figure 2 Random forest algorithm has least RMSE of 0.172 whereas SVM has highest error of 0.323. Referring to Figure 1 and Figure 2, Random forest algorithm is performing best with 89.15% accuracy and 0.172 RMSE.

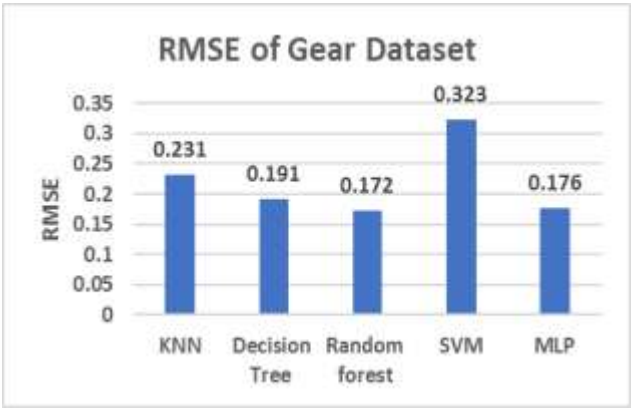

Figure 2: Root Mean Square Error of Gear Dataset

As experimental results show that random forest algorithm is best performing algorithm. Table 3 shows the confusion matrix where each class has 2002 total instance. According to random forest algorithm, how these instances are classified is shown in Table 3. Number in diagonal of matrix and in bold font shows the correct classification.

Table -3: Confusion matrix of Random forest algorithm

Actual/ predicted	Chipped (C)	Miss (M)	Root (R)	Surface (S)	Healthy (H)
Chipped (C)	**1637**	19	319	0	27
Miss (M)	10	**1902**	6	84	0
Root (R)	346	4	**1652**	0	0
Surface (S)	0	86	0	**1865**	51
Healthy (H)	34	0	0	100	**1868**

For example, class label Chipped, random forest algorithm correctly classified 1637 instance out of 2002 instances. Similarly, for class label healthy, 1868 instances are correctly classified out of 2002 with random forest algorithm and 100 instances are misclassified as surface and 34 are misclassified as chipped.

Table 4 shows the class level accuracy of the random forest algorithm. It is important to track the class level accuracy of the machine learning algorithm. Referring to Table 4, Class label Miss achieved maximum accuracy of 95%. Class label surface and Healthy achieved around ~93% accuracy.

Table -4: Class Label Accuracy of Random Forest Algorithm

Class Label	Class Label Accuracy
Chipped (C)	81.77%
Miss (M)	95.00%
Root (R)	82.52%
Surface (S)	93.16%
Healthy (H)	93.31%

4. CONCLUSION

Experimental results show the result of the different machine learning algorithm named K-NN, decision tree, random forest, SVM and MLP algorithms. According to results Random forest algorithm is performing best with 89.15% accuracy and 0.172 RMSE value. After that MLP algorithm is performing with 88.8% accuracy and 0.176

RMSE value. SVM algorithm is not performing well because it encountered higher error. SVM achieves higher accuracy of 88.79% but RMSE value is high 0.323. These results have sown that tree-based algorithms are performing best.

5. FUTURE SCOPES

This work can be extened by using deep learning algorithm like using convolution neural network-based algorithms. It may also be extended to develop a simulation model to predict and diagnose the health condition for gear in an industrial gearbox.

REFERENCES

[1] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

[2] Praveenkumar, T., M. Saimurugan, P. Krishnakumar, and K. I. Ramachandran. "Fault diagnosis of automobile gearbox based on machine learning techniques." Procedia Engineering 97 (2014): 2092-2098.

[3] Samanta, Biswanath, Khamis R. Al-Balushi, and Saeed A. Al-Araimi. "Artificial neural networks and genetic algorithm for bearing fault detection." Soft Computing 10, no. 3 (2006): 264-271.

[4] Sreenath, P. G., Gopalakrishnan Praveen Kumare, Sundar Pravin, K. N. Vikram, and M. Saimurugan. "Automobile Gearbox Fault Diagnosis Using Naive Bayes and Decision Tree Algorithm." In Applied Mechanics and Materials, vol. 813, pp. 943-948. Trans Tech Publications, 2015.

[5] Shao, Siyu, et al. "Highly-accurate machine fault diagnosis using deep transfer learning." IEEE Transactions on Industrial Informatics (2018).

[6] S. K. Murthy, Automatic construction of decision trees from data: A multi-disciplinary survey, Data mining and knowledge discovery, vol. 2, no. 4, pp. 345-389, 1998.

[7] A.-L. Jousselme and P. Maupin, Distances in evidence theory: Comprehensive survey and generalizations, International Journal of Approximate Reasoning, vol. 53, no. 2, pp. 118-145, 2012.

[8] J. C. Platt, 12 fast training of support vector machines using sequential minimal optimization, Advances in kernel methods, pp. 185-208, 1999.

[9] Wu, Xiujun. "Support Vector Machine SMO Algorithms and Their Optimization." In 3rd International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2019). Atlantis Press, 2019.

[10] Marsland, Stephen. Machine learning: an algorithmic perspective. Chapman and Hall/CRC, 2011.