IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)
e-ISSN: 2278-1684,p-ISSN: 2320-334X, Volume 16, Issue 5 Ser. III (Sep. - Oct. 2019), PP 21-24
www.iosrjournals.org

Bearing Fault Diagnosis by Multinomial Logistic Regression Classifier

Sandip Kumar Singh [1]
[1] *Associate professor, Department of Mechanical Engineering, V B S Purvanchal University, Jaunpur,222001*
Corresponding Author: Sandip Kumar Singh

Abstract: *The bearing is an essential part of all rotating machinery. The early detection of faults in bearing by using vibration signals saves a significant amount of financial loss. Many approaches have been used to overcome this problem in the past, but they succeeded to only some extent to identify the faults occurring on the outer race, inner race, and rollers/balls. None of them is capable of diagnosing these faults accurately. Machine Learning-based Data-driven methods have shown better results as compared to signal processing-based techniques. In this paper we are presenting a Multinomial Logistic Regression (MLR) method, which is capable of detecting faults with an accuracy of 99.80%. The results are shown in terms of Area Under Curve (AUC) of the precision-recall curve. A comparison with the Support Vector Classifier (SVC) is also shown.*

Keywords: *Support Vector Classifier (SVC), Multinomial Logistic Regression (MLR), Area Under Curve (AUC)*

Date of Submission: 02-10-2019 Date of Acceptance: 18-10-2019

I. Introduction

Bearing fault detection, using vibration data has gained the focus of mechanical engineers. Samanta and Al-Balushi [1] and Samanta *et al.* [1] analyzed the third and fourth central moments, i.e., skewness and kurtosis. The authors employed ANN and SVM for the diagnosis of bearing faults and summarized that both the even and odd moments are equally capable of representing bearing health effectively. Abbasion *et al.* [2] classified the single level fault severities in rolling element bearings. The authors employed the wavelet transform (WT) for the denoising of vibration signals. The classification has been performed using SVM, and faults in various components have been classified. Liu *et al.* [5] and Bordoloi *et al.* [3] employed SVM for the fault classification in rolling contact elements. Various statistical features viz. kurtosis, standard deviation, range, mean value are also utilized for the diagnosis of rolling element bearings by Kankar *et al.* [7]; Gangsar *et al.* [4] and Yaqub *et al.* [6]. These investigations utilized SVM and ANN (Artificial Neural Networks) and proposed that the selected features are sensitive to provide considerable fault identification accuracy. Now, we propose a new approach based on augmented data processing before the MLR. The sequential augmentation of data improves the performance of MLR significantly.

II. The Proposed Method

Multinomial Logistic Regression (MLR) is a further extension of Logistic regression, which is a binary classifier. MLR classifies multiclass data both linearly and nonlinearly. Primary Logistic regression classifier is useful in binary situations like on /off, fail /pass, live/ dead, etc., where classes are discretely separable. In such situations, the classes can be labeled between "0" or "1". Multinomial Logistic Regression is used for the dependent variables, which are categorically equivalent. This categorical equivalence of dependent variables is also called nominality. This multiclass Logistic Regression is also called Multinomial Logistic Regression. When we have to choose one of the variables out of many, which cannot be ordered in any way, multinomial Logistic regression is very useful to predict the result. MLR is a linear predictor, which uses different features as explanatory variables, with a set of weights multiplied linearly. The utility score is defined for a person i chooses some outcome k in the following way:

$$\text{Utility Score } (\boldsymbol{X_i}, \boldsymbol{k})) = \boldsymbol{\beta_k} . \boldsymbol{X_i}$$

The softmax function (k,$x_1 x_2, \dots \dots x_n$) is defined as,

$$\text{softmax } (\mathbf{k}, \boldsymbol{x_1}, \boldsymbol{x_2}, \dots \dots, \boldsymbol{x_n}) = \frac{e^{x_k}}{\sum_{i=1}^{n} e^{x_i}}$$

The softmax function (k, $x_1 x_2, \dots \dots x_n$) returns a value close to 0, whenever x_k is significantly less than the highest of all the values, and returns a value near to 1 when applied to the highest value unless it is extremely close to the next-largest value.

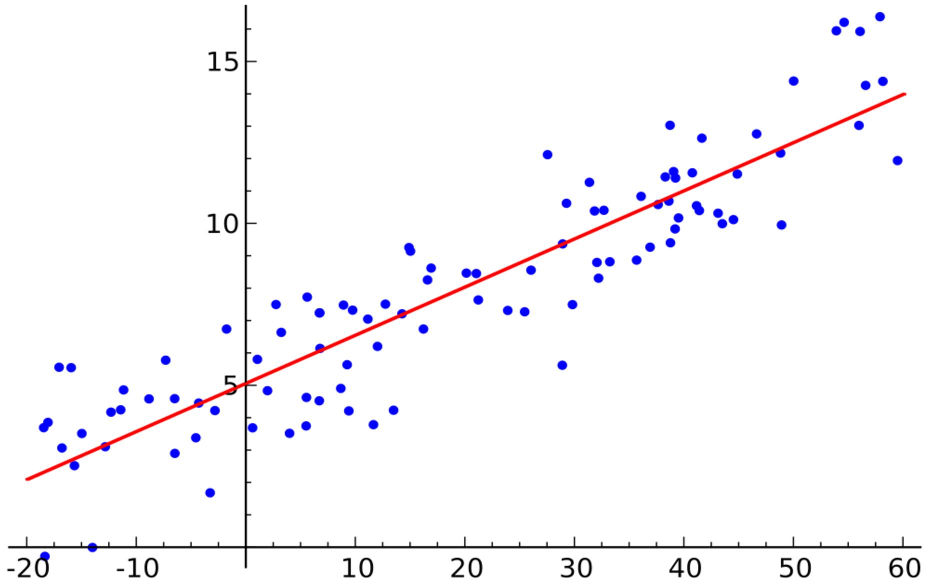

Fig 1. Linear Regression Model

The prediction probability is calculated as,

$$P_r(Y_i = K) = \frac{1}{1 + \sum_{k=1}^{k=K-1} e^{\beta_k' X_i}}$$

III. Results and Analysis

Table-1 shows the precision-recall parameters along with training and testing accuracy of the proposed method. We observe that all performance parameters improve significantly with augmented data SVC. The Figure-2 shows the precision-recall curve for MLRwith 100%,200% and 400% data augmentation. We get maximum AUC (area under the curve), and testing accuracy (99.80%) for 400% augmented data MLR. The Support Vector Classifier (SVC) shows the accuracy of 95.5%.

Thus, the MLR based method provides better classification performance. Table-2, showing the confusion matrix for test data**,** also validates the capability of bearing fault classification. The results have been validated by applying on open source data of Case Western University data repository. Class 0 is for normal bearing, 1 is for outer race fault, 2 is for inner race fault, and 3 is for roller fault.

Table -1: Performance of SVC and MLR

MLR (Multinomial Logistic Regression) 100% data					**MLR (Multinomial Logistic Regression) 200% data**				
Training Accuracy: 0.955					Training Accuracy: 0.982				
Class	**precision**	**recall**	**f1-score**	**support**	**Class**	**precision**	**recall**	**f1-score**	**support**
0	0.96	1	0.98	55	0	1	1	1	148
1	0.97	0.97	0.97	147	1	0.98	0.99	0.99	298
2	0.92	0.92	0.92	200	2	0.99	0.96	0.98	419
3	0.95	0.95	0.95	310	3	0.98	0.99	0.99	559
avg / total	0.95	0.95	0.95	712	avg / total	0.98	0.98	0.98	1424
MLR (Multinomial Logistic Regression) 400% data					**SVC**				

Training Accuracy: 0.994					Training Accuracy: 0.958				
Class	**precision**	**recall**	**f1-score**	**support**	**Class**	**precision**	**recall**	**f1-score**	**support**
0	1	1	1	276	0	0.96	1	0.98	55
1	1	1	1	582	1	0.97	0.96	0.97	147
2	0.98	1	0.99	868	2	0.94	0.94	0.94	200
3	1	0.99	0.99	1122	3	0.96	0.96	0.96	310
avg / total	0.99	0.99	0.99	2848	avg/ total	0.96	0.96	0.96	712

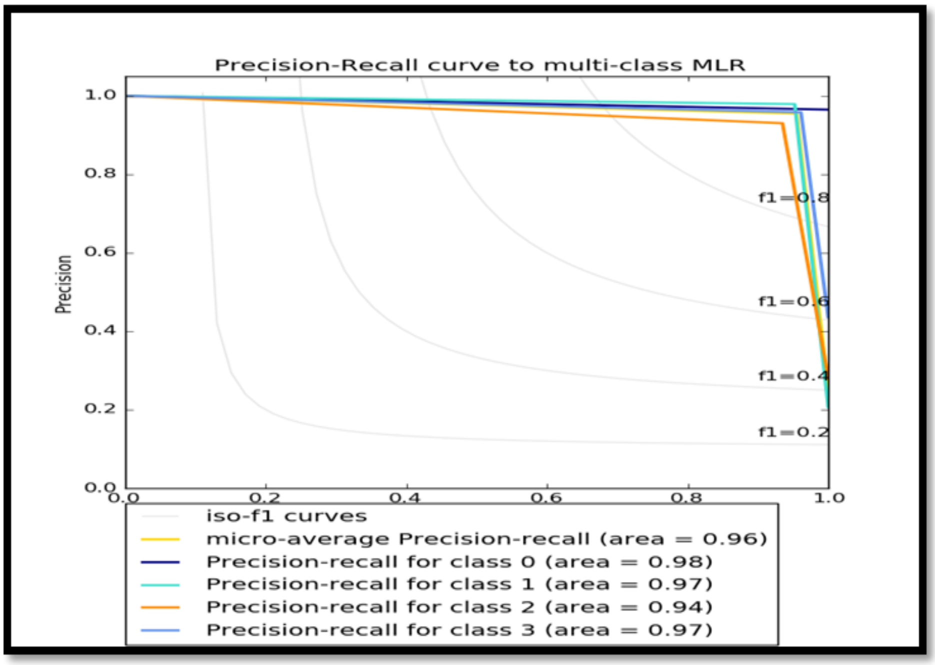

Fig 2.Precission -recall curve for MLR

Table-2: Confusion matrix for MLR and SVC for bearing fault

CONFUSION MATRIX										
MLR 100% data						MLR 200% data				
Class	0	1	2	3		Class	0	1	2	3
0	55	0	0	0		0	148	0	0	0
1	0	142	2	3		1	0	296	0	2
2	0	5	184	11		2	0	7	403	9
3	2	0	13	295		3	0	0	4	555
Testing accuracy =94.9438202247191						Testing accuracy= 98.455056179775283				
MLR 400% Data						SVC				
Class	0	1	2	3		Class	0	1	2	3
0	276	0	0	0		0	55	0	0	0
1	0	582	0	0		1	0	141	3	3
2	0	0	867	1		2	0	3	187	10
3	0	0	14	1108		3	2	1	10	297
Testing accuracy = 99.77733146067415						Testing accuracy = 95.50561797752809				

IV. Conclusion

From the tables and figures shown above, it is observed that the proposed method is a versatile approach for bearing fault identification at the early stages. It is equally capable of detecting faults on the outer and inner race along with roller faults. The machine learning performance metrics show high precision and recall simultaneously. It validates the predictive capability of the SVC based proposed approach. The method is simple and less time-consuming. The pre-processing technique used is also novel and can be used for the smaller size of available data. For future work, the proposed method can be tested for compound faults of bearings.

References

[1]. Samanta, B., Al-Balushi, K.R. and Al-Araimi, S.A., Artificial Neural Networks and Support Vector Machines with Genetic Algorithm for Bearing Fault Detection, *Engineering Applications of Artificial Intelligence*, Vol. 16(7-8), 2003, 657-665.
[2]. Abbasion, S., Rafsanjani, A., Farshidianfar, A. and Irani, N., Rolling Element Bearings Multi-Fault Classification Based on the Wavelet Denoising and Support Vector Machine, *Mechanical Systems, and Signal Processing*, Vol. 21(7), 2007, 2933-2945.
[3]. Bordoloi, D.J., and Tiwari, R., Optimum Multi-Fault Classification of Gears with Integration of Evolutionary and SVM Algorithms, *Mechanism and Machine Theory*, Vol. 73, 2014 (a), 49-60.
[4]. Gangsar, P. and Tiwari, R., Multiclass Fault Taxonomy in Rolling Bearings at Interpolated and Extrapolated Speeds Based on Time Domain Vibration Data by SVM Algorithms, *Journal of Failure Analysis and Prevention*, Vol. 14(6), 2014, 826837.
[5]. Liu, Z., Cao, H., Chen, X., He, Z. and Shen, Z., Multi-Fault classification based on Wavelet SVM with PSO Algorithm to Analyze Vibration Signals from Rolling Element Bearings, *Neurocomputing*, Vol. 99(1), 2013, 399-410.
[6]. Yaqub, M.F., Gondal, I. and Kamruzzaman, J., Machine Fault Severity Estimation based on Adaptive Wavelet Nodes Selection and SVM, in Proceedings of the IEEE International Conference on Mechatronics and Automation, Beijing, China, 2011, 1951-1956.
[7]. Kankar, P.K., Sharma, S.C. and Harsha, S.P., Fault Diagnosis of Ball Bearings using Machine Learning Methods, *Expert Systems with Applications*, Vol. 38(3), 2011(b), 1876-1886.

Sandip Kumar Singh. "Bearing Fault Diagnosis by Multinomial Logistic Regression Classifier." IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE) , vol. 16, no. 5, 2019, pp. 21-24