https://doi.org/10.1051/itmconf/20246301004

Bearing Health Evaluation Model using Segmentive Technique and Cosine KNN in Different Loading Situations

J.S. Yap[1], M.H. Lim[1], and M. Salman Leong[1]
[1]1Institute of Noise and Vibration, Universiti Teknologi Malaysia, Kuala Lumpur 54100, Malaysia

Abstract. Bearing faults are a common cause of machinery failure, and bearing vibration analysis is critical in preventing any unacceptable consequences of such failures. Advancements in smart data and computing make Artificial Intelligence (AI) preferable for bearing vibration analysis. Typically, signal processing and feature engineering are essential for achieving satisfactory classification accuracy. Additionally, a drop in classification accuracy is commonly observed during different loading situations due to the vastly varying vibration characteristics under different loads. This paper evaluates an AI model in variable loading situations using raw vibration signals, devoid of signal processing and feature engineering. The proposed AI model, Segmentive Cosine K-Nearest Neighbours (SCosKNN), demonstrated a higher overall classification accuracy of 90.6-94.3% in same loading situations, and 72.1-84.2% in different loading situations. An improvement of around 9% in same loadings and 10-14% in different loadings were observed compared to a model without Segmentive Technique

1 Introduction

Bearing faults can lead to machinery breakdowns, disrupting operations and incurring financial costs and downtime[1], [2]. The bearing plays a crucial role in machinery by enabling the efficient rotation of a shaft. A typical bearing comprises several elements, as shown in **Fig. 1**. Faults in these bearing elements typically develop over time, affecting the Inner Race, Outer Race, and Ball elements [3], [4].

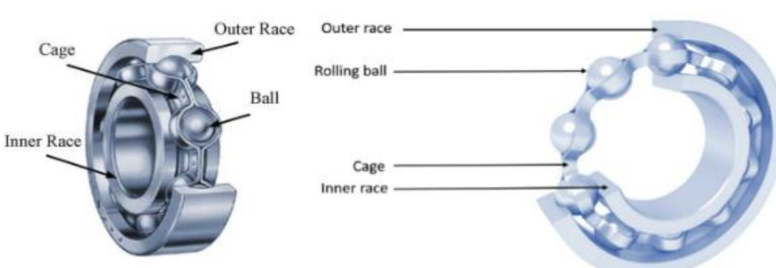


Fig. 1. The elements of bearing: Outer Race, Inner Race, Cage and Ball [3], [4]

Statistically, the primary cause of equipment failures is bearing faults, accounting for as much as 55%[3]–[5]. Equipment failures can lead to financial losses, service downtime, and even health and safety concerns[5], [6]. Fault analysis and condition monitoring techniques, such as bearing vibration analysis, are commonly applied in prevention and predictive maintenance[3], [6]. Artificial Intelligence (AI) is gaining attention in bearing vibration analysis, as human-based analysis heavily relies on individual knowledge and experience[3], [6]. This paper evaluates an AI method using Segmentive Cosine K-Nearest Neighbours (SCosKNN). The evaluations are conducted on raw vibration data without signal processing and feature engineering to minimize human involvement.

2 Challenges of Bearing Fault Analysis

Most AI models are developed using a single or limited set of operating loads, which is not reflective of the realistic and noisy conditions encountered in actual operations[3]. Conventionally, stable operating conditions are assumed in the analysis of vibration signals, while working operations are complex and subject to constant change in the real world[7], [8]. Transfer learning has been explored to address this issue. However, it still requires target domain data or Deep Domain Adaptation (DDA), with the exception of zero-shot methods that create samples of unseen classes through generative models[9], [10].

3 Methodology

This paper proposes an AI-based model using the Segmentive Technique on raw vibration data and evaluates the fault classification performance against a model without the Segmentative Technique. The classification performance is assessed in same and different loadings to reflect practical applicability in the industry.

3.1 Segmentive Technique

A Segmentive CosKNN (SCosKNN) classification model is projected to enhance classification accuracy. This model integrates a Segmentive Technique with a CosKNN classifier. The data samples are segmented and arranged as illustrated in **Fig. 2**.

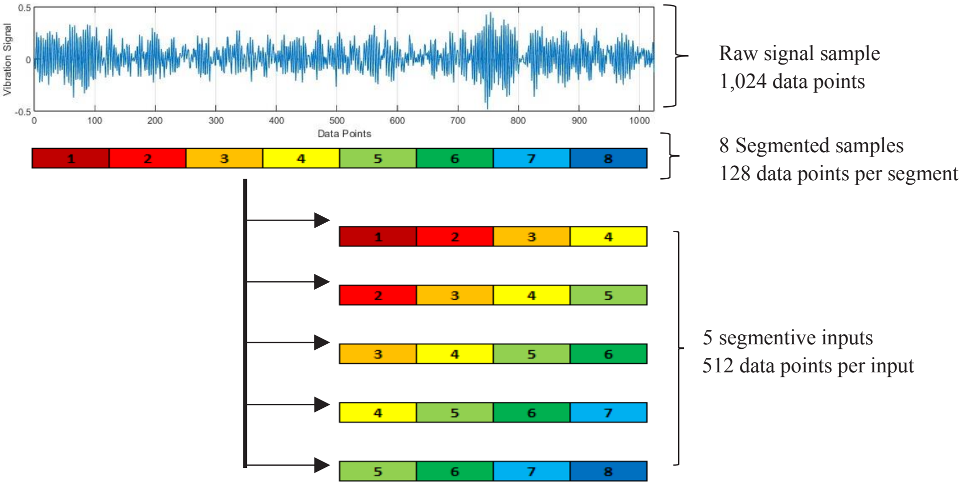


Fig. 2. Segmentive technique of the proposed SCosKNN model

As shown in **Fig. 2**, the raw signal sample which containing 1024 data points is segmented into 8 divisions. The segmented divisions are clustered with into 5 segmentive input which containing 512 data points each. As an outcome of Segmentive Technique, the training data is multiplied by a factor of 5.

3.2 Cosine KNN

Cosine K-Nearest Neighbours (CosKNN) is based on the distance metric of Cosine Distance [11]as follows

$$
\begin{aligned}
Dc\ (A, B) &= 1 - Sc(A, B) \qquad (1) \\
&= 1 - \frac{A\,.B}{||A||\ .\ ||B||} \\
&= 1 - \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2\ .\ \sum_{i=1}^{n} B_i^2}}
\end{aligned}
$$

Dc = Cosine Distance

Sc = Angular Similarity

A = Tested Sample in Vector form

B = Trained Sample in Vector form

i= *i*-th components of Vectors

The CosKNN algorithm is described in **Table 1**, the Cosine Distance (*Dc*) are calculated. The default nearest number value (K) is set as 10 according to MATLAB's Classification Learner App [12]. The top 10 Ks with lowest *Dc* value are selected, subsequently the label of selected Ks are called out. Each called out label are counted and the final predicted class is based on the highest occurrence.

Table 1. The algorithm of CosKNN

CosKNN algorithm
Input: *A, l, B* // *A*: testing data; *l*: label; *B*: trained data **for** ***j*** **to training data size do:** Calculate the cosine distance *Dc*(A,*B*) as described in *equation (1)* **end for** Select the desired K (number of nearest neighbors) Sort the distances by increasing order Count the occurrences of each label among the top K **Output:** Assign *A* to the highest frequency of label *l*

3.3 Datasets Description

The study utilizes the Case Western Reserve University Bearing Data Center(CWRU)[13] which incorporates two distinct bearings positioned at the Drive End (DE) and Fan End (FE), subjected to variable loading conditions. The dataset details are provided in **Fig. 3** and **Table 2**. The loading conditions of 0, 1, 2 and 3hp were simulated by adjusting the dynamometer.

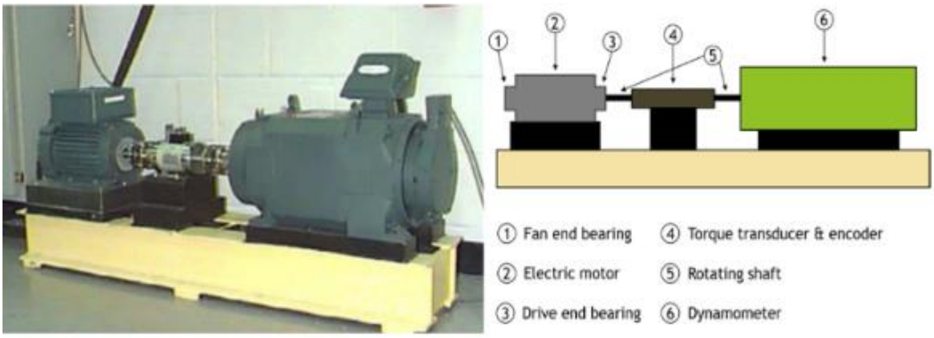


Fig. 3. The CWRU bearing datasets used in this paper.

Table 2. Datasets descriptions in this paper.

Dataset	Drive End (DE)	Fan End (FE)
Bearing Model	6205-2RS JEM SKF, deep groove ball bearing	6203-2RS JEM SKF, deep groove ball bearing
Variable Loading	0hp, 1hp, 2hp, & 3hp	
Labelled Classes	Normal, Outer Race Fault, Inner Race Fault, & Ball Fault	

3.4 Assessment of Fault Classification Performance

The fault classification performance is conventionally assessed by classification accuracy as explained as following equation.

$$Accuracy\ (\%) = \frac{TP+TN}{TP+FP+FN+TN}\ X\ 100 \qquad (2)$$

TP = number of true positive classifications
TN = number of true negative classifications
FP = number of false positive classifications
FN = number of false negative classifications

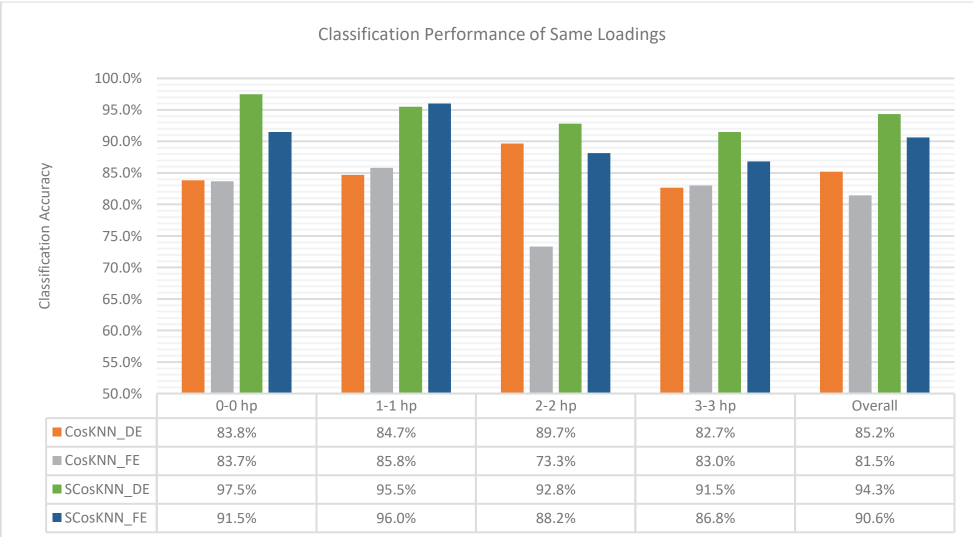
The evaluation is assessed in two scenarios: A. training and testing data are in the same loading conditions; B. training and testing data are in the different loading conditions. The SCosKNN's performance is compared with classical Cosine KNN without Segmentive Technique (CosKNN) and elaborated in section 4.

4 Results and Discussions

4.1 Results

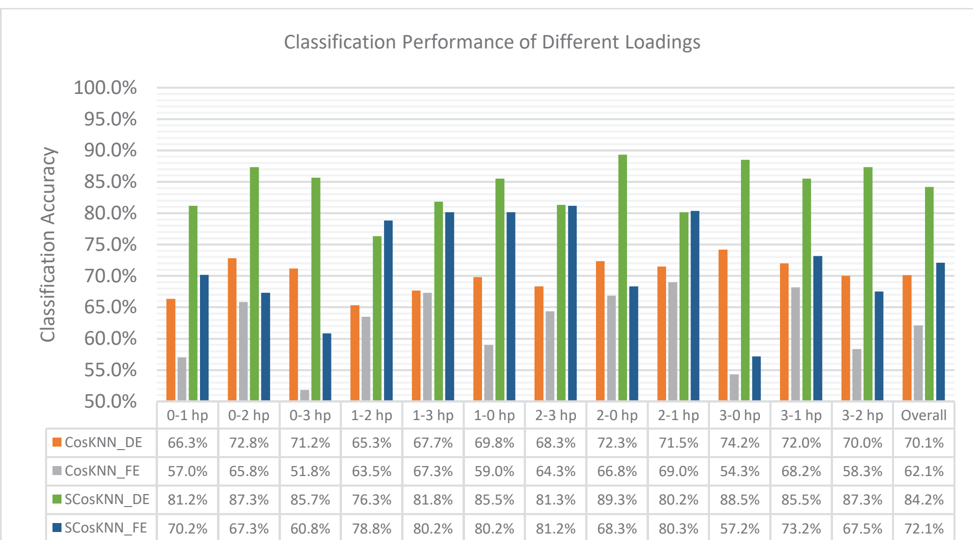
Fig. 4 presents the classification performance of CosKNN and SCosKNN of same loadings situation, while **Fig. 5** presents the situations of different loadings. As observed, the general classification results of different loading's situations are inferior compared to same loading situations, with reduced overall accuracy from 10.2% to 19.3%. This study finds that the

SCosKNN outperformed CosKNN due to assistance from Segmentive Technique, an average increase in accuracy from 9.1% to 14%.



	0-0 hp	1-1 hp	2-2 hp	3-3 hp	Overall
CosKNN_DE	83.8%	84.7%	89.7%	82.7%	85.2%
CosKNN_FE	83.7%	85.8%	73.3%	83.0%	81.5%
SCosKNN_DE	97.5%	95.5%	92.8%	91.5%	94.3%
SCosKNN_FE	91.5%	96.0%	88.2%	86.8%	90.6%

Fig. 4. The classification performance of CosKNN and SCosKNN of same loadings situation.



	0-1 hp	0-2 hp	0-3 hp	1-2 hp	1-3 hp	1-0 hp	2-3 hp	2-0 hp	2-1 hp	3-0 hp	3-1 hp	3-2 hp	Overall
CosKNN_DE	66.3%	72.8%	71.2%	65.3%	67.7%	69.8%	68.3%	72.3%	71.5%	74.2%	72.0%	70.0%	70.1%
CosKNN_FE	57.0%	65.8%	51.8%	63.5%	67.3%	59.0%	64.3%	66.8%	69.0%	54.3%	68.2%	58.3%	62.1%
SCosKNN_DE	81.2%	87.3%	85.7%	76.3%	81.8%	85.5%	81.3%	89.3%	80.2%	88.5%	85.5%	87.3%	84.2%
SCosKNN_FE	70.2%	67.3%	60.8%	78.8%	80.2%	80.2%	81.2%	68.3%	80.3%	57.2%	73.2%	67.5%	72.1%

Fig. 5. The classification performance of CosKNN and SCosKNN of different loadings situation.

4.2 Discussion

The possible reasons of fault classification in different loadings situations is due the rotational speeds are slightly different in different loadings due to higher resistance created by the dynamometer. The rotational speeds (rpm) correspond to loadings are shown in **Table 3**.

Table 3. The rotational speed corresponding to different loadings situation.

Loadings	Rotational Speeds (rpm)
0 hp	1797
1 hp	1772
2 hp	1750
3 hp	1730

The proposed SCosKNN exhibits a classification accuracy of over 90.6-94.3% in same loading situations, and 72.1-84.2% in different loading situations. This enhanced performance can be attributed to the Segmentive Technique and Cosine Distance characteristics, supported by the following reasons:

i. Reduces input dimensionality
The data points per input are halved from 1024 data points to 512 points as explained in **Fig.** ***2***, effectively reducing dimensionality. Referring to equation (1), higher dimensions (*i-th*) potentially increase combinatorial effects. For visualization, the *Dc* were calculated comparing 1024 data points and 512 points as shown in **Fig. 6**, one vibration sample of each class were randomly selected. The labels of different classes are shown in **Table 4**.

Table 4. The label correspondent to type of fault.

Label, *l*	Type of fault
B	Ball fault
I	Inner Race fault
N	Normal (no fault)
O	Outer Race fault

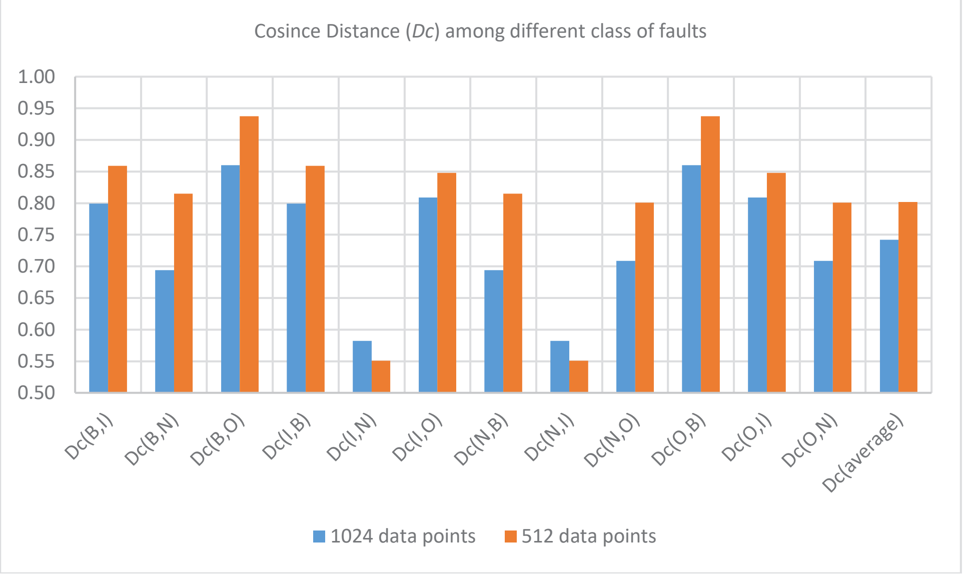


Fig. 6. An example of calculated Cosine Distance (*Dc*) for 1024 data points and 512 data points.

Since the calculated Dc are subjected to different classes for example, the higher value of *Dc* shows higher distinguishing capability between classes. Referring to **Fig. 6**, majority of *Dc* for 512 data points are higher than 1024 data points, hence reducing dimensionality could enhance the classification accuracy.

ii. Support collective multi-output.
Although the data points are reduced after Segmentive Technique, the remaining data points are still utilized by clustering into multiple divisions. As explained in section 3.1, the training and testing data are increased by 5-fold. Computation of multiple outputs from segmented inputs allows for subsequent consolidation to establish a collective predicted output. The framework of SCosKNN is summarized in **Fig. 7**.

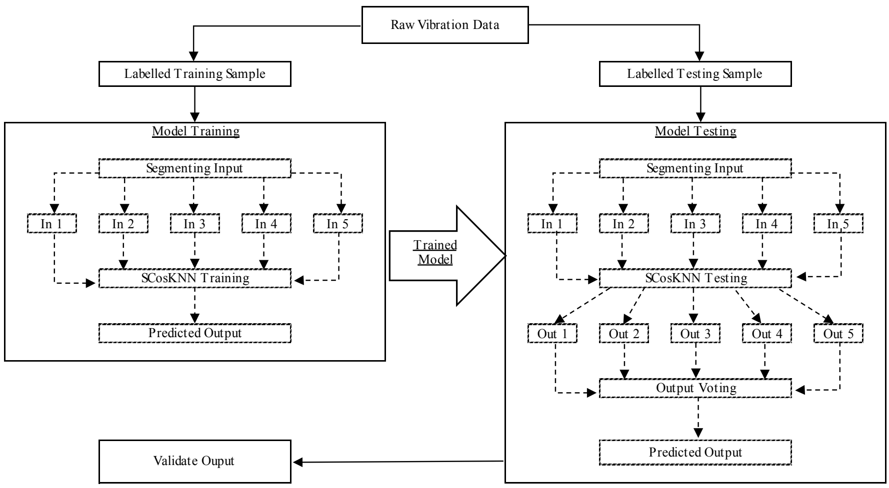


Fig. 7. Framework of proposed SCosKNN model and collective output process

5 Conclusions

This paper presents a bearing fault classification evaluation using a proposed AI model without signal processing and feature engineering under variable loading conditions. An enhanced model, incorporating the Segmentive Technique (SCosKNN), is investigated, am average classification accuracy of 90.6-94.3% in same loading situations, and 72.1-84.2% in different loading situations. An improvement of around 9% in same loadings and 10-14% in different loadings were observed compared to a model without Segmentive Technique. However, given the observed classification drop in loading situations, it is recommended that further improvements be made to SCosKNN in order to enhance its consistency and align it with the performance of transfer learning approaches.

The authors would like to extend their greatest gratitude to the Institute of Noise and vibration UTM for funding the current study under the Higher Institution Centre of Excellence (HICoE) Grant Scheme of Advanced Health Monitoring for Turbomachinery (R.K130000.7843.4J227) and Health Monitoring and Integrity Assessment of Ageing Assets (R.K130000.7843.4J228).

References

[1] A. Anwarsha and T. N. Babu, ‘Recent advancements of signal processing and artificial intelligence in the fault detection of rolling element bearings: a review’, *Journal of Vibroengineering*, vol. 24, no. 6, pp. 1027–1055, Sep. 2022, doi: 10.21595/jve.2022.22366.

[2] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, ‘Machine Learning and Deep Learning Algorithms for Bearing Fault Diagnostics -- A Comprehensive Review’, Jan. 2019, doi: 10.1109/ACCESS.2020.2972859.

[3] M. Hakim, A. A. B. Omran, A. N. Ahmed, M. Al-Waily, and A. Abdellatif, 'A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations', *Ain Shams Engineering Journal*, vol. 14, no. 4. Ain Shams University, Apr. 05, 2023. doi: 10.1016/j.asej.2022.101945.

[4] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, 'Deep Learning Algorithms for Bearing Fault Diagnosticsx - A Comprehensive Review', *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 29857–29881, 2020. doi: 10.1109/ACCESS.2020.2972859.

[5] M. Mohiuddin and M. S. Islam, 'Rolling Element Bearing Faults Detection and Classification Technique Using Vibration Signals †', *Engineering Proceedings*, vol. 27, no. 1, 2022, doi: 10.3390/ecsa-9-13339.

[6] S. R. Saufi, Z. A. Bin Ahmad, M. S. Leong, and M. H. Lim, 'Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review', *IEEE Access*, vol. 7. Institute of Electrical and Electronics Engineers Inc., pp. 122644–122662, 2019. doi: 10.1109/ACCESS.2019.2938227.

[7] C. Weng, B. Lu, Q. Gu, and X. Zhao, 'A novel hierarchical transferable network for rolling bearing fault diagnosis under variable working conditions', *Nonlinear Dyn*, vol. 111, no. 12, pp. 11315–11334, Jun. 2023, doi: 10.1007/s11071-023-08405-x.

[8] X. Liu, X. Zhao, and K. He, 'Feasibility Study of the GST-SVD in Extracting the Fault Feature of Rolling Bearing under Variable Conditions', *Chinese Journal of Mechanical Engineering (English Edition)*, vol. 35, no. 1, Dec. 2022, doi: 10.1186/s10033-022-00806-0.

[9] B. Yang and H. Sun, 'A zero-shot learning fault diagnosis method of rolling bearing based on extended semantic information under unknown conditions', *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 45, no. 1, Jan. 2023, doi: 10.1007/s40430-022-03965-2.

[10] R. Zhang and Y. Gu, 'A Transfer Learning Framework with a One-Dimensional Deep Subdomain Adaptation Network for Bearing Fault Diagnosis under Different Working Conditions', *Sensors*, vol. 22, no. 4, Feb. 2022, doi: 10.3390/s22041624.

[11] X. Liu, H. Huang, and J. Xiang, 'A personalized diagnosis method to detect faults in a bearing based on acceleration sensors and an fem simulation driving support vector machine', *Sensors (Switzerland)*, vol. 20, no. 2, Jan. 2020, doi: 10.3390/s20020420.

[12] M. Senthamil Selvi and S. Jansi Rani, 'Classification of Admission Data Using Classification Learner Toolbox', in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1979/1/012043.

[13] Case School of Engineering, 'Bearing Data Center', Case Western Reserve University Bearing Data Center. Accessed: Sep. 10, 2023. [Online]. Available: https://engineering.case.edu/bearingdatacenter