ITM Web of Conferences **63**, 01010 (2024) https://doi.org/10.1051/itmconf/20246301010
AMICT2023

Enhanced Bearing Fault Analysis under Inconstant Loads Conditions by Cosine Weighted K-Nearest Neighbours Model

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 often lead to machinery failures, underscoring the importance of analyzing bearing vibrations to avert undesirable consequences. Leveraging Artificial Intelligence (AI) in this context benefits from the strides in intelligent data processing and computing capabilities. Traditionally, signal processing and feature engineering play pivotal roles in achieving accurate classifications. However, classification accuracy can decline notably during variable loading scenarios due to the diverse vibration patterns exhibited under different loads. This study assesses an AI model's performance under variable loading conditions using raw vibration signals, without recourse to signal processing or feature engineering. Introducing an enhanced AI model, known as Cosine Weighted K-Nearest Neighbours (CWKNN), resulted in a slightly improved 85.2-88.7% under stable loading conditions and 64.3-72.6% under variable loading conditions.

1 Introduction

Bearing faults have the potential to cause machinery breakdowns, leading to operational disruptions, financial losses, and downtime[1], [2]. The bearing is a pivotal component in machinery, facilitating the smooth rotation of a shaft. Illustrated in **Fig. 1**, typical bearing encompasses several crucial elements. Faults in these components tend to manifest over time, predominantly affecting the Inner Race, Outer Race, and Ball elements [3], [4].

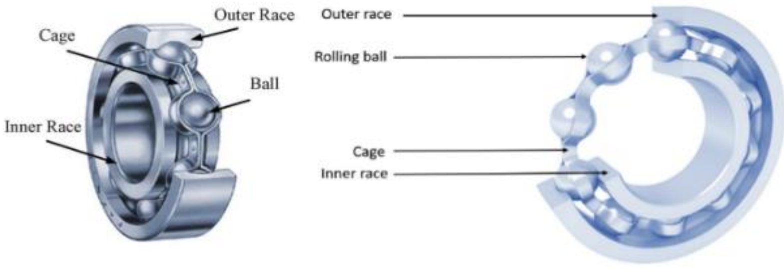


Fig. 1. The elements of bearing: Outer Race, Inner Race, Cage and Ball [3], [4]

Statistically, bearing faults account for a significant portion of equipment failures, with estimates as high as 55%[3]–[5]. Such failures can result in substantial financial losses, service interruptions, and even pose health and safety risks[5], [6]. Consequently, fault

analysis and condition monitoring techniques like bearing vibration analysis are commonly employed for preventive and predictive maintenance[3], [6]. Given the limitations of human-based analysis, there is a growing interest in the application of Artificial Intelligence (AI) in bearing vibration analysis[3], [6].
This study, we assess an AI approach by applying Weighting factor in Cosine K-Nearest Neighbors (CWKNN). Evaluations are performed on raw vibration data, deliberately eschewing signal processing and feature engineering to minimize human intervention.

2 Challenges of Bearing Fault Analysis

Many AI models are typically trained on a single or limited set of operating loads, which may not adequately represent the dynamic and noisy conditions prevalent in real-world operations[3]. Traditional analyses of vibration signals often assume stable operating conditions, whereas actual operational scenarios are characterized by complexity and continual fluctuations[7], [8]. Transfer learning has been explored as a solution, but it still necessitates target domain data or Deep Domain Adaptation (DDA), except for zero-shot methods that generate samples for unseen classes through generative models[9], [10].

3 Methodology

This paper introduces an AI-based model utilizing a segmentation process applied directly to raw vibration data. The study evaluates the model's performance in fault classification compared to common AI-based models. Furthermore, the classification performance is examined under varying load conditions to ensure practical relevance in industrial applications.

3.1 Cosine KNN

Cosine K-Nearest Neighbours (CKNN) 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

3.2 Weighted KNN

Weighted KNN (WKNN) represents a modified iteration of the conventional K-Nearest Neighbors algorithm, offering potential adjustments to bearing vibration analysis sensitivity. While retaining the same number of nearest neighbors (K) as traditional KNN[12], it introduces a weighting scheme in the voting process, contingent on the ranking of neighbors from closest to furthest. The specific weightages employed in this model are itemized in **Table 1**. A visualization of weighting procedure is explained in section 4.

Table 1. The nearest neighbours K's sequences and corresponding weightage of CWKNN.

K	1	2	3	4	5	6	7	8	9	10
Weightage	0.21	0.16	0.16	0.11	0.11	0.05	0.05	0.05	0.05	0.05

The CKNN algorithm is described in **Table 2**, the Cosine Distance (*Dc*) are calculated. The default nearest number value (K) is set as 10 according to MATLAB's Classification Learner App [13]. The 10 number of K with lowest *Dc* value are selected, subsequently the label of selected Ks are called out. Each called out label are counted and multiplied by its corresponding weightage. The final predicted class is based on the highest occurrence.

Table 2. The algorithm of CWKNN

CWKNN 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 Multiple the occurrence with correponding weightage **Output:** Assign *A* to the highest vote of label *l*

3.3 Datasets Description

This study leverages data from the Case Western Reserve University Bearing Data Center (CWRU)[14] encompassing two distinct bearings situated at the Drive End (DE) and Fan End (FE), subject to varying loading conditions. Comprehensive dataset particulars are elucidated in **Fig. 2** and **Table 3**.

Table 3. 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	

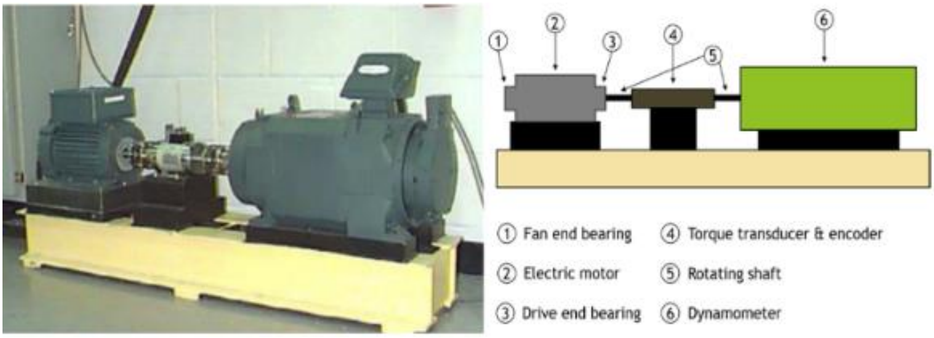


Fig. 2. The CRWU bearing datasets used in this paper.

3.4 Assessment of Fault Classification Performance

The fault classification performance is conventionally assessed by classification accuracy as explained as following equation[15].

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

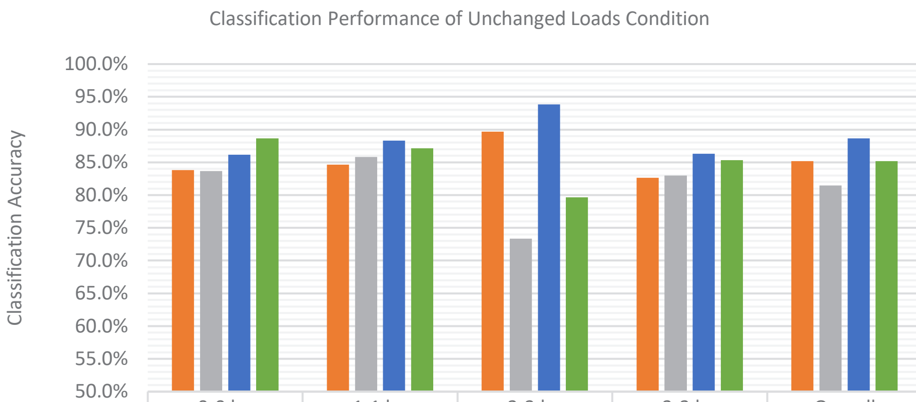
TP = number of true positive classifications
TN = number of true negative classifications
FP = number of false positive classifications
FN = number of false negative classifications

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 CWKNN's performance is compared with classical Cosine KNN without weightage (CKNN) and elaborated in section 4.

4 Results and Discussions

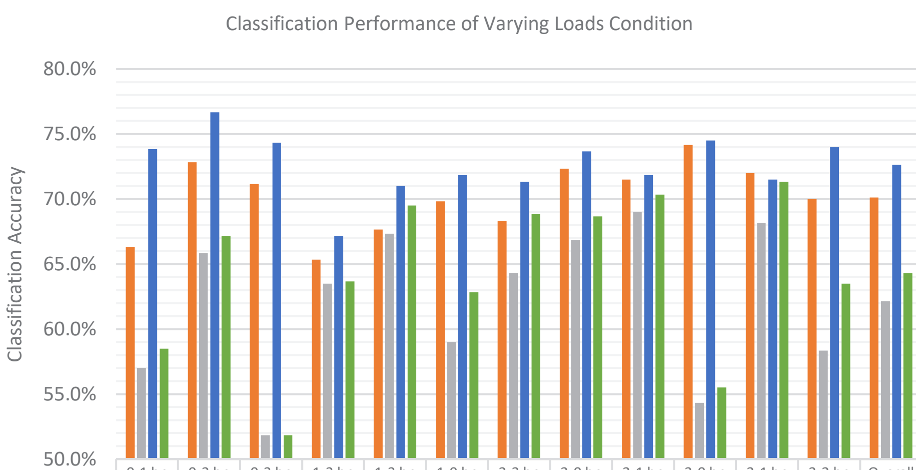
4.1 Results

Fig. 3 presents the classification performance of CKNN and CWKNN of unchanged loads condition, while **Fig. 4** presents the situations of varying loads condition. As observed, the general classification results of varying loads condition are inferior compared to unchanged loads condition, with reduced overall accuracy from 15.1% to 20.9%. This study finds that the CWKNN outperformed CKNN due to assistance from inclusion of weighting process, an average increase in accuracy from 2% to 3.7%.



	0-0 hp	1-1 hp	2-2 hp	3-3 hp	Overall
CKNN_DE	83.8%	84.7%	89.7%	82.7%	85.2%
CKNN_FE	83.7%	85.8%	73.3%	83.0%	81.5%
CWKNN_DE	86.2%	88.3%	93.8%	86.3%	88.7%
CWKNN_FE	88.7%	87.2%	79.7%	85.3%	85.2%

Fig. 3. The classification performance of CKNN and CWKNN of unchanged loads condition



	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
CKNN_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%
CKNN_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%
CWKNN_DE	73.8%	76.7%	74.3%	67.2%	71.0%	71.8%	71.3%	73.7%	71.8%	74.5%	71.5%	74.0%	72.6%
CWKNN_FE	58.5%	67.2%	51.8%	63.7%	69.5%	62.8%	68.8%	68.7%	70.3%	55.5%	71.3%	63.5%	64.3%

Fig. 4. The classification performance of CKNN and SWKNN of varying loads condition.

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 4**.

Table 4. The rotational speed corresponding to different loads condition

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

The proposed CWKNN showcases improved classification accuracy levels of over 85.2-88.7% under stable loading conditions and 64.3-72.6% under variable loading conditions.

By introducing the Weighting of K corresponding K potentially allows the K with lower angular distance to gain higher influence in the overall bearing vibration analysis instead of the classical non-weighting KNN. To visualize the benefits of weighting process, an example of calculated Cosine Distance (*Dc*) is demonstrated as follows.

Table 5 Show the labelling of classes for different bearing's faults/normal states. A randomly selected testing signal was used to compute the *Dc* with a small-size trained and labelled dataset, the computed *Dc* is shown in **Fig. 5**. 10 numbers of Ks with lowest Dc are selected and its class's label are referred as shown in **Table 6** together with the unweighted factor for CKNN and weighted factor for CWKNN. The voted outputs computed by multiplying the class's occurrences and unweighted/weighted factors as highlighted in **Table** *7*. Both CKNN and CWKNN have the same and true classification of Ball fault in this example, however CWKNN procedures higher confidence of Class 'B' at 0.58 compared to 0.50 of CKNN. As shown in the results at **Fig. 3** and **Fig. 4**, CWKNN's classification performances are slightly higher than CKNN by 2% to 3.7%.

Table 5. The labelling of fault/normal classes

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

Table 6. The selected 10 Nearest Neighbors (K); unweighted factors for CKNN and weighted factors WCKNN

K	1	2	3	4	5	6	7	8	9	10
Unweighted factor in CKNN	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Weighted factor in CWKNN	0.21	0.16	0.16	0.11	0.11	0.05	0.05	0.05	0.05	0.05
Corresponded Label, *l*	IR	B	B	B	B	OR	OR	OR	OR	B

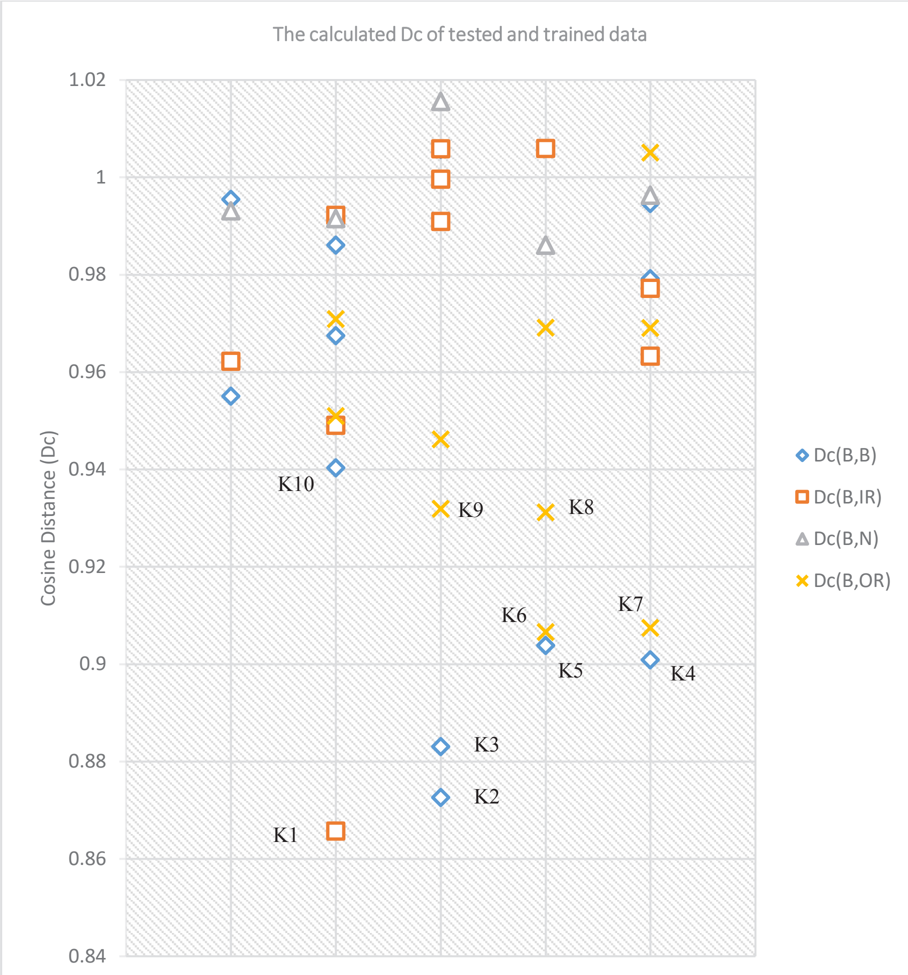


Fig. 5. The visualisation of calculated *Dc* example.

Table 7. The example of voted ouput for CKNN and CWKNN

Label, *l*	B	IR	N	OR	Voted Output
CKNN's vote	0.50	0.10	0.00	0.40	B
CWKNN's vote	0.58	0.21	0.00	0.21	B

5 Conclusions

This paper introduces an evaluation of bearing fault classification using a novel AI model that eschews signal processing and feature engineering, specifically under variable loads condition. The enhanced model, which integrates a weighting factor in a traditional cosine KNN, attains an 85.2-88.7% under stable loading conditions and 64.3-72.6% under variable loading conditions. Nevertheless, due to the observed reduction in classification performance in the context of variable loads scenarios, it is advised to consider further refinements to

CWKNN to bolster its robustness and bring it in line with the performance of transfer learning approaches.

Acknowledgement

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] Q. Lu, X. Shen, X. Wang, M. Li, J. Li, and M. Zhang, 'Fault Diagnosis of Rolling Bearing Based on Improved VMD and KNN', *Math Probl Eng*, vol. 2021, 2021, doi: 10.1155/2021/2530315.

[13] 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.

[14] 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

[15] N. Fathiah Waziralilah, A. Abu, M. H. Lim, L. K. Quen, and A. Elfakharany, 'Bearing fault diagnosis employing Gabor and augmented architecture of convolutional neural network', *Journal of Mechanical Engineering and Sciences*, vol. 13, no. 3, pp. 5689–5702, 2019, doi: 10.15282/jmes.13.3.2019.29.0455.