Quest Journals
Journal of Software Engineering and Simulation
Volume 10 ~ Issue 3 (2024) pp: 47-53
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

A Comparative Study of Statistical Machine Learning Methods for Condition Monitoring Of Electric Drive Trains

Ms. Pooranam N
Assistant Professor - Department of Computer Science and Engineering Sri Krishna College of Engineering and Technology

Abirami Sivakumar
Department of computer science and engineering Sri Krishna college of Engineering and Technology
20eucs002@skcet.ac.in

Harish Velumani
Department of computer science and engineering Sri Krishna college of Engineering and Technology
20eucs050@skcet.ac.in

Gantharaj NN
Department of computer science and engineering Sri Krishna college of Engineering and Technology
20eucs040@skcet.ac.in

ABSTRACT
This project describes the design and implementation of a non-contact vibration picker for capturing data from spinning machinery in order to detect bearing faults early. The Hilbert transform is used to denoise the collected vibration signals, and the dataset is then analysed by Principal Component Analysis (PCA) and Sequential Floating Forward Selection (SFFS) for dimensionality reduction and feature selection, respectively. The most essential attributes are then used to identify and categorize various bearing issues using Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP) algorithms. The entire methodology provides an effective and proactive way for bearing health monitoring and maintenance, emphasizing fast defect identification and leading to significant savings in time, effort, and equipment maintenance costs.
KEY WORDS
Machine Learning, Electric Drive Trains, Condition Monitoring

***R**eceived 08 Mar., 2024; **R**evised 20 Mar., 2024; **A**ccepted 22 Mar., 2024* *© The author(s) 2024.*
P**ublished with open access at **www.questjournals.org

I. INTRODUCTION

The increasing reliance on electric drive trains in a variety of sectors highlights the vital need for robust condition monitoring to maintain optimal performance and dependability. As these systems get more sophisticated, there is a greater need for advanced statistical machine learning algorithms to improve condition monitoring accuracy and efficiency. The purpose of this comparison research is to investigate and assess several statistical machine learning algorithms in the context of electric drive train status monitoring. This study aims to give useful insights into the creation of resilient and adaptable monitoring systems by analysing the strengths and limits of various methodologies.

1.1 MACHINE LEARNING

Machine learning, a type of artificial intelligence, has transformed how we tackle complicated issues in a variety of disciplines. It is a data-driven method that allows computers to learn and forecast or make judgments without being explicitly programmed. This game-changing technology has implications ranging from healthcare and banking to driverless cars and natural language processing. Machine learning is based on algorithms and models that can analyse and extract patterns from massive information, providing answers that were previously considered to be science fiction. Machine learning has become a vital tool for making sense of the massive and complicated information accessible to us in this era of data abundance, frequently outperforming human skills in tasks such as picture identification, language translation, and even game- playing. A number of reasons have contributed to the expansion of machine learning, including the availability of large datasets, advancements in processing power, and breakthroughs in algorithm development.

1.2 ELECTRIC DRIVE TRAINS

Electric drive trains represent a paradigm shift in transportation and industrial machinery, ushering in a new era of better efficiency, lower environmental impact, and improved performance. Electric drive trains, as opposed to traditional mechanical drive systems, use electrical power to propel cars and run machines. The worldwide goal to reduce the environmental footprint of transportation and industrial activities is driving this trend toward electrification. Electric drive trains have a wide range of applications, including electric automobiles and hybrid vehicles, as well as industrial robots and renewable energy systems

IAs the world pursues more sustainable technologies, the need for dependable, high-performance electric drive trains has increased.

1.3 CONDITION MONITORING

Condition monitoring is an important aspect of contemporary engineering and maintenance procedures, since it provides a proactive method to assessing the health of machinery and systems. It entails the constant and real- time evaluation of numerous parameters to assure optimal equipment operation, forecast probable defects, and prevent unexpected failures. This strategy contrasts sharply with typical reactive maintenance, moving the emphasis from correcting problems after they occur to avoiding them before they impair performance. Condition monitoring is a strategic technique of improving dependability, reducing downtime, and optimizing operating efficiency in the context of machinery and industrial systems. Condition monitoring gives a full picture of equipment health by utilizing modern sensors, data analytics, and machine learning approaches. This enables early interventions and cost-effective maintenance plans

II. LITERATURE REVIEW

CONNOR Abbreviate et.al., [1] has proposed in this framework Profound convolutional brain networks have performed astoundingly well on numerous PC Vision assignments. To avoid overfitting, these networks, on the other hand, heavily rely on big data. Overfitting alludes to the peculiarity when an organization learns a capability with extremely high difference, for example, to show the preparation information impeccably. Sadly, big data is unavailable to many application domains, including medical image analysis. Data Augmentation, a data-space solution to the issue of limited data, is the focus of this survey. Information Increase incorporates a set-up of methods that improve the size and nature of preparing datasets to such an extent that better Profound Learning models can be constructed utilizing them.

Mateusz Budaet.al.,[2] has proposed in this system In this study, we systematically investigate the impact of class imbalance on classification performance of convolutional neural networks (CNNs) and compare frequently used methods to address the issue. Class imbalance is a common problem that has been comprehensively studied in classical machine learning, yet very limited systematic research is available in the context of deep learning. In our study, we use three benchmark datasets of increasing complexity, MNIST, CIFAR-10 and ImageNet, to investigate the effects of imbalance on classification and perform an extensive comparison of several methods to address the issue: oversampling.

M. WAQAR AKRAM et.al.[3] has proposed in this framework Imperfection recognition in photovoltaic (PV) modules and their effect evaluation means quite a bit to upgrade the PV framework execution and dependability. To recognize and break down the deformities, a superior open air infrared (IR) thermography plot is introduced in this review. PV modules that are in good working order and those that are defective are used in both the indoor (dark) and outdoor (illuminated) IR experiments. Normal operating modules have comparable measurements for the indoor and outdoor environments. Notwithstanding, the estimations for blemished modules show distinction for example the open air pictures show less or not the slightest bit abandons in contrast with indoor pictures. After this, our improved outdoor thermography scheme is used for outdoor imaging.

XIAOXIA LI and others, [4] has proposed in this system that novel inspection methods and analysis tools are required for efficient condition monitoring and precise module defect detection in large-scale photovoltaic (PV) farms. This paper presents a profound learning based answer for deformity design acknowledgment by the utilization of flying pictures got from automated elevated vehicles (UAVs). The convolutional brain organization (CNN) is utilized in the AI cycle to order different types of module deserts. Such managed educational experience can remove a scope of profound highlights of working PV modules. It essentially works on the proficiency and precision of resource review and wellbeing evaluation for huge scope

R. Pierdiccaet.al. [5] has proposed in this system that there are now a significantly larger number of distributed photovoltaic (PV) plants that produce electricity. As a result, the problem of monitoring and maintaining a PV plant has become a major concern and presents numerous difficulties in terms of efficiency, dependability, safety, and stability. This paper presents the clever way to deal with gauge the PV cells debasements with DCNNs. This is, to the best of our knowledge, the first use of data obtained with a thermal infrared-equipped drone, despite the fact that numerous studies have classified images. The experiments on the "Photovoltaic images Dataset," a dataset that was collected, are shown to demonstrate the degradation issue and provide a comprehensive evaluation of the method that is presented in this study. The proposed method's efficacy and suitability are demonstrated

III. EXISTING SYSTEM

Fault detection and diagnosis are crucial for precise industrial machinery maintenance and management. In this context, data-driven condition monitoring models play a significant role in the diagnosis and management of equipment faults. This research looks on the use of various statistical machine learning techniques in modeling vast amounts of data in the condition monitoring of electric drive trains in supply chains. To discriminate between eleven failure states, large data sets are employed to train linear discriminant analysis, K-nearest neighbor method, naive Bayes, kernel naive Bayes, decision trees, and support vector machine. The decision trees achieved 93.8% accuracy in the testing data set, followed by kernel naive Bayes (91.9%), radial basis function (Gaussian) support vector machine (89.3%), linear discriminant analysis (84.5%), k-NN algorithm (80.5%), and Gaussian naive Bayes (71.3%). As a result, the statistical machine learning technique used effects classification accuracy in electric drive problem diagnostics. Furthermore, decision trees learn and categorize new examples from huge data in a matter of seconds. This simplifies the selection of decision trees for electric drive train status monitoring and management.

IV. PROPOSED SYSTEM

The suggested system acquires data from spinning machinery using a non-contact vibration pickup, allowing for early failure diagnosis in bearings. The Hilbert transform is used to denoise vibration signals, and the dataset is then analysed using Principal Component Analysis (PCA) and Sequential Floating Forward Selection (SFFS) for dimensionality reduction and feature selection, respectively. The most essential attributes are then used to identify and categorize various bearing issues using Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP) algorithms. This complete methodology provides an effective and proactive way for bearing health monitoring and maintenance, emphasizing fast defect identification and leading to significant savings in time, effort, and equipment maintenance costs.

a. LOAD BEARING FAULT DATASET

The project focuses on obtaining and curating a comprehensive dataset especially specialized to load-bearing problems in rotating machinery in this module. This entails gathering vibration data under various load situations. The dataset serves as the foundation for further analysis, ensuring that the system is trained and tested on a wide variety of load-induced bearing defects

The thorough selection and compilation of this dataset are critical to the overall system's accuracy and dependability in recognizing and diagnosing errors connected to fluctuating loads.

b. FEATURE REDUCTION USING PCA BASED ON FEATURE EXTRACTION AND NORMALIZATION

This module addresses the need for fast feature extraction and normalization following the capture of the load- bearing fault dataset using Principal Component Analysis (PCA). PCA is used to minimize the dimensionality of a dataset while maintaining crucial information, improving computing performance and reducing the danger of overfitting. This phase is critical in preparing the data for following stages of analysis, as it ensures that the most significant properties are kept, adding to the system's capacity to properly distinguish between different fault scenarios.

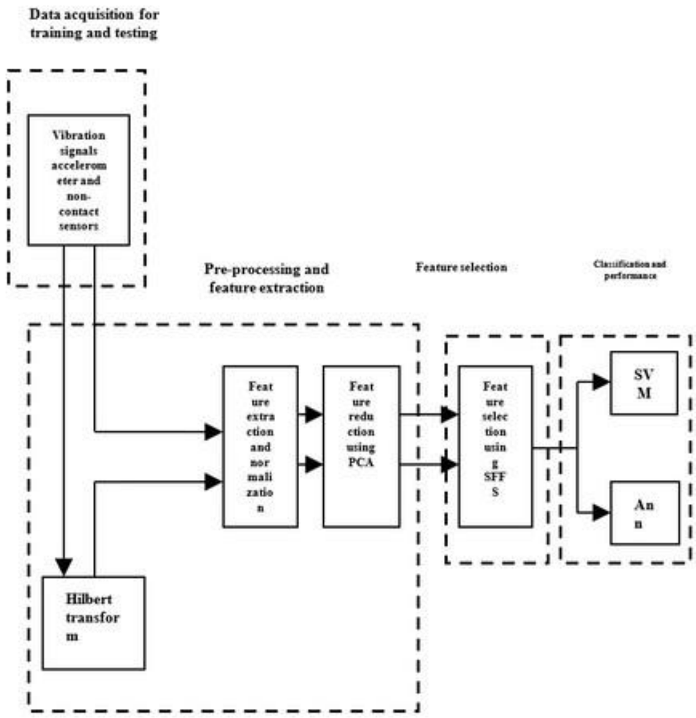

Figure: Block Diagram

c. SVM CLASSIFICATION BASED ON FEATURE SELECTION USING SFFS

The SVM classification module uses Support Vector Machines (SVM) to accurately classify faults based on the pre-processed information. Sequential Floating Forward Selection (SFFS) is used to enhance feature selection, a strategy that iteratively discovers and adds the most discriminative features to improve the model's performance.

This guarantees that the SVM classifier is trained on the most relevant data, improving its capacity to detect and categorize load-bearing problems in the rotating gear under examination.

d. MLP CLASSIFICATION BASED ON FEATURE SELECTION USING SFFS

This part, like the SVM classification module, incorporates Multi-Layer Perceptron (MLP) for fault classification. The SFFS feature selection procedure is used to discover and include the most useful characteristics, hence maximizing the MLP model's ability to distinguish between distinct load-bearing faults. The use of MLP adds a neural network-based method, providing flexibility in collecting nuanced patterns within data, thus improving the system's overall accuracy and resilience in detecting and diagnosing load-induced bearing failures.

V. ALGORITHM DEATILS

A. *Multi Layer Perceptron (MLP)*

This Model The pre-processed dataset was visualized using Matplotlib. The realization led to the insights of a classification problem upon the data pre-set with class labels (human presence or absence). These insights led to the decision to train an MLP model over other different options available. The data collected in the first phase of the experimentation was used to train an MLP over Tensor flow library in python. A sequential neural network was modelled with five hidden layers of varied neuron. The input dimensions of the data were 1429*16. Model training was divided into training, validation, and testing.

The training matrix was 1143*16; a 30% validation set was derived from the training set itself, whereas a matrix of 230*16 was utilized for the testing purpose. A 30% validation split was given to the dataset. Several network topologies with various layers and units in each layer were tested, and the bestperforming model was worked upon for further computations. A five-layer highly optimized MLP was built with the first four layers having the activation function as Relu, whereas the fifth layer uses the Sigmoid activation function to get a binary classification output

B. Support Vector Machine

Support Vector Machine (SVM) is a popular supervised machine learning model that is used for classification and prediction of unknown data. It is asserted by several researchers that SVM is a very accurate technique for text classification. It is also widely used in sentiment classification. For instance, if we have a dataset in which data is pre-labeled into two categories:

positive and negative reviews, then we can train a model to classify new data into these two categories. This is exactly how SVM works. It is the model that we train on a dataset, so it can analyze and classify unknown data into the categories that were present in the training set. SVM is a linear learning method. It finds an optimal hyper-plane to differentiate two classes. Being a supervised classification model, it tries to maximize the distance between the closest training point and either class so as to achieve better classification performance on test data. The process for classification functions is as follows:

1. It takes the labeled sample of data, and draws a line separating the two classes. This line is called the decision boundary. The solution is based only on those training data points which are really close to the decision boundary. The data points are called Support Vectors. For example, if we are categorizing movie reviews (in our case), one side of the boundary will have positive reviews while the other side has negative reviews.
2. Now when new data needs to be classified, it goes either into the left or right side of the decision boundary. Depending on which side the data enters, it is classified under that category. To classify our data with the best precision, we need to split the two categories such that the decision boundary separates the two classes with maximum space between them.

VI. RESULT ANALYSIS

Two machine learning techniques, Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP), were used to evaluate the suggested bearing failure detection system. The SVM method was 84% accurate overall, with precision, recall, and F-measure values of 0.79, 0.75, and 0.84, respectively. The MLP algorithm, on the other hand, outperformed the others, obtaining an accuracy of 92%. Notably, MLP demonstrated good accuracy (0.99) and recall (0.92) metrics, but with a somewhat lower F-measure of 0.89. These findings support the suggested methodology's efficacy, with the MLP algorithm beating SVM in terms of accuracy and precision, making it a potential choice for bearing defect recognition and classification in spinning equipment.

Algorithm	Accuracy	Precision	Recall	F-measure
SVM	0.84	0.79	0.75	0.84
MLP	0.92	0.99	0.92	0.89

Figure: Comparison Table

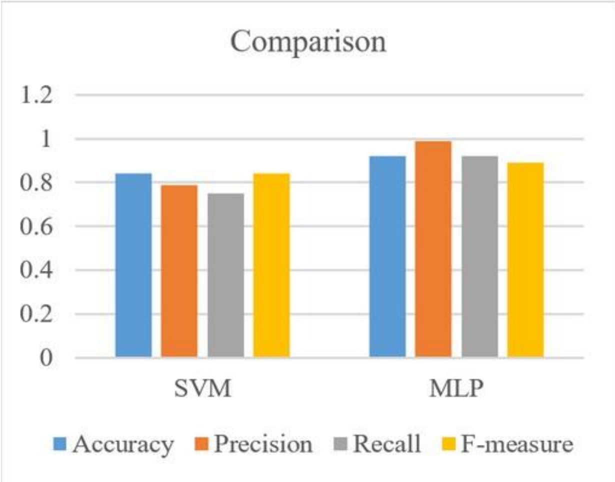

Figure: Comparison Graph

VII. ADVANTAGE

The proposed system can detect bearing faults early, before they cause catastrophic failures. This can prevent downtime, production losses, and costly repairs. By enabling early fault detection, the proposed system can help to reduce maintenance costs. This is because it can prevent unnecessary maintenance interventions and extend the lifespan of bearings. The proposed system can improve the reliability of equipment by preventing unexpected breakdowns. This can lead to increased productivity and reduced downtime. The proposed system uses a non- contact vibration pickup to collect data, which eliminates the need for direct contact with the machinery. This can reduce the risk of damage to the machinery and make it possible to monitor the machinery without disrupting operations.

VIII. FUTURE WORK

It is critical for future work to investigate and enhance the suggested system to fit a greater range of industrial environments and machinery types.

Further research into the non-contact vibration pickup's flexibility across diverse operational circumstances and settings will increase the system's versatility. Incorporating real-time monitoring capabilities and investigating the incorporation of future technologies like as edge computing or the Internet of Things might also give a more dynamic and responsive approach to bearing health monitoring. Continuous research efforts should be directed toward optimizing machine learning algorithms, with the possibility of adding deep learning models for enhanced pattern identification and fault detection.

IX. CONCLUSION

Finally, the created non-contact vibration pickup, along with modern data processing and machine learning approaches, has shown to be a reliable and practical method for monitoring bearing health in rotating machinery. The Hilbert transform was used for denoising, PCA was used for dimensionality reduction, and SFFS was used for feature selection, which simplified the dataset and allowed for the exact detection and categorization of various bearing faults. The effective deployment of SVM and MLP algorithms illustrates the system's capacity to detect faults in real time. This complete technique not only improves the proactive nature of maintenance procedures, but it also offers significant savings in time, resources, and equipment upkeep expenses.

REFERENCE

[1]. Q. M. Rahman, N. Sünderhauf, and F. Dayoub, ''Online monitoring of object detection performance post- deployment,'' 2020, arXiv:2011.07750. [Online]. Available: http://arxiv.org/abs/2011.07750

[2]. Q. M. Rahman, N. Sünderhauf, and F. Dayoub, ''Per- frame map prediction for continuous performance monitoring of object detection during deployment,'' in Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV) Workshops, Jan. 2021, pp. 152–160

[3]. S.S. Moosavi, A. Djerdir, Y. Ait-Amirat, D.A. Khaburi, N'Diaye, Artificial neural network-based diagnosis in the AC-DC converter of the power supply of series hy brid electric vehicle. IET Electrical Systems in, Transportation 6 (2016) 96–106.

[4]. X. Ji, X. He, C. Lv, Y. Liu, J. Wu, Adaptive-neural- network-based robust lateral motion control for autonomous vehicle at driving limits, Control Eng. Pract. 76 (2018) 41–53.

[5]. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, Liu, and M. Pietikäinen, ''Deep learning for generic object detection: A survey,'' Int. J. Comput. Vis., vol. 128, no. 2, pp. 261–318, Feb. 2020.

[6]. X. Huang, D. Kroening, W. Ruan, J. Sharp, Y. Sun, E. Thamo, M. Wu, and X. Yi, ''A survey of safety and trustworthiness

of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability,'' Comput. Sci. Rev., vol. 37, Aug. 2020, Art. no. 100270.

[7]. Z. Shao, J. Yang, and S. Ren, ''Increasing trustworthiness of deep neural networks via accuracy monitoring,'' in Proc. Workshop Artif. Intell. Saf., 2020, pp. 1–8

[8]. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, Liu, and M. Pietikäinen, ''Deep learning for generic object detection: A survey,'' Int. J. Comput. Vis., vol. 128, no. 2, pp. 261–318, Feb. 2020.

[9]. X. Huang, D. Kroening, W. Ruan, J. Sharp, Y. Sun, E. Thamo, M. Wu, and X. Yi, ''A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability,'' Comput. Sci. Rev., vol. 37, Aug. 2020, Art. no. 100270.

[10]. Z. Shao, J. Yang, and S. Ren, ''Increasing trustworthiness of deep neural networks via accuracy monitoring,'' in Proc. Workshop Artif. Intell. Saf., 2020, pp. 1–8

[11]. Q.P. He, J. Wang, Fault detection using the K- nearest neighbor rule for semi conductor manufacturing processes, IEEE Trans. Semicond. Manuf. 20 (2007) 345–354.

[12]. L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and regression trees. Belmont, CA, Wadsworth, 1984.

[13]. H. Helmi, A. Forouzantabar, Rolling bearing fault detection of electric motor using time domain and frequency domain features extraction and ANFIS, IET Electr. Power Appl. 13 (2019) 662–669.

[14]. B. Ravikumar, D. Thukaram, H.P. Khincha, Application of support vector machines for fault diagnosis in power transmission system, IET Gener., Transm. Distrib. 2 (2008) 119–130.

[15]. S. Lahmiri, S. Bekiros, A. Giakoumelou, F. Bezzina, Performance assessment of ensemble learning systems in financial data classification. Intelligent systems in accounting, Financ. Manag. 27 (2020) 3–9.

[16]. A. Holzinger, E. Weippl, A.M. Tjoa, P. Kieseberg, Digital Transformation for Sustainable Development Goals (SDGs)- A Security, Safety and Privacy Perspective on AI.'', in: A. Holzinger, P. Kieseberg, A.M. Tjoa, E. Weippl (Eds.), Machine Learning and Knowledge Extraction. Lecture Notes in Computer Science, vol 12844, Springer, Cham, 2021.

[17]. P. Antonante, D. I. Spivak, and L. Carlone, "Monitoring and diagnosability of perceptual systems," 2020; arXiv:2005.11816. [Online], http://arxiv.org/abs/2005.11816.

[18]. W. Zhou, J. S. Berrio, S. Worrall, and E. Nebot, "Automated evaluation of semantic segmentation robustness for autonomous driving," IEEE Transactions on Intelligence and Transactional Systems, vol. 21, no. 5, pp. 1951-1963, May 2020.

[19]. A. Gupta and L. Carlone, "Online monitoring for neural network based monocular pedestrian stance estimation 2020; arXiv:2005.05451. [Online]: http://arxiv.org/abs/2005.05451.