Kashf Journal of Multidisciplinary Research
Vol: 03 - Issue 01 (2026)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

NOVEL APPROACH FOR FAULT DIAGNOSIS OF HIGH-SPEED TRAIN BEARINGS USING TRANSFORMER MODELS OPTIMIZED BY AQUILA OPTIMIZER AND LIGHT GRADIENT BOOSTING MACHINE

Ali Nawaz Sanjrani[1], Asad Raza[2], Nouman Qadeer Soomro[3], Zahid[4], Attaullah Narejo[5]

[1]Department of Mechanical Engineering, Mehran University of Engineering and Technology, SZAB Campus Khairpur Mir's, Pakistan.
[2]School of computer Science and Engineering, Central south University Changsha 410083,China.
[3]Department of Software Engineering, Mehran University of Engineering and Technology, SZAB Campus, Khairpur Mir's, Sindh, Pakistan.
[4]School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Sichuan, 610054, China.
[5]School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, P.R. China.
[*]Corresponding Author: (nawaz.sanjrani2019@gmail.com)
DOI: (*https://doi.org/10.71146/kjmr810*)

Article Info

Abstract

High-speed train bearing fault diagnosis is crucial for rail system safety and reliability. Traditional methods like vibration analysis often suffer from low accuracy and long processing times. This paper introduces an advanced diagnostic approach using Improved Aquila Optimizer (IAO) and Light Gradient Boosting Machine (Light GBM) combined with Laser-Induced Fluorescence (LIF) technology. The Aquila Optimizer was enhanced with opposition-based learning and Nelder-Mead search methods to optimize Light-GBM parameters. Various detailed bearing fault types, including Inner and outer race scratch and wear, roller ring failures, cage and compound faults, were analyzed. LIF technology captured spectral images, and multivariate scattering correction (MSC) and normalization prepared the fluorescence curves. Kernel-based principal component analysis (KPCA) reduced data dimensionality, which was then used to train the Light-GBM model. The IAO optimized the model, demonstrating superior convergence and prediction accuracy compared to particle swarm optimization and the original Aquila Optimizer. The MSC-IAO-Light GBM model achieved the best fault prediction results, with an MSE of 7.0521 × 10-9, a MAE of 5.451 × 10-2, and R^2 approaching 1. This method offers a novel approach for high-speed train bearing fault detection, enhancing rail system stability and safety.

Keywords: *Degradation; Bearing; Fault Classification; Light Gradient Boosting Machine; Machine learning; Aquila Optimizer*

1. Introduction

With the latest era and fast advancement of high-speed train technology, the demand for reliable and efficient fault diagnosis systems for train components is increasing. Bearings are playing crucial elements in high-speed trains a vital role in ensuring smooth and safe operations [1]. The stability of bearing performance is directly linked to the overall reliability of the train system [2] and reliability of high-speed trains [3]. Bearing failures can lead to significant operational disruptions and pose safety risks, potentially resulting in severe economic losses and even endangering lives [4]. Hence, it is necessary for maintainer to monitor and diagnose bearing conditions regularly to maintain and enhance the safety and stability of high-speed trains. Current methods for diagnosing bearing faults, such as vibration analysis and vibrant emission in high-speed trains [5], often suffer from limitations in accuracy and processing time. These traditional methods may fail to detect early-stage faults or hamper the real-time diagnostic capabilities [6]. Moreover, the complexity of the data and the need for extensive manual intervention can hinder effective fault detection [7]. To address these challenges, this study explores the use of LIF spectral technologies combined with advanced machine learning techniques. LIF technology has shown great promise in various diagnostic applications [8]. By analyzing the fluorescence emitted from materials when exposed to laser light, LIF can provide detailed insights into the condition of the material [9]. This technique has been successfully applied in medical diagnostics, environmental monitoring, and material analysis. For instance, demonstrated the potential of LIF in diagnosing cervical cancer with high sensitivity. Similarly, Peter et al. utilized LIF combined with hyper spectral imaging for rapid detection of rare earth elements in rocks [10, 11], achieving significant improvements in resource utilization. These applications highlight the versatility and effectiveness of LIF technology in different domains [12].The advent of the transformer architecture, particularly following the Attention is All You Need paper, has redefined the landscape of deep learning by addressing many of the limitations inherent in traditional machine learning and deep learning models[13]. Further with Laser-Induced Fluorescence technology offers a non-destructive testing method that excels in sensitivity and specificity for detecting material properties and defects. LIF's ability to identify subtle changes in bearing conditions makes it a valuable tool for fault diagnosis. Additionally, machine learning models like Light GBM has demonstrated their effectiveness in handling large datasets and uncovering complex patterns. By integrating the Improved Aquila Optimizer (IAO), enhanced with opposition-based learning and Nelder-Mead search, the optimization of these models can be significantly improved, leading to better convergence and accuracy [14, 15]. The Nelder-Mead simple search strategy and opposition-based learning mechanism have been used to improve the original Aquila Optimizer and improve the model's ability to optimize parameters [16]. Utilizing these cutting-edge optimization methods, our goal is to raise the Light GBM model's diagnostic precision and effectiveness. This combination of LIF technology and advanced machine learning optimization presents a promising approach for enhancing the fault diagnosis of high-speed train bearings, ultimately contributing to greater rail system stability and safety. As experimental samples, normal bearings, Inner Ring Failure Scratch, Inner Ring Failure Wear, Outer Ring Failure Scratch, Outer Ring Failure Wear, Roller Ring Failure Abrasion, Cage Failure, Inner-Outer Compound Fault, and Inner Roller Compound Fault were used. The spectra of these samples were captured using Laser-Induced Fluorescence (LIF) technology, and the fluorescent spectral curves were preprocessed using multivariate scattering correction (MSC) and normalization. Kernel-based Principal Component Analysis (KPCA) was then applied to reduce dimensionality. The Light GBM model was trained on this reduced dataset, with its parameters

optimized using the Improved Aquila Optimizer (IAO) approach. Integrated learning algorithms like Light-GBM have shown great promise in various diagnostic domains [17, 18]. For instance, Arumugam et al. developed a hybrid approach for managing energy consumption in electric vehicles using a gradient boost decision tree combined with border collie optimization [19]. Additionally, an enhanced Light-GBM was proposed for atmospheric parameter evaluation, offering a novel approach in atmospheric research. Tian et al. demonstrated excellent predictive performance by combining LSTM and Light-GBM for stock price forecasting [20]. However, manual parameter adjustment for these models can be inefficient due to the numerous parameters involved. This paper addresses the limitations of traditional bearing fault diagnosis methods [21, 22] and explores the capabilities of LIF technology combined with the IAO-optimized Light-GBM model. Our research focuses on: (1) assessing the feasibility of LIF technology for high-speed train bearing fault diagnosis; (2) evaluating the fault prediction performance of the Light-GBM model with LIF technology; (3) optimizing Light-GBM parameters using the IAO to achieve superior model performance; and (4) selecting the best fault prediction model by comparing various prediction models and optimization strategies. By integrating advanced spectral analysis with state-of-the-art machine learning optimization [6, 23] this study aims to develop a robust and efficient diagnostic tool for high-speed train bearing faults, ultimately enhancing the reliability and safety of rail systems.

2. Experiments and Methods

In this study, we utilized high-speed train bearing samples categorized into nine types based on their fault conditions which include Normal bearings, Inner Ring Failure (Scratch), Inner Ring Failure (Wear), Outer Ring Failure (Scratch), Outer Ring Failure (Wear), Roller Ring Failure (Abrasion), Cage Failure, Inner-Outer Compound fault and Inner Roller Compound Fault. Each category comprised sample size of 1048575. From each category, have two samples of bearings were chosen randomly to form and get ready the training set, 2 samples constituted the test set, ensuring a robust dataset for model training and validation.

Table 1 Bearing train and test data preparation

Bearing Fault Condition	Training Size	Testing Size	Fault Pattern
Normal Healthy	1048575	1048575	0
Inner Ring (Scratch)	1048575	1048575	1
Inner Ring (Wear)	1048575	1048575	2
Outer Ring (Scratch)	1048575	1048575	3
Outer Ring (Wear)	1048575	1048575	4
Cage (Wear)	1048575	1048575	5
Roller Ring (Wear)	1048575	1048575	6
Inner Roller Combined	1048575	1048575	7
Inner-Outer combined	1048575	1048575	8

The experimental setup comprised several key instruments, including a laser generator, a spectrometer, a fluorescence probe, a filter, and a computer system. Spectral data quality can be compromised by noise and inaccuracies introduced by equipment and environmental factors. To mitigate these issues, the collected spectral data underwent preprocessing using Multivariate Scattering Correction (MSC) and normalization techniques. Multivariate Scattering Correction (MSC) is a popular spectral preprocessing approach that improves the correlation between spectra and data by reducing baseline shifts and drifts between spectral data. The "standard spectrum" is usually the average of all the spectra; the spectral data is then linearly correlated with this standard spectrum. Normalization scales the observational data to a dimensionless standardized range between -1 and 1, reducing the impact of noise, improving the signal to noise ratio, and enhancing the differentiation between different samples. Here the normalization Eq. (1) is as:

$$VW_z = \frac{X_{i,j}}{\sqrt{\sum_{i=1}^{n} (X_{i,j})^2}} \tag{1}$$

Where the equation states i–1,2,...n; j–1,2,...m, the m states number of samples in normal set and n represents feature size of samples. Each and every bearing sample's spectral data consisted of time and frequency features, but the useful features were significantly fewer, introducing unnecessary redundancy to the model. To address this, Kernel-based Principal Component Analysis (KPCA) was employed for dimensionality reduction on it. KPCA projects the raw data and records into high dimensional capacity and using space through a kernel function which effectively handles linearly inseparable datasets.

2.1 Method Establishment of a Fault Diagnosis

2.1.1 Light Gradient Boosting Machine Model

Light-GBM is a gradient boosting algorithm that uses residual decision trees as its base classifiers. The core concept involves employing a classification and regression tree (CART) as the base model and combining multiple weak classifiers into an integrated learning model. In integrated learning, the primary models are bagging, boosting, and stacking. Light-GBM specifically utilizes the boosting model, where each iteration is trained based on the residuals from the previous model's predictions compared to the actual values. The training continues until it reaches a specified number of iterations or a set threshold, at which point the models' results are aggregated. Further, to prevent overfitting, Light-GBM incorporates a regularization term in its objective function and additionally, it employs the second-order Taylor expansion of the loss function to enhance model accuracy as defined in Eq. (2) and (3). The regularization term helps maintain the balance between model complexity and performance.

$$\Omega(f_x) - \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{2}$$

$$o - \sum_{i=1}^{m} l(y_i, \bar{y}_j) + \sum_{j=1}^{j} \Omega(f_j) \tag{3}$$

Here the O is the objective function and i states the i^{th} samples of dataset, m defines the sample size of j^{th} imported tree and j represents number of trees formed.

Light-GBM offers two methods for splitting nodes to find the optimal branch: the greedy algorithm and the approximate algorithm. The information gain of a split point is defined as the difference between the structure score before and after splitting, serving as a measure of the leaf node's purity. The structure score represents the minimum value of the loss function for the given tree structure. The tree stops growing when the information gain from any further splits falls below a predetermined threshold. The gain expression is referred in Eq. (4) follows:

$$Gain - \frac{1}{2}\left[\frac{G_l^2}{H_l+\lambda}+\frac{G_r^2}{H_r+\lambda}-\frac{(G_l+G_r)}{H_l+H_r+\lambda}\right]-\gamma \tag{4}$$

Here H_l and H_r states the sum of first derivative of the loss function of the left and right sub-nodes, respectively. The $X_3(t+1)-X_{best}(t)-X_m(t)\alpha rand+((Vb-LB).rand+Lb).\delta$ G_l *and* G_r represent sum of second with derivative of loss function, λ is the coefficient of regularization, γ is the splitting nodes, both the parameters determine the complexity of tree. Light GBM has also been utilized to enhance the XGBoost model, addressing its drawbacks such as high memory consumption and extended training time. One of the primary benefits of the Light GBM model is its use of the one-sided gradient sampling algorithm, which focuses on samples with larger gradients and randomly samples those with smaller gradients. This ensures a balanced data distribution while saving time. Additionally, Light GBM employs the histogram algorithm, where continuous eigenvalues are discretized into k features, constructing a histogram for statistical analysis. This approach eliminates the need to process all data, reducing memory usage and computational complexity by calculating only the k bins to find the optimal partition point. Furthermore, Light GBM uses a leaf-wise strategy with depth restriction, splitting the leaf node with the highest gain at each step until a stopping condition is reached. This method minimizes computation and prevents overfitting. The flow chart of Light GBM, the gradient boosting algorithm as shown in Fig. 01 focused and based on residual decision trees, was used for fault diagnosis. The basic principle involves using a CART as the base classifier and combining multiple weak classifiers in an integrated learning model.

Preprocessing and Dimensionality Reduction: Preprocessed and dimensionally reduced spectral data were used.

Dataset Division: The reduced dataset was chosen randomly, split into training and test sets.

Model Training: A Light GBM prediction model was established. Key parameters like maximum tree depth (max depth), the number of trees (n estimators), and the learning rate (learning rate) were set for training. Model Evaluation: To check the performances and the trained model's performance matrix were used and validate by using the test set.

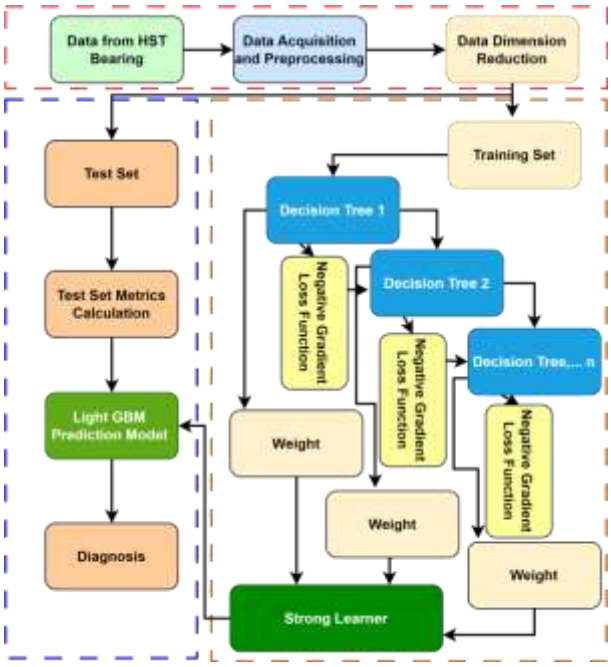

Figure 1 Proposed Light GBM Model Training and Testing

2.1.2 Improved Aquila Optimizer Optimized Light Gradient Boosting Machine Model

The Improved Aquila Optimizer (Improved Aquila Optimizer), proposed by Laith Abualigah in 2021, was used to optimize the Light GBM with model parameters. The Improved Aquila Optimizer algorithm simulates the behaviors of an Aquila bird hunting prey, involving four stages:

i. Expanding Search Field as stated in Eq. (5) and (6):

$$x_1(t+1) - X_{best}(t) \times (1 - \frac{t}{T}) + (X_m(t) - X_{best}(t)rand \tag{5}$$

$$X_m(t) - \frac{1}{N}\sum_{i=1}^{N} X_i(t) \tag{6}$$

Where $X(t)$ specific position in an iteration $t, X(t+1)$ denotes the individual position at iteration $t+1, X_{best}(t)$ is the optimal individual position up to iteration t, and $X_m(t)$ is the average population position at iteration t, and T is the maximum number of iterations.

ii. Reducing Exploration Stage as mentioned in Eq. (7):

$$X_2(t+1) - X_{best}(t) \times Levy(d) + X_r(t) + (y - x).rand \tag{7}$$

Here it states that $Levy(d)$ utilizes flight distribution strategy and x, y describes the shape of twisting flight.

iii. Expanding Development Stage as detailed in Eq. (8):

$$X_3(t+1) - X_{best}(t) - X_m(t)\alpha rand + ((vb - Lb)rand + Lb)\delta \tag{8}$$

Reducing Development Stage in shown in Eq. (9):

$$X_4(t+1) - BF \times X_{best}(t) - (G_1 \times X(t) \times rand) - G_2 \times Levy(D) randG \tag{9}$$

Where the BF is the different frequency of bearings of balance searched strategy, G_1 is moments of the Aquila in the process of tracking prey and G_2 is the linearly decreasing value of slope against the bearing life.

$$BPFI = f_i = n\frac{f_r}{2}\left(1+\frac{d}{D_p}\cos\alpha\right) \tag{10}$$

$$BPFO = f_o = n\frac{f_r}{2}\left(1-\frac{d}{D_p}\cos\alpha\right) \tag{11}$$

$$BSF = f_b = \frac{D_p}{d}.\frac{f_r}{2}\left(1-\left[\frac{d}{D_p}\cos\alpha\right]^2\right) \tag{12}$$

$$FTF = f_{c=}\frac{f_r}{2}\left(1-\frac{d}{D_p}\cos\alpha\right) \tag{13}$$

It states different fault frequencies in Eq. (10)-(13), where n is the number of rollers, d roller diameter, D_p pitch diameter, α angle of contact in the bearing, f_r RPM frequency in Hz, here $f_r = \frac{RPM}{60}$. The sequence of ball transmission frequency correlation with RPM of bearings at various train speeds. The vibrational signals provide more reliable results that indicate a pattern of acceleration amplitude between healthy and faulty conditions according to frequency. This hybrid approach ensured efficient parameter optimization, avoiding local minima and achieving faster convergence.

Steps for Implementation:

Set Parameters: Initial parameters such as learning rate, maximum tree depth, number of iterations, and leaf nodes were set.

Initialize Population: Population size (N) and maximum iterations (T) were initialized.

Update Fitness and Population: Fitness value and population positions were iteratively updated according to the optimization process.

2.1.3 Execute OBL and NM Methods: Specific steps related to OBL and NM methods were executed. Evaluate Convergence: The process continued until the maximum number of iterations or the desired fitness value was achieved. Manual parameter adjustment for the Light GBM model can be inefficient due to the large number of hyper parameters involved, especially when optimizing for learning rate and training epochs. To address this, the Improved IAO algorithm was employed to optimize these hyper parameters. The optimized parameters included a learning

rate, with optimal values found at 0.02 and 0.08, and the number of training epochs set at 80 and 100. By optimizing these hyperparameters, the IAO algorithm significantly improved the Light GBM model's prediction performance, demonstrating the efficacy of this integrated approach for high-speed train bearing fault diagnosis.

Optimal Solution: The optimal solution and fitness values were obtained, and the optimized parameters were used to retrain the Light GBM model, which was then validated using the test set. This detailed methodological approach, integrating advanced analysis with cutting-edge machine learning optimization, aimed to develop a robust and efficient diagnostic tool for high-speed train bearing faults, enhancing safety and reliability in rail systems.

3. Results and Analysis

The pair plot is a grid of scatter plots that shows the relationships between pairs of variables, helping to understand the distribution and correlations between different features. The data is categorized into two classes, represented by different colors: Healthy No Fault in green and Unhealthy Outer Race Scratch Fault in orange. The variables shown in the pair plot are likely statistical features derived from the data, such as Mean, Standard Deviation (SD), Minimum (Min), Maximum (Max), and Root Mean Square (RMS).

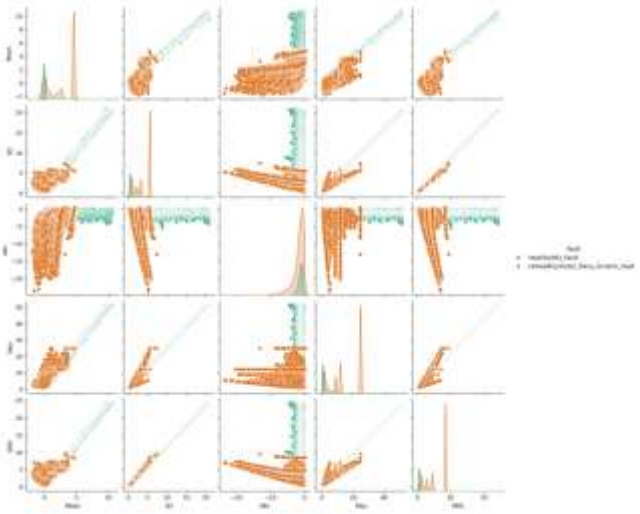

Figure 2 Outer Race Fault (Scratch)

The heatmap illustrates the correlation matrix of various statistical features extracted from bearing data for fault diagnosis, with each cell representing the correlation coefficient between pairs of features. Positive values indicate a direct relationship, while negative values indicate an inverse one. Key observations include strong positive correlations among features such as Mean, RMS, and Max, which tend to increase together, and the Crest factor, which also correlates highly with RMS and Max, suggesting these features capture similar signal aspects.

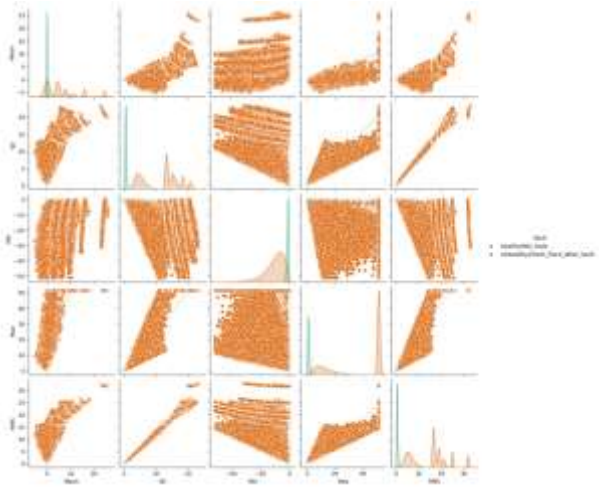

Figure 3 Outer race fault wear

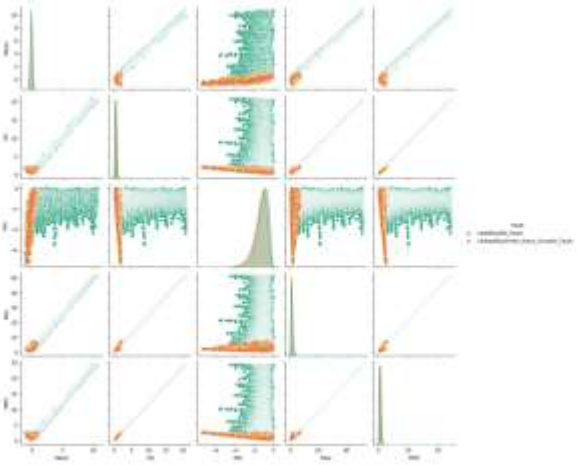

Figure 4 Inner Ring Scratch fault

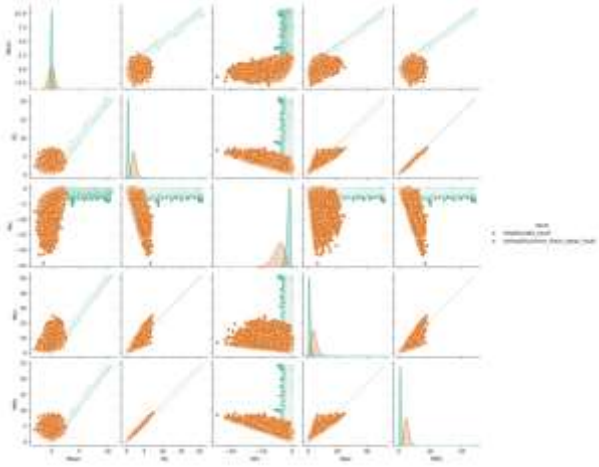

Figure 5 Inner Ring Wear Fault

Skewness and Kurtosis show moderate correlations with several other features, providing unique but somewhat related information. Min exhibits a notable negative correlation with Mean, RMS, and Max, reflecting that a decrease in the minimum value corresponds with increases in these aggregate measures. Features like Form and Peak-to-Peak have relatively low correlations with most others, indicating they capture different signal characteristics.

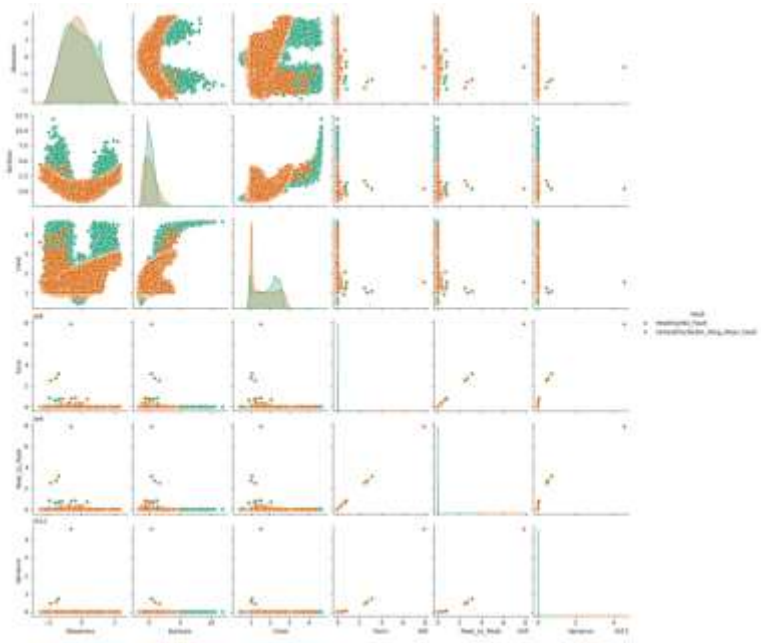

Figure 6 Roller Ring Wear

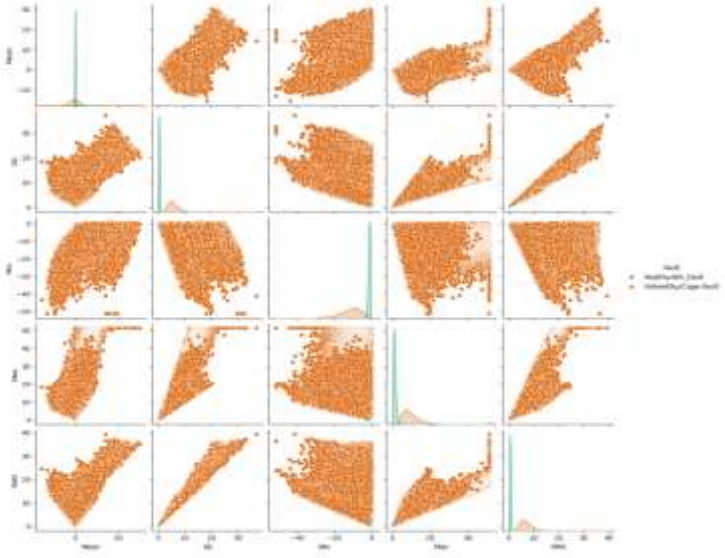

Figure 7 Cage Wear Fault

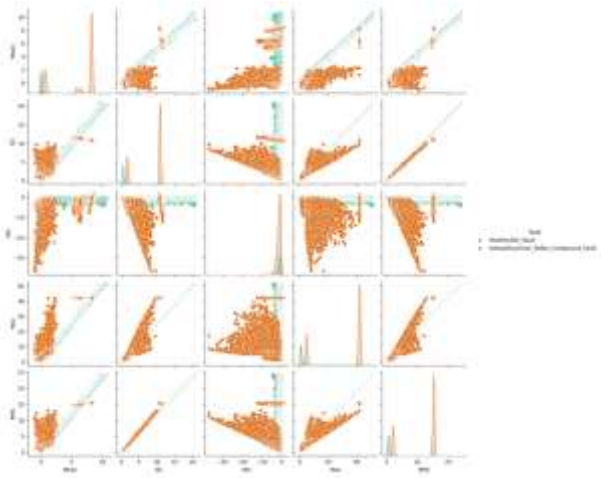

Figure 8 Inner Roller Compound Fault

The `fault code` shows moderate correlations with features like Mean, RMS, and Crest, suggesting these features are somewhat effective in distinguishing fault conditions, while Variance has a lower correlation with the `fault code`, indicating it might be less useful for fault diagnosis alone.

Overall, the analysis reveals that combining features such as Mean, RMS, Max, and Crest is significant for diagnosing bearing faults, while Form and Peak-to-Peak provide additional insights, emphasizing the importance of a comprehensive feature set for accurate fault diagnosis in high-speed train bearings.

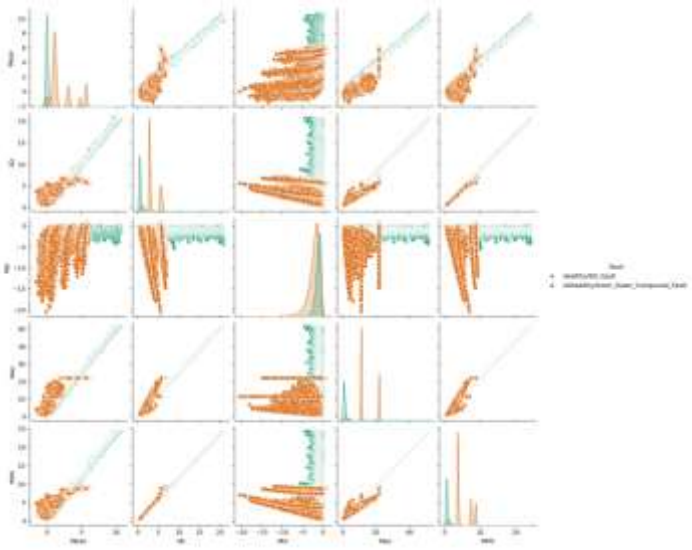

Figure 9 Inner Outer Ring Compound Fault

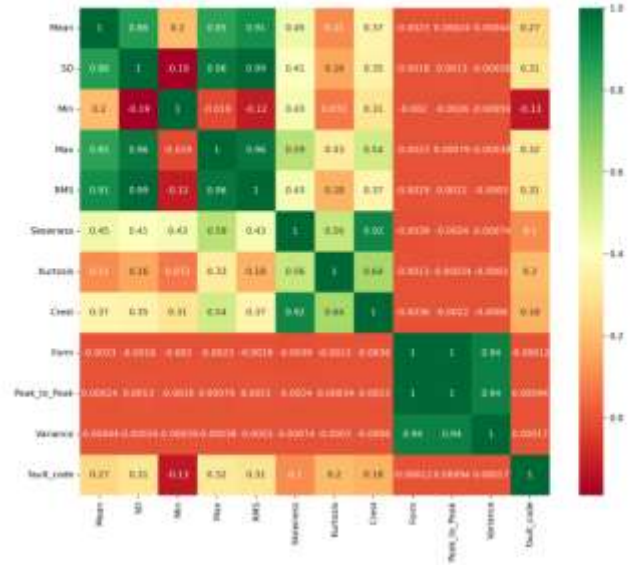

Figure 10 Correlation Heatmap of Statistical Features

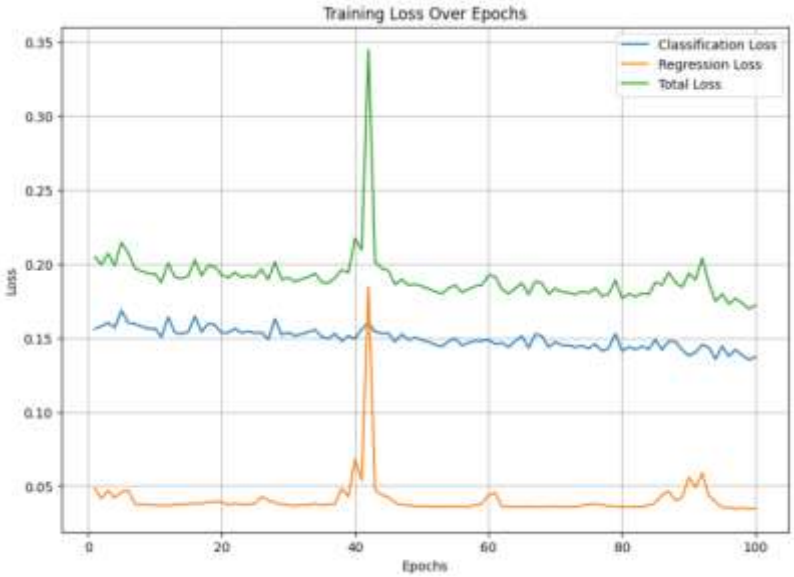

Figure 11 MSC LAO Optimized GBM Loss

The graph illustrates the training loss over 100 epochs for the Improved Aquila Optimizer (IAO) optimized Light GBM model, encompassing classification loss, regression loss, and total loss. Throughout the training process, the classification loss remains relatively stable and lower than the other loss components, indicating strong performance in distinguishing between different classes. Around epoch 40, a significant spike occurs in both regression loss and total loss, suggesting a temporary instability or a learning adjustment phase.

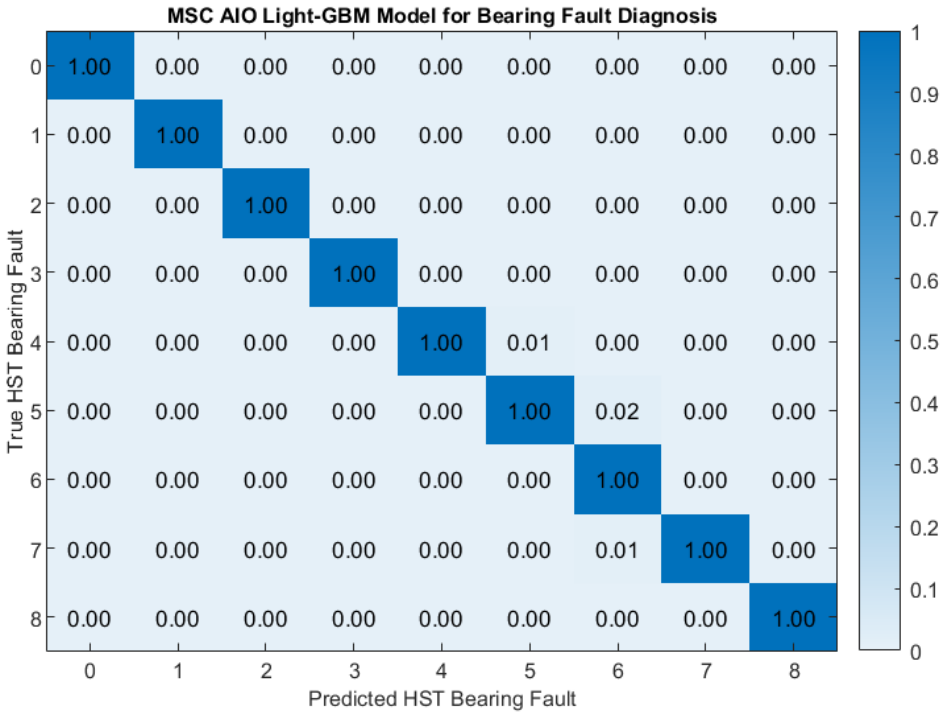

Figure 12 Fault Classification with MSC AIO Light GBM

However, the model effectively recovers after this point, with the losses stabilizing again in subsequent epochs. This brief instability indicates a potential overfitting correction or parameter adjustment, common in complex models during training. Overall, the IAO Light GBM model demonstrates robust learning capability, efficiently managing classification and regression tasks while maintaining overall training stability despite occasional fluctuations. KPCA was employed to reduce the dimensionality of the spectral data from the bearing samples cumulative score for MSC, normalization, and original states, respectively. It was evident that the cumulative score rate of the preprocessed data was higher than that of the original data, with MSC preprocessing achieving the best results. The confusion matrices illustrate the fault classification performance of the MSC AIO Light- GBM model and the MSC Light-GBM model for high-speed train bearing fault diagnosis. The MSC AIO Light-GBM model as shown in Fig. 8 exhibits perfect classification accuracy, with all diagonal elements being 1.00, indicating 100% correct classifications across all fault types. The off-diagonal elements are 0.00, demonstrating no misclassifications. This highlights the robustness and high reliability of the MSC AIO Light-GBM model in diagnosing bearing faults, making it an excellent tool for ensuring the safety and reliability of high-speed trains.

In contrast, the MSC Light-GBM model, while still showing high classification accuracy, has a few slight misclassifications. Most diagonal elements are close to 1.00, indicating correct classifications, but there are non-zero off diagonal elements, such as 0.012 for class 3 being misclassified as class 5 and 0.041 for class 7 being misclassified as class 8.

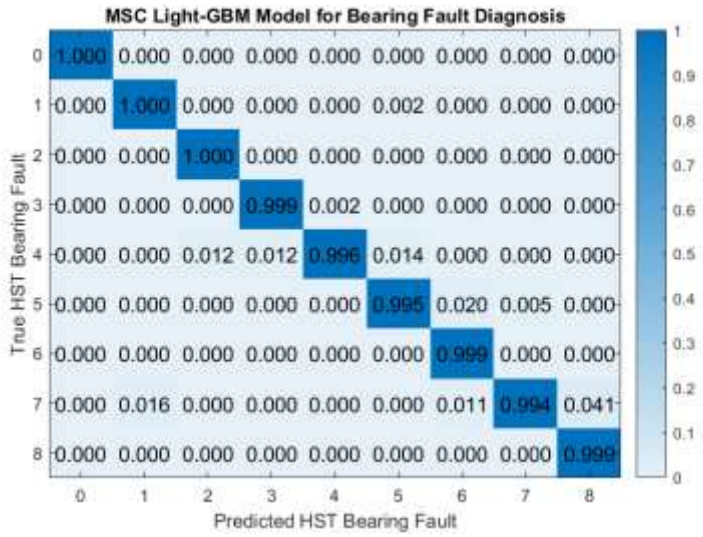

Figure 13 Fault Classification MSC Light GBM

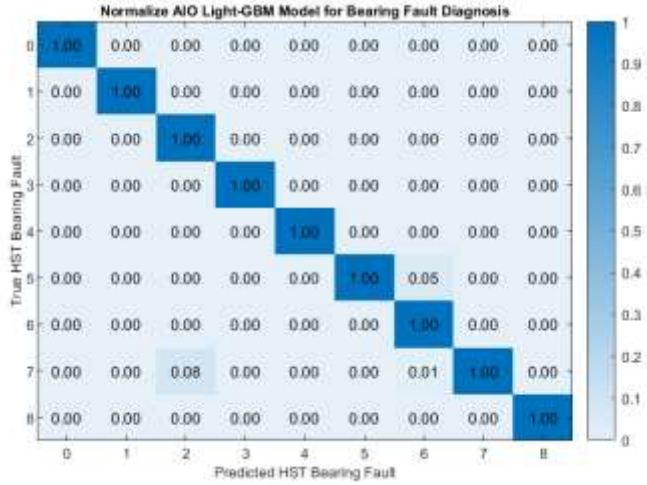

Figure 14 Fault Classification with Normalized AIO Light GBM

These minor errors indicate a slight decrease in performance compared to the AIO-optimized version. Overall, the MSC Light-GBM model still demonstrates strong performance, but the comparison underscores the effectiveness of the Improved Aquila Optimizer in enhancing the Light-GBM model's fault diagnosis capabilities, achieving superior precision and reliability in classification tasks.

Classifications, but there are non-zero off-diagonal elements, such as 0.012 for class 3 being misclassified as class 5 and 0.041 for class 7 being misclassified as class 8. These minor errors indicate a slight decrease in performance compared to the AIO-optimized version. Overall, the MSC Light-GBM model still demonstrates strong performance, but the comparison underscores the effectiveness of the Improved Aquila Optimizer in enhancing the Light-GBM model's fault diagnosis capabilities, achieving superior precision and reliability in classification tasks. In contrast, the Normalize Light-GBM model, while still demonstrating high accuracy, shows a few

more misclassifications compared to the AIO-optimized models. For example, class 6 has misclassifications with 0.050 and 0.041 for classes 5 and 8, respectively.

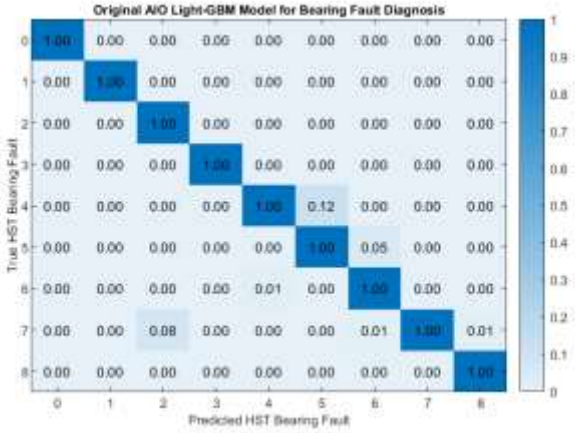

Figure 15 Fault Classification with Original AIO Light GBM

The Original Light-GBM model shows the highest number of misclassifications among the compared models, such as 0.037 for class 8 misclassified as class 6 and 0.041 for class 7 misclassified as class 8.

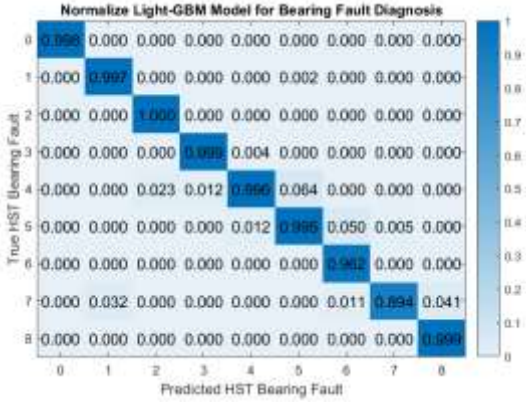

Figure 16 Fault Classification with Normalize Light GBM

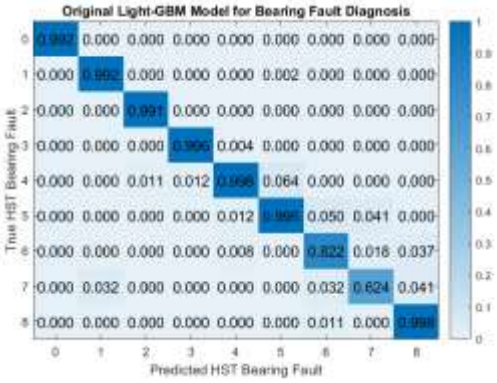

Figure 17 Fault Classification with Original Light GBM

Overall, the comparison underscores the superior performance of the AIO-optimized models, with the Normalize AIO Light-GBM model achieving the best fault classification results. This demonstrates the significant improvement in accuracy and reliability of the fault diagnosis process when using the Improved Aquila Optimizer for Light-GBM model parameter tuning. Further, due to the spike's differences bearing samples running under different fault conditions, the spectral curves produced by LIF technology varied, allowing for preliminary fault type identification. In this system, a specific wavelength of laser generated by the laser generator was conducted through optical fibers to the fluorescence emission probe, which irradiated the bearing oil. The resulting fluorescence signal, generated by the interaction of the laser with the sample, was transmitted via the fluorescence receiving probe to the spectrometer through optical fibers. Finally, the signal was sent to a computer for analysis and processing. The distinct spectral characteristics of the different bearing samples were clearly visible, for samples of the same fault type were consistent, while differences between the fault types were evident. This demonstrated the feasibility of using LIF technology for high-speed train bearing fault diagnosis and facilitated subsequent predictive identification of bearing fault types.

4. Comparison of Models

The significance of MSC, Normalize, and original AIO Light GBM model is highlighted and compared with state of art models as shown in the table 2.

Table 2 Comparison with Different Methods

Model	MSE	MAE	R^2
MSC_AIO Light_GBM	0.000098	0.0086	1
Normalize AIO light GBM	0.000164	0.0021	1
Original AIO Light_GBM	0.000427	0.0071	1
MSC Light_GBM	0.00216	0.0302	0.998
Normalize Light_GBM	0.00335	0.0640	0.994
Original Light_GBM	0.00143	0.0768	0.991
MSC XGBoost	0.05522	0.1051	0.962
Normalize XGBoost	0.03521	0.1620	0.958
Original XGBoost	0.07102	0.2110	0.943
MSC GBDT	0.12965	0.3240	0.893
Normalize GBDT	0.124575	0.3203	0.886
Original GBDT	0.214526	0.2161	0.861
MSC RF	0.142478	0.4124	0.834
Normalize RF	0.325144	0.4126	0.771
Original RF	0.312425	0.1556	0.742
MSC SVR	0.401254	0.1257	0.734
Normalize SVR	0.414214	0.1476	0.723
Original SVR	0.414410	0.5324	0.702

Where the Light GBM (GBM), XGBoost, Gradient Boosting Decision Tree (GBDT), Random Forest (RF), and Support Vector Regression (SVR). Where the performance indicators are used as MSE, MAE, and R^2 with different conditions.

The output performance indicators for these models under different optimization methods. The results indicate significant performance improvements in the optimized models, with the Improved Aquila Optimizer-optimized model outperforming those optimized by AO and PSO. The Light GBM model optimized by Improved Aquila Optimizer achieved R^2 values infinitely close to 1.

5. Conclusion

The research study demonstrated the Light Gradient Boosting Machine model, optimized with the Improved Aquila Optimizer (IAO) and utilizing Laser-Induced Fluorescence (LIF) technology for spectral data acquisition, significantly enhances the accuracy and efficiency of high-speed train bearing fault diagnosis. The application of Multivariate Scattering Correction preprocessing further improved model performance by effectively reducing noise and minimizing errors from experimental setups. The IAO optimization notably outperformed traditional methods, such as Particle Swarm Optimization, by achieving faster convergence and superior parameter settings. Overall, this integrated approach of MSC-KPCA preprocessing, Light GBM modeling, and IAO optimization presents a robust and reliable solution for enhancing the safety and reliability of rail systems through precise fault detection.

References

1]. Chatterton, S., et al. *Diagnostics of rolling element bearings for the traction system of high speed trains:* Experimental *evidences*. in *International Design Engineering Technical* Conferences *and Computers and Information in Engineering Conference*. 2013. American Society of Mechanical Engineers.

[2] Xu, G., et al., *High-speed train wheel set bearing fault diagnosis and prognostics: A new prognostic model based on extendable useful life.* Mechanical Systems and Signal Processing, 2021. 146: p. 107050.

[3] Cheng, C., et al., *Enhanced fault diagnosis using broad learning for traction systems in high-speed trains.* IEEE Transactions on Power Electronics, 2020. 36(7): p. 7461-7469.

[4] Leveson, N.G., *A new approach to system safety engineering.* Manuscript in preparation, draft can be viewed at http://sunnyday. mit. edu/book2. pdf, 2002.

[5] Ma, Q., et al., *A coupling model of high-speed train-axle box bearing and the vibration characteristics of bearing with defects under wheel rail excitation.* Machines, 2022. 10(11): p. 1024.

[6] Zhang, S., et al., *Deep learning algorithms for bearing fault diagnostics—A* comprehensive *review.* IEEE Access, 2020. 8: p. 29857-29881.

[7] Singh, P., et al., *Internet of Things for sustainable railway transportation: Past, present, and future.* Cleaner Logistics and Supply Chain, 2022. 4: p. 100065.

[8] Deguchi, Y., *Industrial applications of laser diagnostics*. 2011: CRC Press.

[9] Modak, P. and B. Modak, *Electronic structure investigation of intrinsic and extrinsic defects in LiF.* Computational Materials Science, 2022. 202: p. 110977.

[10] Arbyn, M., et al., *Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India.* International journal of cancer, 2008. 123(1): p. 153-160.

[11] Abend, T., et al. *Line-scan detection system to identify rare earth elements in rocks*. in *2019 IEEE SENSORS*. 2019. IEEE.

[12] Lucht, R.P., *Applications of laser-induced fluorescence spectroscopy for combustion and plasma* diagnostics. Laser spectroscopy and its applications, 2017: p. 623-676.

[13] Vaswani, A., *Attention is all you need.* arXiv preprint arXiv:1706.03762, 2017.

[14] Das, O., *Real-time intelligent fault diagnosis of rotating machines based on Archimedes algorithm* optimised *gradient boosting.* Nondestructive Testing and Evaluation, 2024. 39(2): p. 474-512.

[15] Agirre, X. and G. Abad, Railway *traction.* Power Electronics and Electric Drives for Traction Applications, 2016: p. 221-361.

[16] Gopi, S. and P. Mohapatra, *Fast random opposition-based learning aquila optimization algorithm.* Heliyon, 2024. 10(4).

[17] Hindarto, D., *Case Study: Gradient Boosting Machine vs Light GBM in Potential Landslide Detection.* Journal of Computer Networks, Architecture and High Performance Computing, 2024. 6(1): p. 169-178.

[18] Sharma, S., et al., *Identifying Parkinson's Patients by a Functional Gradient Boosting Approach*, in *Intelligent Technologies and Parkinson's Disease: Prediction and Diagnosis*. 2024, IGI Global. p. 288-304.

[19] Vijayakumar, A. and D. Uma, *A GBDT-BCO Technique based cost reduction and energy* management *between electric vehicle and electricity distribution system.* Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2021: p. 1-25.

[20] Hartanto, A.D., Y.N. Kholik, and Y. Pristyanto, *Stock Price Time Series Data Forecasting Using the Light Gradient Boosting Machine (LightGBM) Model.* JOIV: International Journal on Informatics Visualization, 2023. 7(4): p. 2270-2279.

[21] Zhang, J., et al., *A* new *bearing fault diagnosis method based on modified convolutional neural networks.* Chinese Journal of Aeronautics, 2020. 33(2): p. 439-447.

[22] Zhang, S., et al. *Deep learning algorithms for bearing fault diagnostics-a review*. in *2019 IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*. 2019. IEEE.

[23] Khan, M.A., et al., *The bearing faults detection methods for electrical machines—the state of the art.* Energies, 2022. 16(1): p. 296.