International Research Journal of Multidisciplinary Scope (IRJMS), 2024; 5(3):798-808

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2024.v05i03.0857

Equipment Health Monitoring Using Machine Learning Techniques

Pankaj V Baviskar*, Chitresh Nayak

Department of Mechanical Engineering, Medicaps University, Indore, MP, India. *Corresponding Author's Email: pankaj.rcpit@gmail.com

Abstract

The rapid growth of science and technology in modern civilisation has led to an increase in the size, complexity, and automation of machinery and equipment. Two of the most important aspects of modern industrial production are problem identification and machinery condition monitoring. Early problem detection is made possible by effective condition monitoring, which is crucial when taking into account variables like production efficiency, operational dependability, maintenance costs, and downtime. Research on the identification of problems and the health monitoring of machinery has practical implications. For the purposes of equipment monitoring and fault diagnostics, information on the temperature, vibration, noise level, and lubrication state of the equipment is recorded. After that, the information is utilised to identify the issue's primary source and put remedial measures in place. A condition monitoring system's core elements are fault prediction, feature extraction, and problem diagnostics. Feature extraction and fault diagnostics are essential for normal detection, problem localisation, and failure severity prediction. This paper includes fault diagnosis and applications of computational intelligence in condition monitoring and fault detection also this paper presents a method for equipment status monitoring using Machine Learning (ML) techniques. Popular machine learning (ML) classification methods like Random Forest (RF), Random Tree (RT), Naive Bayes (NB), XG Boost (XGB), and Logistic Regression (LR) are used for assembling. The pressing need to increase machine reliability and reduce the possibility of production losses due to machine breakdowns is the reason behind the growing emphasis on machine condition monitoring.

Keywords: Failure Prediction, Machine Learning, Naive Bayes, Production, Random Forest, Random Tree, Unexpected Downtime.

Introduction

A number of core ideas underpin predictive maintenance and equipment health monitoring, which are intended to maximize maintenance efforts, minimize downtime, and guarantee operational dependability. The main ideas consist of Constant Monitoring: It's critical to continuously gather data from sensors and monitoring systems. This entails gathering data in real time on a variety of factors, including vibration, temperature, pressure, and performance measures. Data analytics and machine learning: To find patterns, anomalies, and trends in the gathered data, advanced analytics techniques, such as machine learning algorithms, are used. These methods aid in spotting early indications of equipment deterioration or failure. State-Based Maintenance: Maintenance tasks are planned according to the equipment's actual state rather than following a standard schedule, which may be expensive and ineffective. By being proactive, this method lowers needless maintenance and increases the lifespan of equipment. Predictive Models: To forecast equipment performance and foresee possible faults, predictive models are constructed using real-time inputs as well as past data. To produce precise forecasts, these models combine machine learning, artificial intelligence, and statistical techniques.Vibration signals were gathered from the gearbox of an all-terrain vehicle (ATV) to conduct the condition monitoring on an assembly of a gear train. The goal of this monitoring was to ensure that the gear train was in good working order. The results of the finite element analysis allowed for the specific location of the defect in gear to be located and mapped out. The vibration data was collected with the use of an accelerometer, and this was done both while the gear was running regularly and when it was being simulated to be performing improperly. An algorithm for deep learning that was derived from

(Received 27th March 2024; Accepted 17th July 2024; Published 30th July 2024)

the tree family was used in that study to classify the various gear train scenarios (1). This family of algorithms includes several different classification approaches, including choice trees, random forests, and random trees, amongst others. The random forest strategy generated the highest degree of classification accuracy, which approached 99% when compared to other tree family algorithms. This was achieved by producing the largest number of trees. The criteria for making decisions were used throughout the process of creating an online monitoring system to display the present condition of the gear. That study might be helpful in the deployment of online gear health monitoring in ATVs, which would contribute to the safety of drivers (2).

The problem of maintenance has risen to the top of the priority list as a direct result of the widespread use of electrical apparatus in a wide range of settings and environments. It is possible to categorize, on a fundamental level, the pressures that influence the functioning of electrical equipment into four distinct categories. TEAM stresses are an abbreviation that stands for thermal, electrical, ambient, and mechanical stressors. This term is used to refer to all these different kinds of pressures under a single umbrella term. The machine is more likely to develop flaws as a direct result of these pressures, which increases the likelihood that the machine could break down. The statistics suggest that stator winding problems account for 36% of all motor failures. It's usually because of a short circuit that spans many turns when a winding fails. If timely maintenance is not performed properly, this problem might worsen into a phase-to-phase or phase-to-ground short circuit. It's not easy to see this inter-turn issue while it's still in its infancy, thus spotting it is just one of the many challenges associated with identifying this issue. Therefore, one of the goals of this study is to develop a machine learning approach that permits continuous monitoring of the equipment's health (2). With the help of the provided model, early fault detection could be predicted, which boosts the efficiency of the machine. The objective of this research is to create a system that could monitor the health of machines in real-time utilizing machine learning and IoT technologies, to construct an entirely automated condition monitoring system making use of Internet of Things devices and multi-agent systems and to implement the microcontroller for controlling and transmitting data, so that real-time monitoring might be carried out efficiently. For industrial maintenance and equipment management, the practical implications of combining FFT, time series analysis, and machine learning classifiers in equipment condition monitoring are revolutionary. These developments provide businesses a competitive edge in the market, better predictive maintenance, increased safety and compliance, improved dependability and efficiency, and data-driven decision making. However, resolving issues with technology, data management, and staff development is necessary for successful adoption. The use of sophisticated monitoring technology demands a personnel proficient in predictive maintenance, machine learning, and data analysis. For their staff to have these abilities, companies will need to make training and development programme investments (3).

Methodology

The term "research methodology" refers to the process through which authors outline the specifics of how they plan to conduct their studies, and the name "research methodology" itself refers to this process. It's a way of approaching a study problem that is reasonable and systematic. It is common practice for authors to provide a brief explanation of their technique to guarantee that their research produces accurate and trustworthy results and achieves the stated goals and objectives. The process takes into account, not just the data itself but also its origins, potential uses, and methods of acquisition (3).

Data Description

The dataset that was utilized in this study is obtained from the data collected using IoT sensors, added on an Automobile part making machine. Dataset consisting with nearly 55,000 records, with 28 different features like hydraulic oil temp, timestamp, production count, downtime etc. Supervised Aggregative Feature Extraction employed with, a supervised regression approach that uses a functional learning paradigm to represent learning problems with time-series type inputs and scalar outputs. There are certain datasets that could not be used for study or for financial gain (3). The dataset that is given in this study examines the data that is associated with predictive maintenance. Predictive maintenance, often called condition-based maintenance, is performed to reduce the risk of an unexpected failure by keeping tabs on the equipment's performance and condition even while it's being put through its paces in regular operations. Predictive maintenance aims to anticipate when equipment may break down by using a variety of indicators and data. Parts could be maintained in two ways: either on a regular schedule, even if they are functioning well (Preventive Maintenance), or just when they break (Reactive Maintenance). Predictive Maintenance sidesteps the problems associated with both Reactive Maintenance (which wastes parts before they're used up) and Preventive Maintenance (which wastes time and resources) (unscheduled downtime). In Predictive Maintenance, failure points are anticipated based on historical data on the equipment's state of health. Parts replacements could be pre-emptively arranged in this way. It takes a combination of statistical, machine learning, and domain-specific techniques to deal with missing data and outliers. For equipment condition monitoring systems to be accurate and reliable, these problems must be handled effectively. In order to estimate missing values based on the maximum likelihood estimation, we employ deletion methods, imputation methods, and model-based methods in our approaches for handling missing data. This approach requires a lot of computing power but is statistically sound. We employ a variety of methods to handle outliers, including detection techniques such as statistical, visual, and machine learning techniques; treatment techniques such as transformation and deletion; and hybrid techniques such as model-based imputation for missing data and outliers. In all, the dataset has the following:

Telemetry Time Series Data: The term for the data that is gathered by the equipment is telemetry. This is data that is gathered from the sensors that are located on the IoT devices and are sent to apps by those devices. The term "time series data" refers to a series of data items that have been gathered continuously over an extended time and at regular intervals. Hourly averages of voltage, pressure, temperature and vibration were gathered from an automobile part making machine during 2021, and the report details those findings (3).

Error: They are malfunctions that occur while the machines are in use. As these malfunctions do not cause the machines to power down, they are not deemed failures. Due to the hourly sampling of telemetry data, any discrepancies in dates and hours are rounded to the nearest hour.

Maintenance: This table is updated to reflect the addition of a new record whenever a machine's component undergoes replacement. Two circumstances call for the replacement of components: The replacement was performed by the technician at the appointment that was previously planned (Proactive Maintenance) the failure of a part necessitates an emergency repair performed by the expert (Reactive Maintenance). Data related to this situation is filed away in the Failures section. All the records for maintenance for both 2014 and 2015 are included. Due to the periodic nature of the telemetry data collection, all times are rounded to the nearest hour (3).

Failures: Every entry in the log reflects a component that had to be replaced because it had failed. These data are a part of the larger Maintenance data collection. Since telemetry data is acquired on an hourly basis, these measurements are rounded to the nearest hour for clarity.

Metadata of Machines: Metadata is data that represents other data, providing a structured reference that aids in the sorting and identification of properties of the material that it represents. The metadata associated with the machine contains information on both the sort of model it is and its age.

Technique Used

The suggested technique makes advantage of much technological advancement such as the Fast Fourier Transform, fuzzy logic, Wavelet Power Spectrum, and Machine Learning classifiers, among others.

Fast Fourier Transform: The Power Spectral Density (PSD) could be calculated by modifying the predicted autocorrelation sequence, which is found using nonparametric techniques, dividing the result by two, and then using the Fourier transform and division to find the value (3). Welch's technique is one of these approaches; further information on this method could be found lower down the page. Utilizing the data sequence in combination with the data windowing procedure results in the creation of periodograms that have been brought up to date (4). It is represented by the symbol for the information sequence (m) in equation 1.

$$y_j(m) = y(m + jD),\ m = 0, 1, 2, \ldots\ldots . M - 1 \quad [1]$$

while $j = 0, 1, 2 \ldots \ldots \ldots . L - 1$.

Consider jD to be the point at which the jth Series begins. Afterward, L of length 2M represents the data divisions that have been produced. The output periodograms in equation 2 that resulted provide (5).

$$P_{yy}^{\approx(j)}(f) = \frac{1}{MU}\left|\sum_{m=0}^{M-1} y_j\ (m) y^k w(m) e^{-j2\pi fn}\right|^2 \quad [2]$$

In this case, the window function's normalizing factor of the power is represented by U, which is signified in equation 3.

$$\mathbf{U} = \frac{1}{M}\sum_{m}^{M-1} \mathbf{W}^2\,(\mathbf{m}) \quad [3]$$

For a number of reasons, the Fast Fourier Transform (FFT) is extremely important when it comes to signal processing and equipment state monitoring. Such as Condition Assessment, Fault Detection, Identification of Vibrations and Oscillations, and Frequency Analysis. Finite Form Analysis (FFT) facilitates targeted maintenance by identifying particular frequencies linked to problems or inefficiency. This minimizes downtime and lowers maintenance costs by enabling proactive interventions prior to breakdowns.

Fuzzy Logic: Since fuzzy logic can deal with the inherent uncertainty and imprecision of real-world data, it is important for monitoring the state of equipment. Monitoring the condition of equipment frequently entails handling noisy and ambiguous sensor data. This imprecise data may be efficiently processed using fuzzy logic, which makes it appropriate for real-world scenarios where data may lack clarity. The process of replicating the way that individuals logically think while dealing with fuzzy information is what is known as fuzzy theory. This kind of theory is well-suited for the qualitative investigation of complicated large-scale systems (6). Due to the complexity of engineering practice, developing a reliable mathematical model to express the connection between cause and effect is challenging. Thus, the use of fuzzy logic theory in defect diagnostics is more akin to how humans really think and express themselves verbally. Fuzzy logic is an efficient pattern identification technique that has found widespread use in the fields of electricity (7), transmission lines, transportation, and manufacturing. Since it closely resembles human logical thinking, fuzzy logic is an effective means of information transfer. Most judgments made in the real world are made in such situations when accurate knowledge of limitations, objectives, and consequences is lacking. Fuzzy logic might be a useful tool to aid in this kind of situation (8). It is common knowledge that feature assessment is the central problem when developing a feature selection algorithm and this is because feature selection (FS) is a crucial pre-processing step in machine learning and pattern recognition. Noise could be reduced in the feature selection process, leading to improved classification accuracy. Fuzzy logic is used to deal with the idea of a partial truth, where the true value might be somewhere between false and true. The steps of Fuzzy Logic are Fuzzification, Inference, and Defuzzification.

Fuzzification: First, the sensors collect the clear input data, and then, using membership functions, the data are transformed into a fuzzy input set, linguistic words, and linguistic variables. Finally, the crisp input data are used in the analysis

Inference: The inference could be reached by putting the rules into practice. The system is going to function following a predetermined set of guidelines such as IF-THEN.

Defuzzification: In the end, a crisp output is achieved by mapping the fuzzy output with the assistance of the membership function.

The algorithm used for fuzzy logic is as follows:

1. The establishment of the terminology and the linguistic variables.
2. Build the membership function (MF)
3. The beginning of the rule's application.
4. Membership functions are used to transform sharp data into fuzzy data.
5. Incorporate the principles of fuzzy input into the workflow.
6. Transform the hazy output into a clear one.

Wavelet Power Spectrum: Because of its capacity to offer comprehensive time-frequency analysis, identify transient events, manage non-stationary signals, and improve feature extraction for precise fault diagnosis and early warning, the Wavelet Power Spectrum is important for equipment condition monitoring. Through the use of more efficient maintenance techniques, this results in increased dependability, decreased downtime, and cost savings. The wavelet power spectrum was used to extract the features. The genes have been chosen, and the findings have been listed. They have been evaluated alongside the outcomes of past research initiatives and found to be comparable. There are several benefits to using this approach to feature selection and clustering. The approach is straightforward in contrast to other feature selection strategies, and it requires no specialized software for implementation since wavelets are so readily accessible in popular programs like MATLAB. As a bonus, it could be used in combination with numerous well-established classification techniques, often requiring a smaller sample size. In comparison to more traditional approaches, this one move along quickly (8). Selecting unique genes from inside a cluster of previously chosen genes is one way to get around the issue of redundant genes, which is a problem that arises when using feature selection approaches. Initial results from using wavelet power spectrum for feature selection with microarray data are promising, and more studies into its ease of use, speed, effectiveness, and suitability for a wide range of datasets might lead to applications in genomic signal processing using microarrays.

Machine Learning Classifiers: Machine learning classifiers offer automatic, precise, and real-time diagnostics, they are essential for keeping an eye on the state of equipment. The versatile nature of these tools allows them to be applied to a wide range of tasks, including improving equipment reliability, streamlining maintenance procedures, and cutting operational expenses. Machine learning has seen widespread usage in the context of condition monitoring systems as was discussed in the preceding. In this paper goes through the steps involved in designing an ensemble of classifiers, as well as the benefits of using such a method. The theoretical underpinnings of several different machine learning methods, including RF, RT, NB, XG Boost, and LR, are provided in a condensed form (9-13).

Random Forest (RF): The random forest technique is commonly used in machine learning algorithms that are used to address issues. It is based on the concept of collective learning, which is a method for combining several different classifiers to handle a wide variety of challenging issues and improve the accuracy of the model (9). Figure 1 depicts the RF algorithms' fundamental flowchart in its entirety.

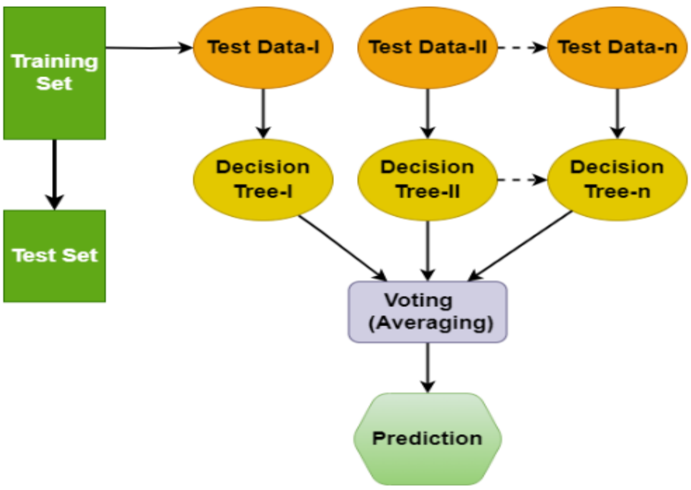

Figure 1: Random Forest Algorithms (10)

Random Tree (RT): It is common to practice in the field of data mining to utilize the supervised learning algorithm random tree to solve classification issues. It creates a network of learners that collaborate to build a tree of choices (11). Bifurcations occur at each node using the optimal set of variable parameters. In this decision tree-building process, N-selected qualities are used to create a tree. Until the node class is reached, the number of attributes chosen is arbitrary, and the attribute with the biggest entropy gain is picked for further bifurcation. It was Leo Breiman and Adele Cutler who first described random trees. The algorithm can deal with both classification and regression problems. The forest, an ensemble of tree predictors, is the basis for the random tree. The random trees classifier uses each tree in the forest to assign a classification to the input feature vector, and then returns the classification that received the most "votes." Whenever a regression is performed, the classifier's answer is the mean of all the individual trees' answers taken together. Machine learning's Random Trees are a hybrid of two separate methods, the single-model-tree-based Decision Forest, and the more generalized Random Forest. Each node in a model tree, which is a kind of decision tree, stores a linear model that has been tuned specifically for the subspace specified by that node. Random Forests have been found to significantly outperform individual decision trees; they do this by using two distinct randomization techniques to build a diverse set of trees. Starting with replacing each tree in the training set with a sample taken in a Bagging style. Second, instead of continually computing the best possible split for each node, a tree is built by randomly selecting a subset of all characteristics to evaluate at each node, and then finding the optimal split for that subset. In a first of its type, random model trees combine model trees with random forests for classification. As a consequence of using this product for split selection, random trees tend to have a fair amount of diversity. Additionally, a single ridge value setting is effective for all leaves, which streamlines the optimization process (12).

Naïve Bayes (NB): Specifying the insurmountable sample difficulty of studying Bayesian classifiers must seek methods to decrease it. The Naive Bayes classifier does this by assuming conditional freedom which decreases the number of constraints to be calculated when modeling P(X|Y) from 2(2n-1) to only 2n.Assumed L, M, and N are 3 sets of random variables. If and only if the probability distribution leading L is independent of the value of M provided N, then L is conditionally independent of M given N. i.e.

$$(\forall_{\mathbf{i,j,k}})\mathbf{P}(\mathbf{L} = \mathbf{l_i}|\mathbf{M} = \mathbf{mj}, \mathbf{N} = \mathbf{nk}) = \mathbf{P}(\mathbf{L} = \mathbf{l_i} \mid \mathbf{N} = \mathbf{nk}) \quad [4]$$

Data is categorized using a looping split in the case space. Directed trees do not have outgoing edges, and the root node of the decision tree is the node from which all other nodes branch out. A single tip has been received by each of the additional nodes (13, 14).

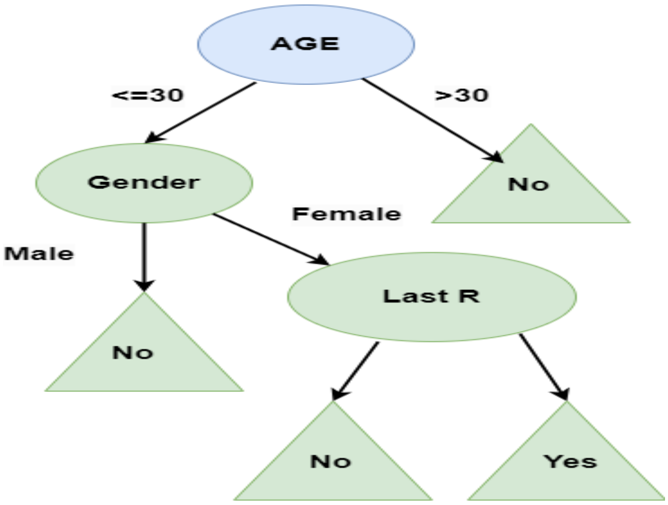

Figure 2: Decision Treed (14)

An internal or evaluation node is a node with outward edges. Figure 2 shows a decision tree for determining if a prospective customer can reply to a direct mailing. The rest nodes are stated as leaves. An internal or evaluation node is a node with outward edges. The rest nodes are mentioned as leaves. It is also called decision nodes or terminals.

$$\boldsymbol{E}(\boldsymbol{s}) = \sum_{i=1}^{c} \boldsymbol{p}_{\boldsymbol{i}\ \log_2 p_i} \quad [5]$$

Where E(s) is entropy and p is probability.

XG Boost: An eXtreme Gradient Boosting (XGBoost) classifier is utilized by the recommended approach to recognize botnet assaults when specific conditions are satisfied. An interesting new tree-based ensemble learning classifier, XGBoost (often abbreviated as XGB), can also be referred to by its acronym. In terms of gradient-boosted decision trees, this is the best implementation to date. Each successive decision tree in a gradient-boosted decision tree series contributes to the improvement of the model and influences the performance of the previous tree in the series (15). Weak classifiers might be coupled with XGBoost to create a strong one. Unlike RF, which relies on pre-existing data, XGBoost considers confirmed decision trees. Gradient boosting is an iterative method of enhancing the quality of the baseline loss function. The goal is to have the previous phase's residue be as small as possible. The term "residual" could be used to refer to the gap between the ideal and actual values. The model is considered ready for deployment whenever the residual value falls below a threshold value. However, if numerous decision trees reach a given value before the residual reaches that value, training is ended and the final model is selected. The XGB model differs most noticeably from gradient boosting in its acceptance of regularization, ability to exploit parallel processing, and execution speed. Here's a quick calculation to remember XGBoost's goal (16). Utilizing this function, an evaluation of the model's overall performance might be made (17).

$$\mathbf{P}(\boldsymbol{\theta}) = \mathbf{t}(\boldsymbol{\theta}) + \mathbf{r}(\boldsymbol{\theta}) \quad [6]$$

Where θ denotes the parameters, r the regularization period, and t the number of training iterations that were discarded.

Logistic Regression (LR): Logistic regression is a form of regression analysis used to determine the optimal values for the logistic model's input variables (18). Logistic regression is a method of discriminative classification that takes as input a vector of values that could be expressed as real numbers. Features or predictors are the aspects of the input vector that would be used for classification. The examination of multiclass classification data could potentially benefit from logistic regression. Logistic regression's use of the probability P of a dichotomous event, which is commonly assumed to have originated from the Bernoulli trial, is often tied to studying the properties of the events themselves (19). The mathematical explanation for the Logit function is that it is the natural log of the likelihood that Y falls into one of the categories (20). Assuming p to represent the probability, the corresponding Logit function for p might be written as (21).

$$\mathbf{Logit}(\mathbf{p}) = \mathbf{1n}(\mathbf{p}/\mathbf{1} - \mathbf{p}) \quad [7]$$

Proposed Process

Figure 3 is an illustration of the architectural components that make up the model. In order to describe the process flow of the proposed technique, the following steps are required:

Step 1: Start.

Step 2: In this step, the current, voltage, frequency, and power factor could all be read with the help of the PZEM-004T sensor.

Step 3: At this stage, the data should be sent via the microcontroller, which should also be used to operate it.

Step 4: In this step, the time series would now undergo the Fast Fourier transformation approach. Time-based information about the signal is converted into a frequency-domain signal.

Step 5: In this step, the process of data mining, the term "pre-processing" refers to the transformation of raw data into a format that can be mined more effectively.

Step 6: Now, Feature Extraction is performed with the aid of the Wavelet Power Spectrum, and Feature Selection is performed with the use of Fuzzy Logic to determine the minimum number of distinct Attribute Combinations.

Step 7: At this point, the selected characteristics include the following: Voltage, Angular Acceleration, Pressure, and Vibration.

Step 8: In this step, additional data is separated into two categories: the testing set, and the training set. The data set that is used in the training of the model is referred to as the "training set." Once the training phase of the process is complete, the model is validated using a separate set of data that is referred to as the test set.

Step 9: In this step of the process, data would classify using the Machine Learning classifiers using the ensemble approach. In the process of classification, Ensembling is used, and it is engaged in by the RF, LR, RT, NB, and X G Boost classifiers.

Step 10: Further, if the model has been trained, go to the next step, which is the trained Machine Learning classifier; otherwise, return to step 9.

Step 11: The trained ml classifier was applied to responses from measurement sites.

Step 12: Finally, a real-time output has been accomplished.

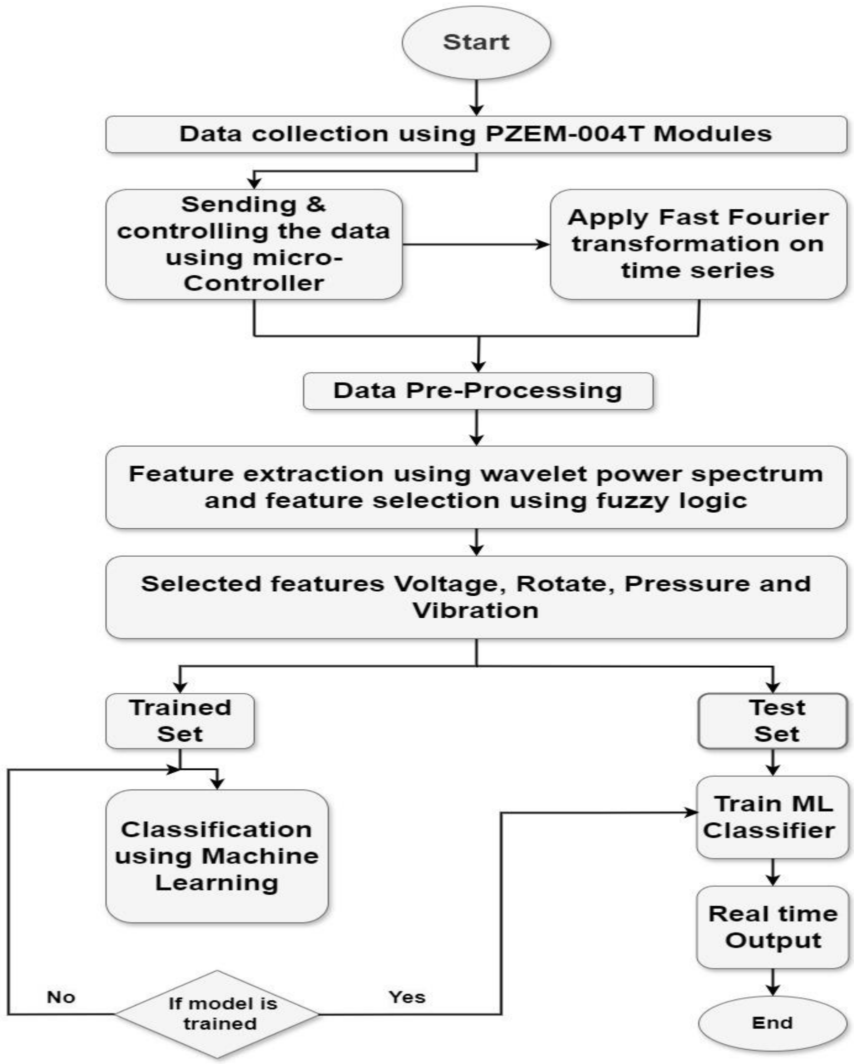

Figure 3: Proposed Process

Advantages and Limitations of the suggested strategy

Time series analysis, machine learning classifiers, and the Fast Fourier Transform are all combined to provide a strong and complete foundation for equipment condition monitoring. By combining the advantages of both strategies, condition monitoring and predictive maintenance become more precise, effective, and dependable. It also gives High Accuracy and Precision, Adaptability, Handling Non-Linear Relationships, Enhanced Fault Detection, Improved Predictive Maintenance, Noise Reduction and Feature Enhancement, Real-Time Monitoring and Alerting, Scalability etc. however, there are drawbacks as well, including difficulties with integration complexity, data requirements, computational resources, model interpretability, and continuing maintenance. In order to overcome these obstacles, thorough planning, sufficient funding, and proficiency with both the individual techniques and their combination are needed (22).

Results and Discussion

The data collected using IoT sensors, added on an Automobile part making machine. Dataset consist of around 55,000 records, with 28 different features like hydraulic oil temp, timestamp, production count, downtime etc. Supervised Aggregative Feature Extraction employed with, a supervised regression approach that uses a functional learning paradigm to represent learning problems with time-series type inputs and scalar outputs.

A horizontal bar chart comparing the accuracy percentages of various classifiers is shown in Figure 4 "Classifier Accuracy". The y-axis lists the classifiers that have been tested, and the x-axis shows the accuracy percentages, which go from 0% to 100%. The accuracy of the related classifier is indicated by the length of each bar. Longer bars on classifiers yield higher accuracy percentages. Figure 5 "Classifier Loss" shows a horizontal bar chart that contrasts the classifiers' respective loss levels. The lowest loss values displayed by KNeighbors Classifier, Decision Tree Classifier, and XGB Classifier indicate that they are the highest-performing classifiers with the least amount of mistake and Quadratic Discriminant Analysis exhibit higher loss values, indicating poorer performance relative to the other classifiers.

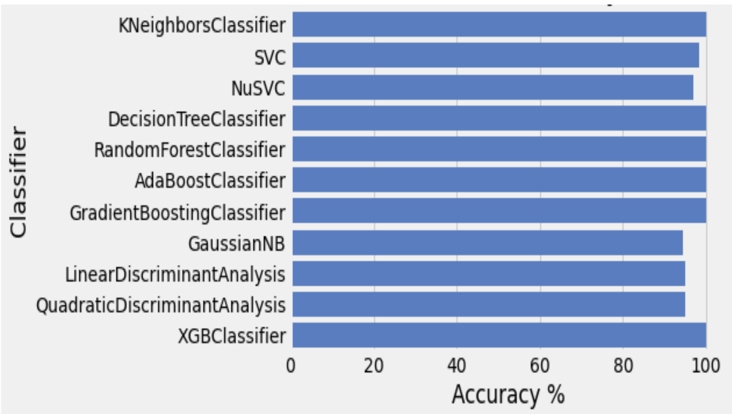

Figure 4: Classifier Accuracy

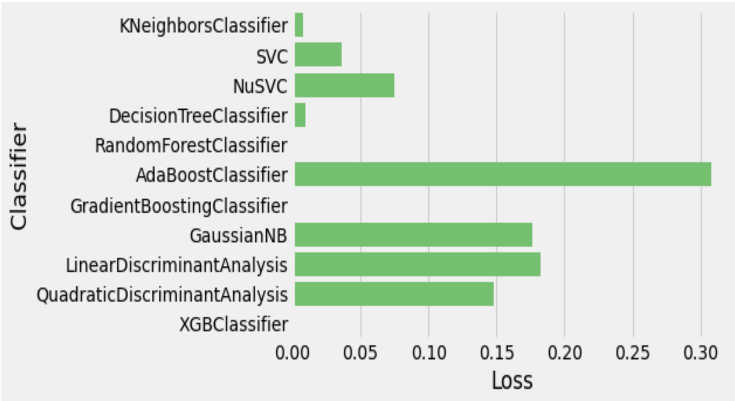

Figure 5: Classifier Loss

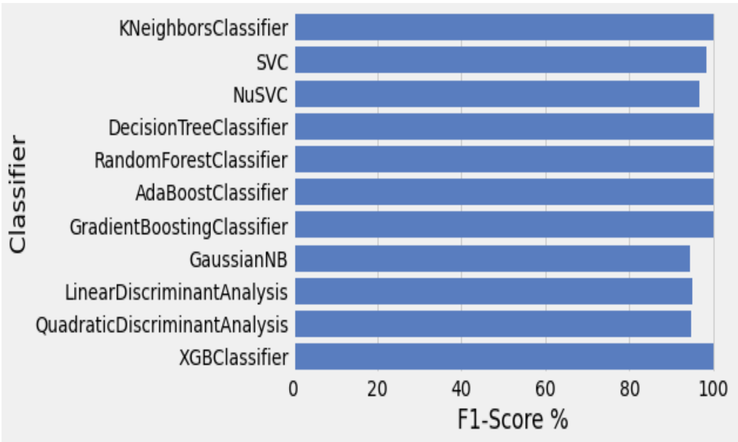

Figure 6: Classifier F1-Score

Figure 6 "Classifier F1-Score" shows a horizontal bar chart that contrasts the various classifiers' F1-scores. A classifier's accuracy that takes into account both precision and recall is expressed as the F1-score. It is especially helpful when you need to account for both false positives and false negatives. It is the harmonic mean of precision and recall. Ensemble learning classifiers are often considered superior to traditional classifiers due to several key reasons: 1.Improved Accuracy and

Generalization: Ensemble methods combine multiple base learners (individual classifiers) to form a stronger composite model. By leveraging diverse models that may have different strengths and weaknesses, ensembles can often achieve higher accuracy and better generalization to unseen data compared to a single classifier. This is because ensemble methods reduce the risk of over fitting to the training data and can effectively handle complex relationships within the data. 2. Robustness to Noise and Outliers: Traditional classifiers can be sensitive to noise and outliers in the data, which may lead to poor performance. Ensemble methods, especially those like Random Forests or Bagging, average out the predictions of multiple classifiers, reducing the impact of noisy data points and outliers. This robustness helps improve the overall reliability of predictions. 3. Handling Complex Relationships: In many real-world datasets, the relationships between input features and the target variable can be complex and nonlinear. Ensemble methods can capture these complex relationships more effectively by combining models that specialize in different aspects of the data. Techniques like Gradient Boosting can sequentially learn from errors of previous models, gradually improving predictive performance. 4. Versatility across Different Tasks: Ensemble methods are versatile and can be applied to various types of machine learning tasks such as classification, regression, and even ranking. They are adaptable to different types of data and can incorporate different types of base learners, including decision trees, neural networks, or simpler models depending on the task requirements. 5. Reduction of Bias and Variance: Ensemble methods address both bias and variance issues. Bias refers to the difference between predicted values and actual values, while variance refers to the variability of model predictions for a given data point. By combining multiple models, ensembles can reduce bias and variance simultaneously, leading to more reliable predictions. 6. Interpretability and Explain ability: While some ensemble methods like Random Forests can provide feature importance's or decision pathways, others like boosting methods may be less straightforward in terms of interpretability. However, overall ensemble models are often as interpretable as their individual components, and techniques exist to interpret ensemble predictions. In summary, ensemble learning classifiers are typically preferred over traditional classifiers because they offer improved accuracy, robustness to noise, better handling of complex relationships, versatility across tasks, and effective reduction of bias and variance. These advantages make ensemble methods a powerful tool in machine learning for achieving high-performance predictive models.

Conclusion

The integration of Machine Learning (ML) and Industry 4.0 has expanded the potential for improving industry quality. The most rapidly developing area in the industries is the early identification and avoidance of failures, as well as the thorough pre-processing and efficient use of a variety of sensor data. Furthermore, it has been demonstrated that applicable machine learning approaches perform better for condition monitoring and failure prediction of any equipment that uses ICT-based technology. The majority of approaches now in use have concentrated on one or more particular system components. It is therefore difficult to create a machine learning-based model with the right attributes for a complicated system. To produce a decision of prediction, SVM model shows poor performance in training phase. To improve the performance in training phase and reduce the error, an XGBoost is used; and the prediction is accomplished for the given range of temperature. In addition, to predict the failure state with respect to down time, LSTM based model creates the long dependency problems. To overcome this, we can also use a weight optimized GRU model in future. The presented framework is well-suited for PdM planning and capable of accurately predicting future component condition for maintenance planning. In order to identify the crucial parameters in failure, it has been observed that it is necessary to assess the precise and timely results-based predictive maintenance technique along with the historical data.

Abbreviations

ML: Machine Learning
RF: Random Forest
RT: Random Tree
NB: Naive Bayes
XB: XG Boost

LR: Logistic Regression
ATV: All-terrain vehicle.

Acknowledgment

The authors would like to thank Medicaps University Indore & RCPIT Shirpur for the support.

Author Contributions

Pankaj V Baviskar handles conceptualisation, methodology, writing-original draft, preparation, data curation, investigation, writing-reviewing and editing, formal analysis, and visualisation. Dr. Chitresh Nayak oversees project administration.

Conflict of Interest

The authors declare no conflict of interest.

Ethics Approval

Not applicable.

Funding

This research received no external funding.

References

1. Gao Z, Cecati C, Ding SX. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. IEEE transactions on industrial electronics. 2015 Mar 26; 62(6):3757-67.
2. Gnanasekaran S, Jakkamputi L, Thangamuthu M, Marikkannan SK, Rakkiyannan J, Thangavelu K, Kotha G. Condition Monitoring of an All-Terrain Vehicle Gear Train Assembly Using Deep Learning Algorithms with Vibration Signals. Applied Sciences. 2022 Oct 27; 12(21):10917.
3. Subasi A, Kiymik MK, Alkan A, Koklukaya E. Neural network classification of EEG signals by using AR with MLE preprocessing for epileptic seizure detection. Mathematical and computational applications. 2005 Apr; 10(1):57-70.
4. Faust O, Acharya RU, Allen AR, Lin CM. Analysis of EEG signals during epileptic and alcoholic states using AR modeling techniques. Irbm. 2008 Mar 1; 29(1):44-52.
5. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A. Pytorch: An imperative style, high-performance deep learning library. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada. 2019; 32-44.
6. Zhang W, Jia MP, Zhu L, Yan XA. Comprehensive overview on computational intelligence techniques for machinery condition monitoring and fault diagnosis. Chinese Journal of Mechanical Engineering. 2017 Jul; 30:782-95.
7. Eristi H. Fault diagnosis system for series compensated transmission line based on wavelet transform and adaptive neuro-fuzzy inference system. Measurement. 2013 Jan 1; 46(1):393-401.
8. Suganthi L, Iniyan S, Samuel AA. Applications of fuzzy logic in renewable energy systems – A review. Renewable and Sustainable Energy Reviews, Elsevier. 2015; 48(C): 585-607.
9. Hengl T, Nussbaum M, Wright MN, Heuvelink GB, Gräler B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ. 2018 Aug 29; 6:e5518.
10. Gangwar MK, Sahu NK, Shukla RK, Nayak C. Machine learning based progressive crack fault monitoring on spur gear using vibration analysis. International Journal of Vehicle Noise and Vibration. 2024; 20(1):89-106.
11. Kalmegh S. Analysis of weka data mining algorithm reptree, simple cart and randomtree for classification of indian news. International Journal of Innovative Science Engineering & Technology. 2015 Feb;2(2):438-46.
12. Kalmegh S. Analysis of weka data mining algorithm reptree, simple cart and randomtree for classification of indian news. International Journal of Innovative Science, Engineering & Technology. 2015 Feb; 2(2):438-46.
13. Shaikhina T, Lowe D, Daga S, Briggs D, Higgins R, Khovanova N. Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. Biomedical Signal Processing and Control. 2019 Jul 1; 52:456-62.
14. Giron-Sierra JM. Digital Signal Processing with Matlab Examples, Volume 3. Singapore: Springer; 2017.
15. Parsa AB, Movahedi A, Taghipour H, Derrible S, Mohammadian AK. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident Analysis & Prevention. 2020 Mar 1; 136:105405.
16. Awal MA, Masud M, Hossain MS, Bulbul AA, Mahmud SH, Bairagi AK. A novel bayesian optimization-based machine learning framework for COVID-19 detection from inpatient facility data. Ieee Access. 2021 Jan 11; 9:10263-81.
17. Hammood BA, Sadiq AT. Ensemble machine learning approach for IoT intrusion detection systems. Iraqi Journal for Computers and Informatics. 2023 Dec 30; 49(2):93-9.
18. Bashir S, Khan ZS, Khan FH, Anjum A, Bashir K. Improving heart disease prediction using feature selection approaches. In2019 16th international bhurban conference on applied sciences and technology (IBCAST) IEEE. 2019 Jan 8; pp. 619-623.
19. Dwivedi AK. Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Computing and Applications. 2018 May; 29:685-93.
20. Sperandei S. Understanding logistic regression analysis. Biochemia medica. 2014 Feb 15; 24(1):12-8.
21. Zeng XQ, Li HF. Application of machine learning in disease prediction. Journal of Biomedical Engineering Research. 2021; 40(2):203-9.
22. Baviskar PV, Nayak C. A Review on Equipment Health Monitoring Using Machine Learning Techniques. In International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022) Atlantis Press. 2023 May 1; pp. 382-396.