EPJ Web of Conferences **355**, 01004 (2026) https://doi.org/10.1051/epjconf/202635501004
STAAAR 2025

Condition Monitoring of Rolling Element Bearings Through Statistical Features and Support Vector Machine

Pravin Gosavi[1], Nikita Chougule[1], SujalKadam[1], Saniya Naik[1], and Vaishnavi Gavade[1]*
[1]Department of Mechanical Engineering, KIT's College of Engineering (Empowered Autonomous) Kolhapur, 416234 Maharashtra, India

Abstract. In rotating machinery, bearings are essential and critical components, and their failures frequently result in expensive downtime. This paper studies the vibration of an SKF 6206 deep groove ball bearing, focusing on predictive maintenance and checking how the bearing performs at different speeds and defect sizes. From vibration signal, statisticalfeatures suchas root mean square (RMS), crest factor, peak factor, peak value, kurtosis, and skewness were extracted to help identify faults at an early stage. These features were then used to train a Support Vector Machine (SVM) model for classifying the health state of bearings. By comparing the predictions with experimental data, it was clear that the selected features with SVM gave reliable accuracy in telling healthy bearings apart from faulty ones. According to the study, this strategy can enhance machinery reliability and assist predictive maintenance.

1.Introduction

Rotating machinery is used in almost every industry, including manufacturing, power generation, and transportation. Bearings are an important part of every rotating machine. The most common type used is the Deep Groove Ball Bearing. The condition of rolling element bearings largely depends on smooth and efficient performance of such machines. Bearings are critical yet are often the most failure-prone parts. Even a very small defect in a bearing can increase vibration and heat, which eventually leads to failure. In industries, this usually means unplanned downtime, reduced productivity, and sometimes even complete breakdown of the machine. Due to these risks, nowadays condition monitoring and fault detection has become a key area of research.

There are different types of bearings, but the SKF 6206 deep groove ball bearing (DGBB) is mostly used in machines and other industries. It has simple construction, low cost and capacity to handle radial and axial load due to this it has become a popular bearing. And also, this bearing is used in studies for fault detection and condition monitoring.

[*] Corresponding Author: nikitachougule3127@gmail.com

detecting faults at an early stage in such bearing is important for predictive maintenance and for keeping machines running without interruption.

In the study of bearings, health vibration signals play an important role because they carry detailed information about the health of the bearing. The condition of machine changes is identified by calculating statistical features such as root mean square (RMS), crest factor, peak factor, peak value, kurtosis, and skewness. They are simple to calculate and give valuable insight into the presence of faults which is why they are commonly used features. Many researchers have found that these features can help in detecting faults at an early stage of bearing life.

In our work, we diagnose the bearing faults by using selected statistical features directly with Support Vector Machine (SVM) classifiers. SVM is one of the machine learning algorithms which is popular for reliability and ability to work with small size data sets. SVM trains separate healthy bearings from faulty bearings by using the extracted features. The results are compared with experimental data to check the performance of algorithms

The objective of this study is to show that we can obtain reliable fault detection results without complex transformation by using a simple set of statistical features with an SVM model. The results show that this approach is both accurate and efficient, making it a suitable option for predictive maintenance and for improving the reliability of rotating machines.

2.Literature Review

Many studies have been carried out on bearing fault diagnosis using different machine-learning methods and advanced signal-processing techniques. In this part of the work, the main aim is to go through some of the important research that used various feature extraction methods, different classifiers, and a few optimization approaches to understand how these techniques have been applied in the past.

A method to find bearing faults by cleaning the vibration signals using spectral kurtosis filtering was proposed by Lai et al. [1] proposed and then they have used Hilbert envelope demodulation for showing the faulty vibrational patterns. From this method they have extracted values that are kurtosis and the amplitude ratios, then they fed this data to Least Square support Vector Machine (LS-SVM) a classifier. This method was tested on two different widely used datasets; CWRU and MFPT and got 95% and 100% accuracy respectively. But they only tested it on standard datasets, not on real-life machinery. A multi-fault diagnosis method for the bearings under different operating conditions was developed by Imane et al. [2]. Then used to apply the Empirical Wavelet Transform (EWT) to the vibration signals. For extracting the features, they have used Gaussian Mixture Model (GMM) and Random Forest classifier. It was showing good accuracy for different conditions, its performance affected by high noise levels. A bearing fault detection approach was presented by Salunkhe et al. [3], where they used the Dimension Analysis Matrix Method (DAMM) along with an SVM classifier. In their work, small defects were made on the bearing using EDM, and vibration data was recorded at different loads and speeds. Basic statistical features such as RMS, skewness, and kurtosis were considered for classification. Their results showed that SVM performed better than ANN, CNN, and SOM models, but the testing was carried out only in a controlled laboratory setup. Jamil et al. [4] compared SVM and K-Nearest Neighbor (KNN) for fault detection using the well-known CWRU bearing dataset. They extracted a total of 18 statistical features from both the time and frequency domains and validated the results using k-fold cross-validation. Their findings showed that frequency-domain features gave higher accuracy. KNN reported 98.8% accuracy, which was slightly higher than the 96.2% achieved by SVM, highlighting how important the choice of features is in fault diagnosis.

Shen et al. [5] improved bearing fault diagnosis by using an Improved Grey Wolf Optimizer (IGWO) to tune the SVM parameters. The IGWO algorithm used nonlinear contraction and dynamic weights, which helped it avoid being stuck in local optima and also made the convergence faster. Their model, tested on both the CWRU dataset and a full life-cycle experiment, reached 98.75% accuracy and performed better than PSO-SVM and the standard GWO-SVM.Wang et al. [6] worked on wheelset bearing fault detection and combined vibration and sound signals inside a GWO-SVM framework. In their approach, the GWO optimized the SVM hyperparameters. Their method achieved 98.3% accuracy, and the prediction time was extremely low (0.0027 ms per segment), making it faster than many traditional methods. Zheng et al. [7] used Wavelet Packet Decomposition for feature extraction and then applied PCA to reduce the feature size. They trained a multiclass SVM with an RBF kernel and achieved 97.22% accuracy, which highlighted how effective wavelet-based features can be in fault classification. Fasana et al. [8] compared Spectral Kurtosis (SK) and SVM for identifying useful frequency bands in bearing fault diagnosis. Their experiments were conducted on vibration signals from a high-speed gearbox under different fault sizes and speeds. SK automatically selected the most informative frequency ranges, while SVM needed training but still gave strong classification performance. Both methods gave similar results. Zhu et al. [9] proposed using Empirical Mode Decomposition (EMD) to extract intrinsic mode functions (IMFs). They then calculated kurtosis values from the first six IMFs and used them as input features for SVM. The model achieved 93.75% accuracy on the CWRU dataset and proved to be effective even in noisy environments. An online fault diagnosis method was developed by Gangsar and Tiwari [10] for induction motors using multiclass SVM. To extract features, they used vibration and current signals and then applied multiclass SVM to it as a classifier. In this, they used grid search with cross-validation to tune parameters. They detected 10 different faults across different loads and speeds. Pule et al. [11] combined PCA to reduce the number of features and then applied SVM as a classifier. They tested it with real life data from Mendeley repository and worked under varying speeds by compounding faults. As a result, they have got 97.4% accuracy. A detailed review of intelligent condition monitoring techniques for rolling bearings was carried out by Kannan et al. [12]. They told us about the different sensing methods, features extraction techniques, machine learning, and deep learning approach. They specifically highlighted that sensor fusion is very useful. They have also mentioned challenges like noise in signals, sensor degradation; suggesting that adaptive classifiers and probabilistic models are good options for future research. From the existing literature, it is clear that many studies rely on the complex feature extraction methods and most of them also use multiple domain features or large datasets to achieve high accuracy. Unlike earlier studies that depend on complicated transformations, our study focuses on checking how far basic statistical features like RMS, Peak value, etc. can go in detecting a small outer race defect. In this, we are comparing single features-based classification with a combined multi feature vector using SVM that makes the findings more practical for simple monitoring systems.

3.Methodology

3.1. Dataset Details

In this study, we have used the vibration data that was obtained by physical experimentation of the Bearing Dataset. This data was taken from a motor-driven mechanical system under operating different conditions. It is a setup where we created artificial defects using electro-discharge machining that replicate real-life bearing defects.

We highly focused on a specific part of that bearing dataset which had the vibration signals from two different conditions; one is with healthy bearings and other with a 0.5 mm defect at outer race. These signals were recorded at different motor speeds; 1200 RPM and 1500 RPM, with no motor load and no shaft tilt or misalignment. We were able to record even small vibrations caused by these faults very effectively because of the wide frequency range of the recordings. This kind of setup makes it possible to clearly observe even small defects that disturb the vibration signal. It plays a key role in detecting faults at an early stage in rotating machinery. In this study, we are focusing on such specific conditions, aiming to identify the statistical features with less effort.

3.2. Experimental Details

We studied two primary approaches that are used to create defects on the bearings for bearing fault analysis. In the first one, we run the bearing until it wears out naturally and over time observe the changes in its vibration signals. It shows results, which are close to real life conditions but this whole process takes a long period of time and also it is not easy to control for experimental operations. The second approach is widely common method of purposefully createsfaults. For this we can use methods like acid etching, spark erosion, etc. These can give precise control on size, shape and the location of the defects which helps to study changes in vibrations under different fault conditions. In this case, we studied about wire cut electro discharge machining (EDM) from earlier researches and implement it to make a defect that simulate real conditions on outer race of the bearing. It gives high precision, reliability and repeatable defect size. For this study, we created a defect of 0.5 mm width and 1 mm depth, along with healthy bearings, to study their changes in vibration response under different controlled conditions. The experimental parameters like fault-type, speed, and applied load are mentioned in Table 1 below.

Table 1. Defect Specifications and Operating Conditions.

Sr. No.	Parameter	Value
1.	Fault Types	Healthy, Outer Race Defect (0.5 mm)
2.	Speeds	1200 RPM, 1500 RPM
3.	Load and Shaft Alignment	0 Load and 0 tilt

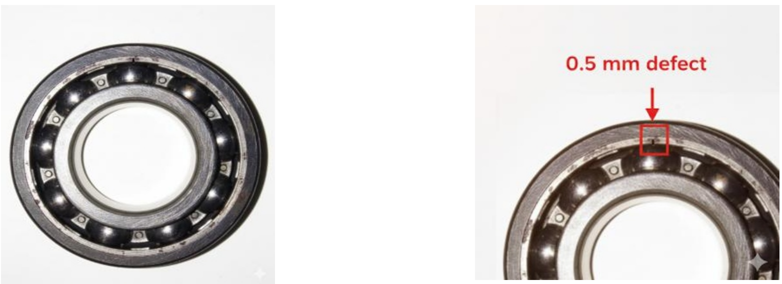

Fig. 1. Outer race fault in the SKF6206 Bearing with a 0.5 mm defect.

In this study, vibration data was collected with the help of a laboratory test rig designed to study bearing faults in rotating machinery. The setup was fixed on a steel frame that was anchored to a concrete base to maintain stability and reduce the effect of outside vibrations. The main parts of the rig included a bearing mounted on a rotating shaft, a DC motor with speed control, and a data acquisition system. The shaft was supported on two pedestals and connected to the motor using a flexible coupling. This arrangement helped in smooth rotation and reduced any misalignment. The motor speed could be adjusted, and experiments were carried out at 1200 RPM and 1500 RPM (common in small industrial motors) under no-load conditions with proper alignment so that the effect of defects could be clearly observed and with such condition of zero load and zero tilt we can study the effect of the defect clearly. To record vibration signals, a high-sensitivity tri-axial piezoelectric accelerometer was mounted in the radial direction on the bearing pedestal. The signals were captured through a fast Fourier transform (FFT) analyzer and dedicated software, with a sampling frequency of 2560 Hz. Data was taken for 1.6 seconds after the rig had been running for about 15 minutes to allow stabilization. The bearings used were 6206 deep groove ball bearings, with their dimensions listed in Table II. Artificial defects were created on the outer race using Wire Cut Electro Discharge Machining (EDM). Each defect measured 0.5 mm in width and 1 mm in depth. Along with testing healthy bearings, this setup gave a stable and controlled environment to observe vibration changes and study how early-stage faults start showing up in the signal.

Table 2. Geometric Parameters of the 6206 Ball Bearing.

Sr. No.	Parameters	Value
1.	Bore Diameter (d)	30 mm
2.	Outside Diameter (D)	62 mm
3.	Shoulder Diameter of Outer Ring (D1)	52.08 mm
4.	Raceway Diameter of Inner Ring (F)	37.5 mm
5.	Pitch Diameter (dp)	46.48 mm
6.	Bearing Width (B)	16
7.	Ball Diameter (dr)	9.5 mm
8.	Number of Balls	9
9.	Defect Location	Outer Race
10.	Defect Size	0.5 mm (width) × 1 mm (depth)
11.	Speeds	1200 RPM, 1500 RPM
12.	Net Weight	0.19 kg
13.	Contact Angle	0 deg.

Defects in bearings, such as those on the outer race, produce unique vibration patterns that depend on the bearing design and its operating speed. These patterns can be used to identify both the type and position of the fault. While advanced techniques often study these patterns in the frequency domain, the present work focuses on a simpler approach. Here, statistical features are extracted from the vibration signals to detect faults. This makes the process effective and easier to apply without the need for complicated models or heavy calculations.

3.2.1. Processing Steps

Before extracting the features, the vibration signal was preprocessed so that all values stayed in a similar range and did not create imbalance during classification. First, the raw data was lightly cleaned using a basic moving–average type smoothing to reduce small random noise. After that, standard scaling was applied to normalize the values, because features like RMS usually have higher magnitude compared to features such as skewness. Keeping everything on a common scale makes the comparison fair. Since the recorded duration was already short, the whole signal was taken as one segment without breaking it further. These simple steps helped the model behave more consistently and kept the classification process stable.

3.2.2. Statistical Features:

In this work, vibration signals from both healthy and faulty bearings were collected and then converted into useful values through feature extraction. When a defect, like an outer race fault, is present, these features start showing clear changes in the vibration pattern. The features we selected are simple to calculate, but they still capture the important variations in the signal, which makes them effective for identifying faults when used with machine learning methods. Each feature highlights a different part of how the signal behaves, and combining all of them together helps in getting a more accurate and dependable classification of the bearing condition.

i. Root Mean Square (RMS)

The RMS value measures nothing but the overall energy present in the vibration signal. It reflects the average power of fluctuations and is typically higher in faulty bearings due to increased vibration caused by defects. By comparing RMS values from healthy and defective bearings, abnormal behavior can be detected and the severity of faults can be assessed.

ii. Kurtosis

Kurtosis measures the sharpness or peaked-ness of the vibration signal’s distribution. A higher kurtosis value indicates the presence of sudden shocks or spikes, which are common in faulty bearings. By observing kurtosis, it is possible to distinguish between normal operation and faults that cause abrupt changes in vibration patterns.

iii. Skewness

Skewness quantifies the asymmetry of the vibration signal’s distribution. Significant skewness suggests that the vibration is biased toward either high or low values, which can result from uneven loading or defect-induced impacts. This feature helps identify fault signatures associated with irregular vibration patterns.

iv. Crest Factor

The ratio of the peak value of the vibration signal to its RMS value is called crest factor. For finding sudden faults crest factor is a useful feature. Defects like cracks or pits cause sharp impacts, which make the peak values high and increasing this ratio crest factor helps to detect fault at an early stage.

v. Peak Factor

The ratio of the peak value to the mean value of the signal is called peak factor. The highest vibration compared to the average behavior is indicated by peak factor. Defective bearings frequently experience sudden spikes much higher than the average vibration level.

vi. Peak Value

The value shows the highest amplitude recorded in the vibration signal over a sampling period called Peak factor. Fault produce impacts frequently cause sudden, sharp increases in vibration, which are captured by this feature. Despite its simplicity, it is used as a quick and effective indicator of abnormalities in bearing operation.

These features give a whole view of the vibration signal. Energy and sudden changes are shown by RMS and peak values, whereas signal shapes are described by kurtosis and skewness. Crest factor and peak factor mainly capture the sharp spikes that appear when a fault suddenly hits the bearing surface. Because of this, they can support machine-learning models like SVM in separating healthy and faulty cases. By paying attention to the features that actually show clear changes in the signal, the overall fault detection process becomes easier, more accurate, and more helpful for predictive maintenance work.

Table 3. Feature Calculation Formulae.

Sr. No.	Feature	Formulae
1.	Root Mean Square (RMS)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
2.	Kurtosis	$\frac{1}{N}\sum_{i=1}^{N} \left(\frac{x_i - \underline{x}}{\sigma}\right)^4$
3.	Skewness	$\frac{1}{N}\sum_{i=1}^{N} \left(\frac{x_i - \underline{x}}{\sigma}\right)^3$
4.	Crest Factor	$\frac{Peak\ Value}{RMS}$
5.	Peak Factor	$\frac{Peak\ Value}{\underline{x}}$
6.	Peak Value	$max(x_i)$

- x_i is each individual sample of the vibration signal
- N is the total number of samples
- $\underline{x}$ is the mean of the signal
- σ is the standard deviation of the signal

3.3. Model Development and Fault Classification Using SVM

3.3.1. Processing Steps

We used Support Vector Machine (SVM) in our study because finding faults in bearings is not easy—changes in vibration signals are very small and can be hidden by noise. SVM performs well even with small datasets. It can clearly separate healthy and faulty conditions and reduce mistakes from random changes in the signal. It also handles new data well. This helps in maintenance, as early detection of faults can prevent machines from breaking down unexpectedly. Since our vibration data came from limited operating conditions, SVM was a good choice to correctly identify both healthy and faulty states.

3.3.2 Principle of SVM Operation

To distinguish between normal and defective bearings, SVM works by learning from labeled vibration data. The hyperplane is a decision boundary constructed by SVM in a space where each axis corresponds to one of the statistical features, such as RMS, kurtosis, or crest factor. The hyperplane is chosen so that the gap between the two classes is as wide as possible, ensuring reliable predictions. The Radial Basis Function (RBF) kernel was applied in this study. The RBF kernel maps the features into a higher-dimensional space, which helps separate patterns that are otherwise difficult to distinguish in their original form.

3.3.3 Comparison of SVM with Other Machine Learning Models

In many earlier studies, models like Random Forest, Artificial Neural Networks (ANN), and k-Nearest Neighbors (k-NN) were also used for bearing fault detection. These models can work well, but they usually need more data and more complicated features to give stable and accurate results. ANN and deep learning models take a long time to train and require a large amount of input data, which was not available in this experiment. Random Forest performs nicely when many features are used, but its performance drops when the feature set is small like in our case. The k-NN method is simple, but it gets easily affected by noise and changes in speed or load. Compared to these methods, SVM works better when the dataset is small and the features are limited. It still makes a clear separation between healthy and faulty signals. Because of these reasons, SVM is more suitable for the type of data used in this study and gives more dependable results.

3.3.4 KNN classifier and Performance evaluation

To compare SVM with another classifier, k-Nearest Neighbor (k-NN) was also considered. k-NN is simple and easy to apply, but it is sensitive to noise and changes in speed. Including k-NN in the methodology shows how different classifiers behave under the same feature set and supports the reviewer's request for model comparison. The final performance was analyzed using accuracy, precision, recall, F1-score and confusion matrix these matrices help understand how good each feature is in detecting the defect.

3.3.5 Model Training and Parameter Tuning

The dataset was split into two parts such as training and testing sets. In the training set, the model learns the patterns by finding the patterns and then in the testing set, the model is tested over new data. Then normalization was done before evaluation, so that no feature can dominate others. For SVM, two key parameters we tuned that are penalty Factor (C) and Kernel Width (γ). This helps to avoid overfitting because overfitting can reduce accuracy after testing it on new data. To make sure reliability, cross-validation was done. In which model is tested on different parts of datasets to ensure stable accuracy. In our study, SVM with an RBF kernel was used because it handles small datasets well and can separate non-linear boundaries created by minor defects.

4. Metrics for Performance Evaluation

SVM was used to test different statistical features to evaluate the classification performance for fault detection in deep groove ball bearings. The performance was measured using different metrics, which helped to observe the accuracy, reliability, and usefulness of each feature in detecting bearing faults.

4.1 Accuracy, Precision, Recall and F-1 score

4.1.1. Accuracy

Overall performance of the model is called accuracy. It is the ratio of the sum of both true positive and true negative values to the total number of values.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

4.1.2. Precision

The value which shows how many of the predicted positive results are actually correct is called precision. It mainly focuses on the quality of positive predictions.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

4.1.3. Recall

It is also called sensitivity. The number of actual positive cases that were accurately predicted is known as recall, or sensitivity. Its main goal is to avoid overlooking the good cases.

$$Recall\ (Sensitivity) = \frac{TP}{TP+FN} \tag{3}$$

4.1.4.F1-Score

The F1-score is calculated by taking the harmonic mean of recall and precision. When we must take into account both, it provides a balanced measure.

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

In the above equations, TP denotes true positives, TN denotes true negatives, FP denotes false positives, and FN denotes false negatives, representing correct and incorrect classifications of bearing conditions.

4.2 Performance Confusion Matrix

By using a confusion matrix, we evaluate and compare the effectiveness of different features in classification. It provides the number of TP, TN, FP, and FN, which helps us understand how well each feature can differentiate between classes and where misclassifications occur. By analyzing these values, we can identify the most informative features and improve overall model performance

Table 4. Performance Matrix for Statistical Features.

Sr.No	Feature	TP	TN	FP	FN
1	RMS	3	8	0	0
2	Krutosis	27	47	0	2
3	Peak Factor	21	42	9	0
4	Peak Value	12	28	3	9
5	Skewness	0	54	18	0
6	Crest Factor	0	51	0	21

5.Results and Discussions

5.1 Statistical Feature Performance Comparison

From the results, RMS is the most effective feature, with accuracy 98%, precision 100%, recall 100%, and F1-score 100%, which shows that the model almost never makes mistakes using RMS. Kurtosis also performs well, with accuracy 97.37%, precision 100%, recall 93.10%, and F1-score 96.43%, missing only a few positive cases. Peak Factor and Peak Value give moderate results. Peak Factor has accuracy 87%, precision 70%, recall 100%, and F1-score 82.4%, meaning it identifies all positive cases but includes some false positives. Peak Value shows accuracy 83.33%, precision 80%, recall 57.14%, and F1-score 66.67%, so it misses several positive cases. Skewness (accuracy 75%) and Crest Factor (accuracy 70.83%) have precision, recall, and F1-score of 0%, as the model could not correctly identify any positive cases. This happens because these features do not provide enough information to separate the positive and negative classes, making the model unable to detect any true positives.

Table 5. Performance Result of Different Statistical.

Sr.No	**Feature**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F-1 Score (%)**
1.	RMS	98	100	100	100
2	Krutosis	97.37	100	93.10	96.43
3	Peak factor	87	70	100	82.4
4	Peak value	83.33	80	57.14	66.67
5	Skewness	75	0	0	0
6	Crest factor	70.83	0	0	0

Skewness and crest factor did not give any useful results in this study. This is mostly because these two features are not good for detecting any type of impulsive fault used here. The 0.5 mm outer race defect produces short and repeated impact impulses that increase the kurtosis value. But these impacts do not shift the signal to one side. Skewness measures only asymmetry not the strength of impact so it says close to zero even if the fault is present. Crest Factor did not help either, mainly because it reacts too much to random noise. In real vibration signals, noise spikes can produce high peak amplitudes even in healthy bearings, causing overlapping crest factor values between healthy and faulty conditions. Similar behavior has been reported in literature, where crest factor becomes unreliable for small localized defects or low-energy impacts because peak amplitudes do not rise significantly unless the defect is severe (Fasana et al., 2010). In our experiment, the small EDM outer race fault generated repeated impacts, but the energy of each impact was low, so the peak/RMS ratio did not change enough to form a separable class boundary for SVM. These theoretical limitations explain why both features failed in classification despite being commonly used in some early fault-detection studies. In summary, RMS and Kurtosis are the most reliable features, Peak Factor and Peak Value are somewhat useful, and Skewness and Crest Factor have very limited effectiveness due to their inability to capture discriminative information. To understand the results more clearly, the TP, TN, FP, and FN values from Table IV were also checked. These values help in seeing how each feature behaves inside the confusion matrix. RMS and Kurtosis have high true-positive and true-negative counts, with almost no wrong predictions, which matches their high precision and recall. Peak Factor, although able to catch all faulty cases, also produced a few false positives, which explains why its precision is lower. Peak Value shows both false positives and false negatives, so its overall F1-score remains in the middle range. On the other hand, Skewness and Crest Factor show zero true-positive values, which means the SVM model could not detect any faulty sample when these features were used alone. This directly supports their zero recall and zero F1-score. Overall, when we look at accuracy, precision, recall, F1-score, and the confusion matrix together, it becomes clear that RMS and Kurtosis are the strongest features for separating healthy and faulty bearings in this experiment.

5.2 Comparison of Feature Performance Using SVM

5.2.1 RMS: With nearly flawless scores on every metric, Root Mean Square (RMS) emerged as the SVM model's most potent and dependable feature. It is an excellent predictor for this task because of its flawless precision, recall, and F1-score, which show that it can distinguish between the classes with remarkable accuracy and without producing any false positives or negatives.

*5.2.2 Kurtosis:*Kurtosis also showed very strong performance. It reached a perfect precision score of 100% and gave the second-highest accuracy overall. Its recall and F1-score were also high, which shows that kurtosis is a very informative and reliable feature for SVM-based classification.

5.2.3 Peak Factor and Peak Value: It gave mixed results. Peak Factor managed to catch all the faulty cases, so its recall was perfect, but its precision was lower because it also marked some healthy samples as faulty. Peak Value, on the other hand, had more trouble with recall and missed a good number of faulty cases, which affected its F1-score.

5.2.4 Skewness and Crest Factor: It did not provide any useful information for the SVM model in this experiment. Both features resulted in zero precision, recall, and F1-score. This clearly shows that the SVM decision boundary formed using these two features could not separate the faulty samples from the healthy ones, making them ineffective for this type of classification.

5.3 Comparison with Other Machine Learning Classifiers

Here we use only the SVM model. Still its use can be explained when compared with other classifiers mentioned in earlier studies. Many studies noted that models like ANN and CNN require big training data and it takes more time for computation. In our case the data set is small, so they are not really practical. Random forest performs well when many statistical features, but here basic statistical features are available that's why its performance becomes limited. In some studies, KNN shows high accuracy but it is highly reactive to noise and also in different operation conditions. On the other hand,SVM works well with a small data set, The RBF kernel also helps it separate the healthy and faulty signals which are required for this study. Due to this using SVM is the most stable option for this faulty detection study.

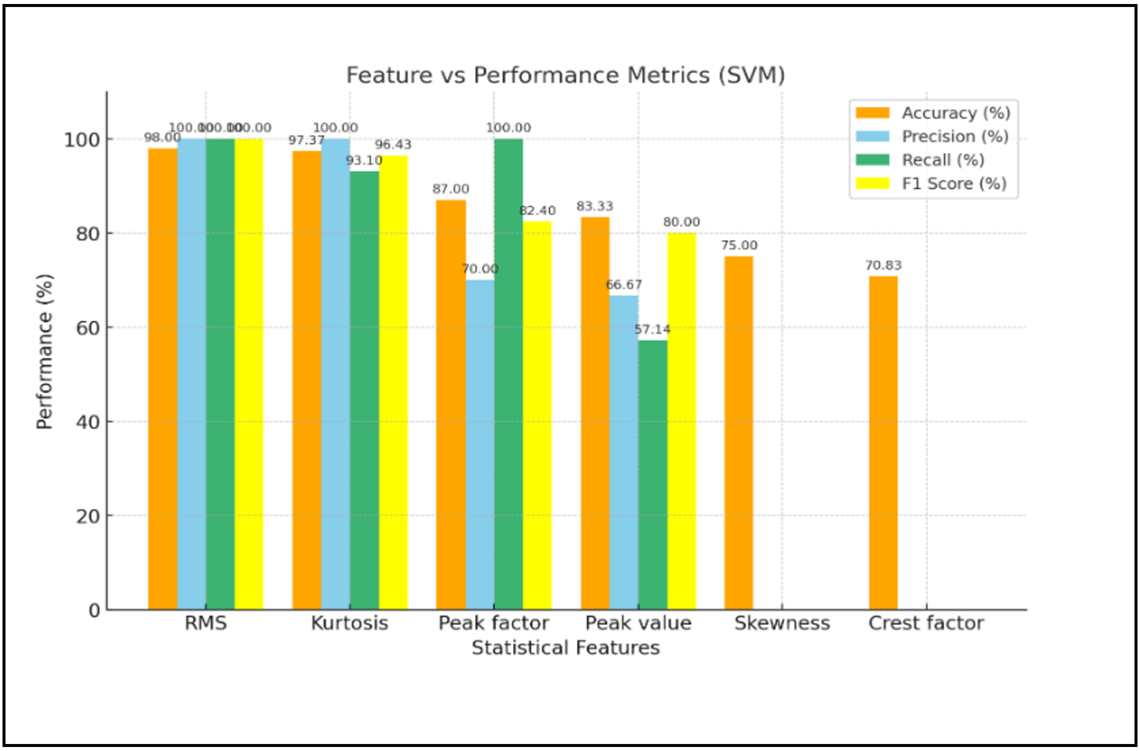

Fig. 2. Graphical comparison of SVM Performance metrics across different features.

6.Conclusion and Future Work

6.1 Conclusion

From this study, we understood that how strongly an SVM model depends on the type of features we select. In the current case, RMS gave the most accurate results and always separated the healthy and faulty bearings correctly. Kurtosis also performed well and made only a few mistakes. Peak Value and Peak Factor gave us the average results; Peak Value missed some faulty samples, while Peak Factor detected all faults but also raised a few wrong alarms. Skewness and Crest Factor did not work in this experiment because our model failed to classify even a single faulty case using these two features. So, these observations clearly show that choosing proper features is one of the most important steps for getting good accuracy in machine-learning-based fault detection. It is also important to note that this experiment was carried out in a fully controlled setup. The defect size, speed, and running conditions were fixed throughout the testing. Because of this, the strong performance of RMS and Kurtosis in this experiment does not guarantee the same results in real industrial machines. In industries, machines operate with changing loads, different speeds, and a lot of background noise. Sometimes more than one type of fault may occur at the same time, which makes the vibration signals much more complicated. In such situations, depending on just one or two features will not give a stable or reliable result. Instead, a real-world condition-monitoring system would need a combination of multiple features or a broader set of vibration indicators. Overall, RMS and Kurtosis performed well for the particular experimental conditions used in this project, but for a complete practical system, a multi-feature approach and further detailed study would be required in the future.

6.2 Future Work

In future work, it would be useful to check whether combining RMS and Kurtosis can improve the model's performance even further. It may also be worth exploring other machine-learning models like Random Forest or simple Neural Networks to see if weaker features, such as Peak Value, can still give better results with different algorithms. Another helpful direction would be to study why Skewness and Crest Factor performed poorly in this experiment. Understanding their limitations could make feature selection stronger and help in designing better models for similar classification problems.

References

1. Lai, Lianyou, Weijian Xu, and Zhongzhe Song. "A Novel Fault Diagnosis Method for Rolling Bearings Based on Spectral Kurtosis and LS-SVM." *Electronics* 14, no. 14 (2025): 2790. https://doi.org/10.3390/electronics14142790.
2. Imane, Moussaoui, Chemseddine Rahmoune, Moahmed Zair, and Djamel Benazzouz. "Multi-Fault Bearing Diagnosis under Time-Varying Conditions Using Empirical Wavelet Transform, Gaussian Mixture Model, and Random Forest Classifier." *Advances in Mechanical Engineering* 16, no. 8 (2024): 16878132241275787. https://doi.org/10.1177/16878132241275787.
3. Salunkhe, Vishal G., and R. G. Desavale. "An Intelligent Prediction for Detecting Bearing Vibration Characteristics Using a Machine Learning Model." *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems* 4, no. 3 (2021): 031004. https://doi.org/10.1115/1.4049938
4. Jamil et al. (2021) evaluated the performance of SVM and KNN classifiers in diagnosing faults in rolling element bearings using feature-based approaches. Their study, published in *Vibroengineering Procedia*, demonstrated how these classifiers can effectively distinguish fault conditions based on extracted features. https://doi.org/10.21595/vp.2021.22307.
5. Shen, Weijie, Maohua Xiao, Zhenyu Wang, and Xinmin Song. "Rolling Bearing Fault Diagnosis Based on Support Vector Machine Optimized by Improved Grey Wolf Algorithm." *Sensors* 23, no. 14 (2023): 6645. https://doi.org/10.3390/s23146645.
6. Wang, Tianhao, Hongying Meng, Fan Zhang, and Rui Qin. "Fault Detection of Wheelset Bearings through Vibration-Sound Fusion Data Based on Grey Wolf Optimizer and Support Vector Machine." *Technologies* 12, no. 9 (2024): 144. https://doi.org/10.3390/technologies12090144.
7. Zheng, Hong, and Lei Zhou. "Rolling Element Bearing Fault Diagnosis Based on Support Vector Machine." *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, IEEE, April 2012, 544–47. https://doi.org/10.1109/CECNet.2012.6201982.

8. Fasana, Alessandro, Stefano Marchesiello, Miriam Pirra, Luigi Garibaldi, and Alessandra Torri. “Spectral Kurtosis against SVM for Best Frequency Selection in Bearing Diagnostics.” *Mécanique& Industries* 11, no. 6 (2010): 489–94. https://doi.org/10.1051/meca/2010056.
9. Zhu, Ke Heng, Xi Geng Song, and Dong Xin Xue. “Roller Bearing Fault Diagnosis Based on IMF Kurtosis and SVM.” *Advanced Materials Research* 694–697 (May 2013): 1160–66. https://doi.org/10.4028/www.scientific.net/AMR.694-697.1160
10. Gangsar, Purushottam, and Rajiv Tiwari. “Online Diagnostics of Mechanical and Electrical Faults in Induction Motor Using Multiclass Support Vector Machine Algorithms Based on Frequency Domain Vibration and Current Signals.” *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering* 5, no. 3 (2019): 031001. https://doi.org/10.1115/1.4043268.
11. Pule, Mushabi, OduetseMatsebe, and Ravi Samikannu. “Application of PCA and SVM in Fault Detection and Diagnosis of Bearings with Varying Speed.” *Mathematical Problems in Engineering* 2022 (April 2022): 1–12. https://doi.org/10.1155/2022/5266054.
12. Kannan, Vigneshwar, Tieling Zhang, and Huaizhong Li. “A Review of the Intelligent Condition Monitoring of Rolling Element Bearings.” *Machines* 12, no. 7 (2024): 484. https://doi.org/10.3390/machines12070484.