INTERNATIONAL JOURNAL OF AUTOMOTIVE AND MECHANICAL ENGINEERING
ISSN: 2229-8649 e-ISSN: 2180-1606
VOL. 20, ISSUE 2, 2023, 10371 – 10385
DOI: https://doi.org/10.15282/ijame.20.2.2023.04.0802

RESEARCH ARTICLE

Bayes' Theorem for Multi-Bearing Faults Diagnosis

Yeo Siang Chuan*, Lim Meng Hee, Hui Kar Hoou and Eng Hoe Cheng

Institute of Noise and Vibration, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia

ABSTRACT - During the process of fault diagnosis for automated machinery, support vector machines is one of the suitable choices to categorize multiple faults for machinery. Regardless of the volume of sampling data, support vector machines can handle a high number of input features. It was learned that support vector machines could only sense binary fault classification (such as faulty or healthy). However, the classification accuracy was found to be lower when using support vector machines to diagnose multi-bearing faults classifications. This is because the multiple classification problem will be reduced into several sub-problems of binary classification when support vector machines adapt to multi-bearing faults classifications. From there, many contradictory results will occur from every support vector machine model. In order to solve the situation, the combination of Support Vector Machines and Bayes' Theorem is introduced to every single support vector machine model to overcome the conflicting results. This method will also increase classification accuracy. The proposed Support Vector Machines - Bayes' Theorem method has resulted in an increase in the accuracy of the fault diagnosis model. The analysis result has shown an accuracy from 72% to 95%. It proved that Support Vector Machines - Bayes' Theorem continuously eliminates and refines conflicting results from the original support vector machine model. Compared to the existing support vector machine, the proposed Support Vector Machines - Bayes' Theorem has proven its effectiveness in diagnosing the multi-bearing faults problem classification.

ARTICLE HISTORY

Received	:	12th May 2022
Revised	:	29th April 2023
Accepted	:	29th May 2023
Published	:	30th June 2023

KEYWORDS

Support vector machines;
Bayes' Theorem;
Classification;
Vibration;
Multi-bearing fault

1.0 INTRODUCTION

Bearings are the main mechanical components that ensure the integrity of rotating machinery. Costly downtime and total machine breakdown could happen when bearing fault occurs. Thus, bearing fault diagnoses are being advanced rapidly in the past decades. Bearing fault diagnosis methods that were developed includes analysis of vibration [1], interpretation of thermal imaging [2] and analysis of acoustic [3]. Analysis of vibration spectra [4] is proven to be very convincing in diagnostic and monitoring for rotating machinery. Besides, a few tools for processing vibration signals, such as the Hilbert-Huang transform, empirical mode decomposition and wavelet analysis, have been introduced. As according to Hui et al. [5], non-adaptive signal processing methods have advanced to self-adaptive signal processing methods. According to Cui et al. [6], qualitative vibration analysis has advanced to quantitative vibration analysis. However, the knowledge and experience of the machine operator will affect the effectiveness of these diagnostic methods.

In current years, it has been noticed the rise of the artificial intelligence (AI) approach used for machinery fault diagnosis. With the planned artificial intelligence structure, the result of diagnosis can be more consistent. Furthermore, an advanced system for fault diagnosis can be established that will help to eliminate human intervention. The relationship between the sensors captured data (which served as inputs) and the machine's condition (which served as outputs can be established by an artificial intelligence algorithm. From there, the output can be provided by a trained artificial intelligence algorithm. The accuracy of artificial intelligence approach fault diagnosis machinery is highly dependent on the artificial intelligence algorithm applied to analyze the input data, although this diagnosis can provide more consistent results. This means the diagnostic accuracy based on the Hidden Markov Model (HMM), support vector machines (SVMs), self-organizing maps (SOMs), artificial neural networks (ANNs), and so on could be very different.

Yan et al. [7,8,9] and Kaisi et al. [10] have conducted extensive research in this particular domain. In the course of their investigations, they have examined a range of models, among which is the innovative multiscale cascading deep belief network (MCDBN) for identifying the location of faults in rotating machinery. Additionally, they have introduced a unique method called multi-domain indicator-based optimized stacked denoising autoencoder, which enables automatic fault identification in rolling bearings. Furthermore, they have proposed a novel hybrid deep learning model designed for multistep forecasting of diurnal wind speed. All of these contributions demonstrate the depth and breadth of their expertise in this field.

Based on studies from Zhang et al.[11], Jedliński and Jonak [12] and Kankar et al. [13], regardless of the volume of sampling data, support vector machines can handle a high number of input features. However, there is a limitation for support vector machines. It was designed to classify binary problems. Research from Jegadeeshwaran and Sugumaran [14] mentioned from fault diagnosis of automobile hydraulic brake systems, the multi-bearing faults classification is

***CORRESPONDING AUTHOR | Yeo Siang Chuan** | ✉ siangchuan.yeo@graduate.utm.my

reduced into a multi-layer binary classification. However, the concern will be if the data was wrongly classified in the first layer, it will continue to be wrong in the next layer. This happens due to the architectural design of the decision tree. Another research from Keskes et al. [15] mentioned from fault diagnosis of rotor bar condition, the multi-bearing faults classification is reduced into binary classification, which is healthy or faulty only. From there, the rotor bar condition can be classified as one or two broken rotor bars. This brings to our attention that what if the broken rotor bars have more than two levels of severity?

Previous studies have reported support vector machines for multi-bearing faults classification have different strategies, namely directed acyclic graphs, error-correcting output code, binary tree, one vs all and one vs one [16]. With all the strategies introduced, the popular strategies are one vs one [17] and one vs all [18, 19]. During the process of multi-bearing faults classification, contradictory results may be provided from different support vector machine models. Therefore, the first result will be treated as the decision by the machine learning models without having the conflicting results refined. Bayes' Theorem is proposed in this paper to eliminate the conflicting results in order to increase multi-bearing faults classification accuracy. In this section, we explained why the artificial intelligence approach is necessary for automatic bearing fault diagnosis. In Section 2, we present the limitations of support vector machines multi-bearing faults classification. In Section 3, we introduce Bayes' Theorem for machinery fault diagnosis. In Section 4, we present the bearing data collection methods. In Section 5, we discuss the features used for faults classification. In Section 6, we compare the different strategies in support vector machine multi bearing faults classification performances. In Section 6 also, we evaluate the Support Vector Machine - Bayes' Theorem model. At last, we discuss and conclude effectiveness of the Support Vector Machine - Bayes' Theorem model.

2.0 LIMITATIONS OF SUPPORT VECTOR MACHINES FOR MULTI BEARING FAULTS CLASSIFICATION

According to Kankar et al.[13], support vector machine is a machine learning method that has proven to be useful for fault diagnosis. Based on studies from Cui et al. [20] and Wenbo Liu et al. [21], as a result of its remarkable success in the realms of text mining and fault diagnosis, support vector machine has emerged as a leading technology among machine learning methods, ultimately contributing to the rapid advancement of statistical learning. In fact, the support vector machine's tremendous achievements played a key role in catalyzing the development of the kernel method, which has since gained widespread popularity and applicability, largely due to the influence and promotion of the support vector machine. This dynamic interplay between the support vector machine and the kernel method underscores the pivotal role that support vector machine has played in shaping the landscape of modern machine learning.

As according to Baccarini et al.[19], support vector machine created a hyperplane to allocate most of the points (from the same class) on the same side. Meanwhile, the distance between the classes is maximized in this hyperplane. According to Konar and Chattopadhyay [22], vector w and scalar b determine the position of the hyperplane. Refer to Eq. (1) as below:

$$f(x) = \mathrm{w}^T x + b \tag{1}$$

Figure 1 presents an example of a support vector machine (with the involvement of Gaussian radial basis function (RBF) kernel) creating a hyperplane for faulty and healthy classes by kurtosis and skewness features. According to Hsu et al. [23], the RBF kernel function is proposed to be the first-try kernel function for a support vector machine model. And according to Chen et al.[24], the accuracy of the RBF kernel is found to be better than the polynomial kernel. In this paper, a few of support vector machine kernel functions, such as polynomial, quadratic and RBF will be used. Among the choices, the most reliable kernel function for bearing fault classification will be determined.

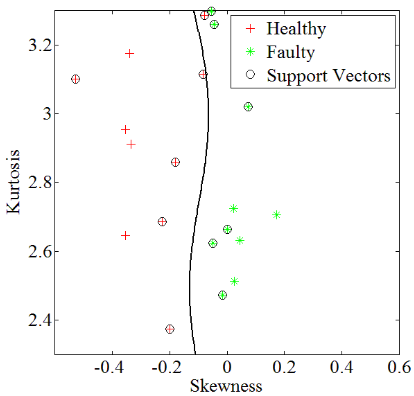


Figure 1. SVM (with involvement Gaussian radial basis function (RBF) kernel) create a hyperplane for faulty and healthy classes by kurtosis and skewness features

The development of a support vector machine is to have two classes of an issue that are to be classified by a few features. In this paper, we involve four types of bearing conditions, which are outer raceway fault, inner raceway fault, rolling element fault and healthy. One vs all and one vs one strategy are the two (2) multi-bearing faults classification strategies. Equation (2) shows the number of models for one vs one, and Eq. (3) indicates the number of models for one vs all. If one vs all is used, one training model is required.

$$\text{Models (for one vs one)} = \frac{\text{Number of Class} \times (\text{Number of Class} - 1)}{2} \tag{2}$$

$$\text{Model (for one vs all)} = \text{Number of Class} \tag{3}$$

A Support Vector Machine library was developed by Chang and Lin [25], named LIBSVM. From there, the one vs one strategy was implemented in multi-bearing faults classification. From each binary classification, a decision can be voted for any class. Then, which class has the highest number of votes, the final decision is made. However, no decision can be made if the highest number of votes happens in two classes or more. This is the drawback of this method. Chang and Lin were aware of the drawback; they chose the first class from all the identical classes. Table 1 shows examples of the conflicting results from one vs one strategy of support vector machine. Table 2 shows the examples of the conflicting results from the one vs all strategy of the support vector machine.

Table 1. One vs one strategy of support vector machine result example

		Sample		
		A	B	C
Votes (6 choices)	Class I	1	1	0
	Class II	2	1	0
	Class III	1	3	0
	Class IV	2	1	6
Decision		Conflict	Class III	Class IV

Table 2. One vs all strategy of support vector machine result example

		Sample		
		A	B	C
Votes	Class I	N	N	N
Yes = Y	Class II	Y	N	N
Or	Class III	N	Y	N
No = N	Class IV	Y	N	Y
Decision		Conflict	Class III	Class IV

In this paper, four types of bearing conditions, which are outer raceway fault, inner raceway fault, rolling element fault and healthy, are classified from the development of multiple one vs one and one vs all support vector machine models. Results from this study show that each of the support vector machine models may not be consistent, and the results could probably become contradicted. Hence, Bayes' Theorem is suggested to solve this concern. With Bayes' Theorem involvement in the support vector machine model, a single decision is resulted.

3.0 INTRODUCTION TO BAYES' THEOREM

Bayes' Theorem, named after the 18^{th}-century British mathematician Thomas Bayes, is a theorem in probability and statistics that helps in determining the probability of an event that is based on some event that has already occurred [26]. In another way, it is a mathematical formula to determine conditional probability. Conditional probability is the probability of an event happening, given that it has some relationship to one or more other events. In short, the conditional probability of event A, given that event B has already occurred, is determined by Bayes Theorem. The Bayes' Theorem is expressed in the following formula:

$$P(A | B) = \frac{P(B | A) \times P(A)}{P(B)} \tag{4}$$

where P(A|B) – the probability of event A occurring, given event B has occurred, P(B|A) – the probability of event B occurring, given event A has occurred, P(A) – the probability of event A, and P(B) – the probability of event B.

Bayes Theorem is the foundation of the Bayesian statistics field. It is also called Bayes Law or Bayes Rules. In the finance realm, the Bayes' Theorem is used to rate the risk of customers to borrow money from banks. Not only the finance realm, but Bayes Theorem is also famous for being used to determine the accuracy of medical test results. This is the process of determining the probability of any given person having an illness (such as cancer rate).

The applications of Bayes theorem can be found in the monitoring of machinery condition [26], fields of medical [27], human language [28], Psychological [29], ecological data [30] and so on. As of today, Bayes' Theorem has given confidence in the relevant event.

4.0 DATA COLLECTION

Case Western Reserve University Bearing Data Center is the website that is allowed to download data for this study. The downloaded data can represent faulty and healthy ball bearing conditions (outer raceway faults, inner raceway faults and rolling element). The test rig for the experiment comprised of a dynamometer, a torque transducer, and a 2 hp motor. The test rig will be arranged to simulate different bearing conditions. Figure 2 shows the test rig for the experiment. Fault diameters of 21 mils (0.53 mm) and 7 mils (0.18 mm) were used in the SKF bearing to have various bearing faults to be simulated in the laboratory. A motor load of 0 HP to 3 HP at 1797 to 1720 rpm was operated. Accelerometer is attached at the bearing housing in order to collect vibration data. Table 3 shows the vibration data collection conditions.

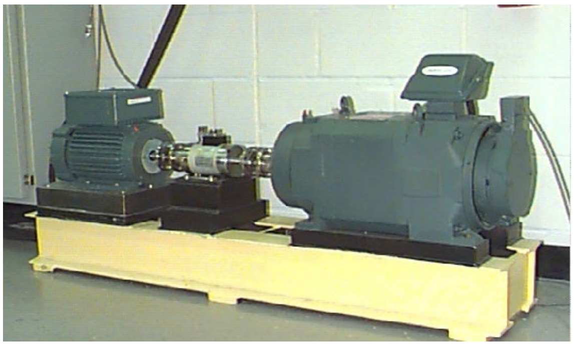


Figure 2. Experimental test rig

Table 3. The vibration data collection conditions

Bearing Conditions	Healthy	Faulty Outer raceway Inner raceway Rolling element
Fault Diameter (Severity)	0 mils (0 mm)	7 mils (0.18 mm) 21 mils (0.53 mm)
Motor Load (HP)	0 – 3	0 – 3

From the 7 mils fault diameter with a 1 HP load, the raw continuous signal collected can extract the 400 sets of time series vibrations. From there, the collected data were categorized into 2 data sets. 1 set of data was needed during the training phase. The purpose of the training phase is to ensure the connection between the input and output of the artificial intelligence model to be established. Another set of data was used during the testing phase. The purpose of the testing phase is to ensure the trained artificial intelligence model is validated. At last, a new set of testing data, which consists of all sorts of testing conditions, such as motor load and different severities, will then be validated by an artificial intelligence model. This last time, a total of 7000 sets of data, and every single bearing condition had 250 sets of data.

The vibration data distribution for the training phase is 50 training data for each bearing condition, namely Inner Raceway Fault, Outer Raceway Fault, Rolling Element Fault and Healthy. The same goes to the testing phase, 50 testing data will be distributed for each bearing condition, namely Inner Raceway Fault, Outer Raceway Fault, Rolling Element Fault and Healthy. Table 4 shows the vibration data distribution used in this study (CWRU bearing fault simulator).

Table 4. Vibration data distribution used in this study (CWRU bearing fault simulator)

Bearing Condition	Training data	Testing data
Healthy	50	50
Rolling element fault	50	50
Inner raceway fault	50	50
Outer raceway fault	50	50
Total number of data samples	200	200

5.0 THE EXTRACTIONS OF THE BEARING FAULT FEATURE

Bearing Fault Features are obtained from time series vibration data. Those features are, namely, margin, impulse, shape, crest, kurtosis and skewness. Statistical analysis was carried out. Next, support vector machine model training and testing purposes use statistical features as input features. Table 4 shows the Bearing Fault Feature equations.

Table 5. Bearing fault features

No.	Bearing fault feature and equation	No.	Bearing fault feature and equation
1	Impulse $\dfrac{\max\lvert x(n)\rvert}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$	4	Kurtosis $\dfrac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^4}{\left(\sqrt{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^2}\right)^4}$
2	Shape $\dfrac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}x(n)^2}}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$	5	Skewness $\dfrac{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^3}{\left(\sqrt{\frac{1}{N}\sum_{n=1}^{N}(x(n)-\bar{x})^2}\right)^3}$
3	Crest $\dfrac{\max\lvert x(n)\rvert}{\sqrt{\frac{1}{N}\sum_{n=1}^{N}x(n)^2}}$	6	Margin $\dfrac{\max\lvert x(n)\rvert}{\left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}\right)^2}$

Figure 3 shows the distribution of data for bearing fault features, namely margin, impulse, shape, crest, kurtosis and skewness. During the motor load condition of 7 mils fault diameter with a 1 HP, the vibration signals are collected. A total of 100 samples for each bearing condition were used. The 50 random samples were used to produce the artificial intelligence model, which functioned as training data. Another 50 random samples were used to validate the trained artificial intelligence model, which functioned as testing data.

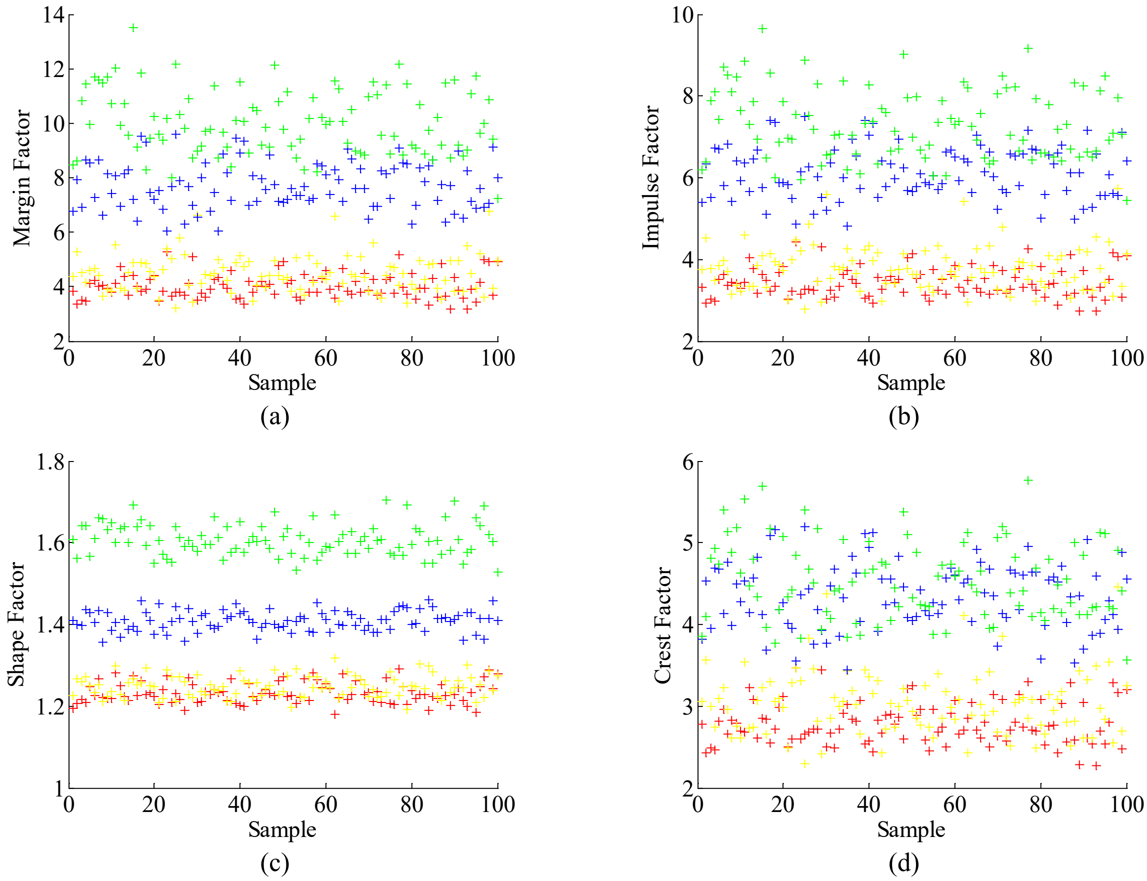
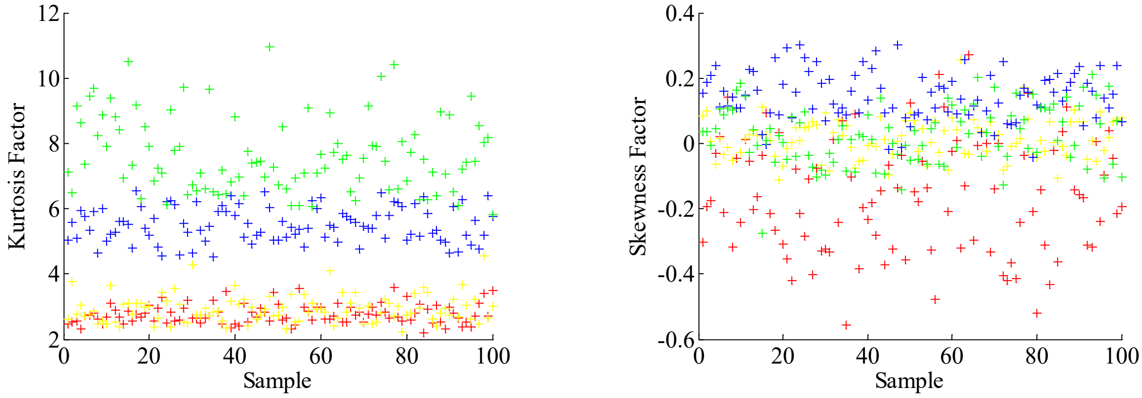


Figure 1. (a) Margin, (b) impulse, (c) shape, (d) crest, (green: outer raceway fault; blue: inner raceway fault; yellow: rolling element fault; red: healthy)



(e) (f)

Figure 2. (cont.) (e) kurtosis, and (f) skewness of bearing fault conditions (green: outer raceway fault; blue: inner raceway fault; yellow: rolling element fault; red: healthy)

6.0 RESULTS AND DISCUSSIONS

6.1 Support Vector Machines Multi Bearing Faults Classification

Support vector machine analysis created a hyperplane to separate the four bearing conditions. The three fault bearing conditions are outer raceway, inner raceway and rolling element. Then, one more bearing condition is healthy. In the following sections, we discuss the support vector machine results from different strategies.

6.2 Strategy: One Vs One

According to Chang and Lin [27], LIBSVM employed a one vs one strategy for multi-bearing faults classification. Six different one vs one support vector machine models are developed to separate six binary classes. The most accuracy among polynomial, quadratic and linear are demonstrated from the RBF kernel functions. This is consistent with Chen et al. [26]'s findings report, in which accuracy is found to be better at RBF kernel but less at polynomial kernel.

i. inner raceway vs outer raceway
ii. rolling element vs inner raceway
iii. healthy vs inner raceway
iv. rolling element vs outer raceway
v. healthy vs outer raceway
vi. healthy vs rolling element

The classification performance of a classifier is often to be measured by receiver operating characteristic (ROC) curves. This means all probable cut-off points are immediately measured their specificity and their sensitivity. The true positive rate represents the ratio of the sample that is correctly classified. The false positive rate represents the ratio of the sample that is incorrectly classified. Figure 4 shows receiver operating characteristic curves for the training phases and testing phases of the one vs one strategy support vector machine model. Four bearing conditions, namely outer raceway fault, inner raceway fault, rolling element fault and healthy, represent Class 1 to 4 in the plot. It showed superior performance of the support vector machine model during the training stage because all training receiver operating characteristic curves overlapped. During the training phase, we can notice the overlapping of the data for Class 1 and Class 2. During the testing phase, we then notice the Class 1 and Class 2 support vector machine model performance has decreased. It means some of the new data in the testing phase was not able to be classified by the trained support vector machine model.

The accuracy of a classifier in the training and testing phases is displayed in the confusion matrixes, as shown in Table 6. During the training phase, the support vector machine model showed a percentage of classification accuracy at 100%. From this support vector machine model, no data generation has conflicting results. During the testing phase, the support vector machine model showed a percentage of classification accuracy at 86%. This is because a different set of data was used compared to the data that was used in the training phase.

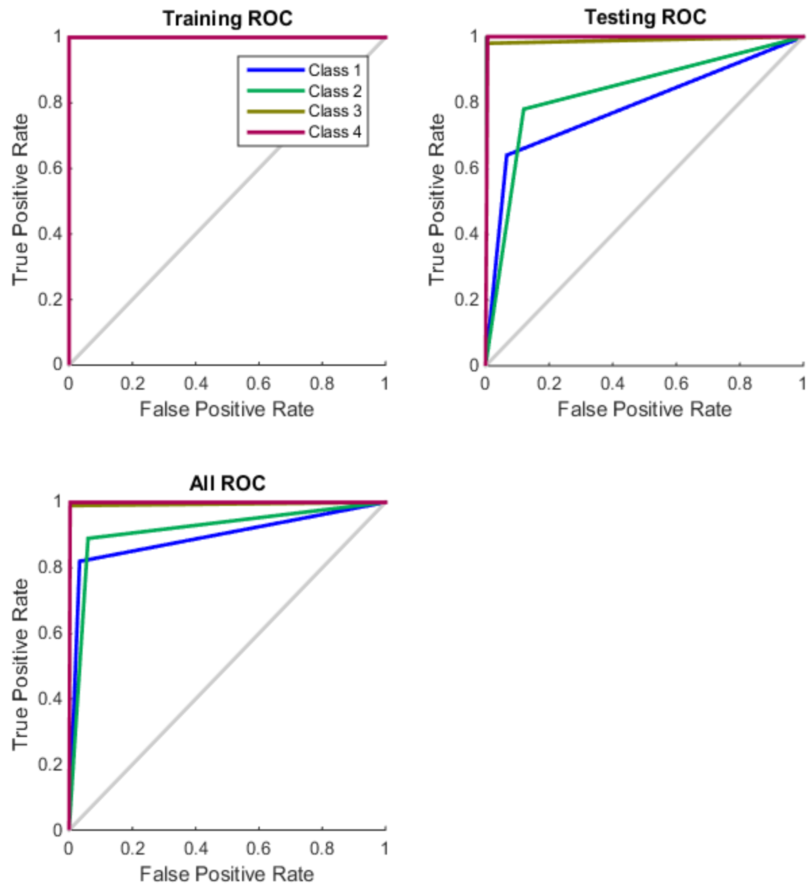


Figure. 3 Training phase and testing phases of Receiver Operating Characteristic (ROC)

Table 6. Training and testing phase confusion matrix

Class		Training actual result				Testing actual result			
		I	II	III	IV	I	II	III	IV
Predicted result	I	50	-	-	-	29	-	-	-
	II	-	50	-	-	15	45	2	-
	III	-	-	50	-	-	3	48	-
	IV	-	-	-	50	6	2	-	50
	Conflict	-	-	-	-	-	-	-	-
Sensitivity in %		100	100	100	100	58	90	96	100
Accuracy in %		100				86			

From the matrix shown in the above tables, we can notice that this support vector machine is unable to improve the performance of classification because no conflicting results to be refined. With this, we can conclude that the one vs one strategy support vector machine model in multi-bearing faults classification is not necessary to be continued in this study.

6.3 Strategy: One Vs All

The one vs all strategy in multi-bearing faults classification involves four support vector machine models. The desired class (for example, rolling element fault) will be separated from the other classes (such as outer raceway fault, inner raceway fault and healthy) from each hyperplane. In summary, a "yes" or a "no" will be the decision from the support vector machine structure. For example, the healthy bearing conditions will represent and generate the result of "Yes". The faulty bearing condition will represent and generate the result of "No" (could be any type of bearing fault). Next, the support vector machine structure will carry out the classification of each bearing fault against all the others. Therefore, in order to completely classify all four bearing conditions in this study, all four support vector machine models were designed for diagnosis. In similar conditions, the RBF kernel function performed better than other kernel functions in the one vs all strategy.

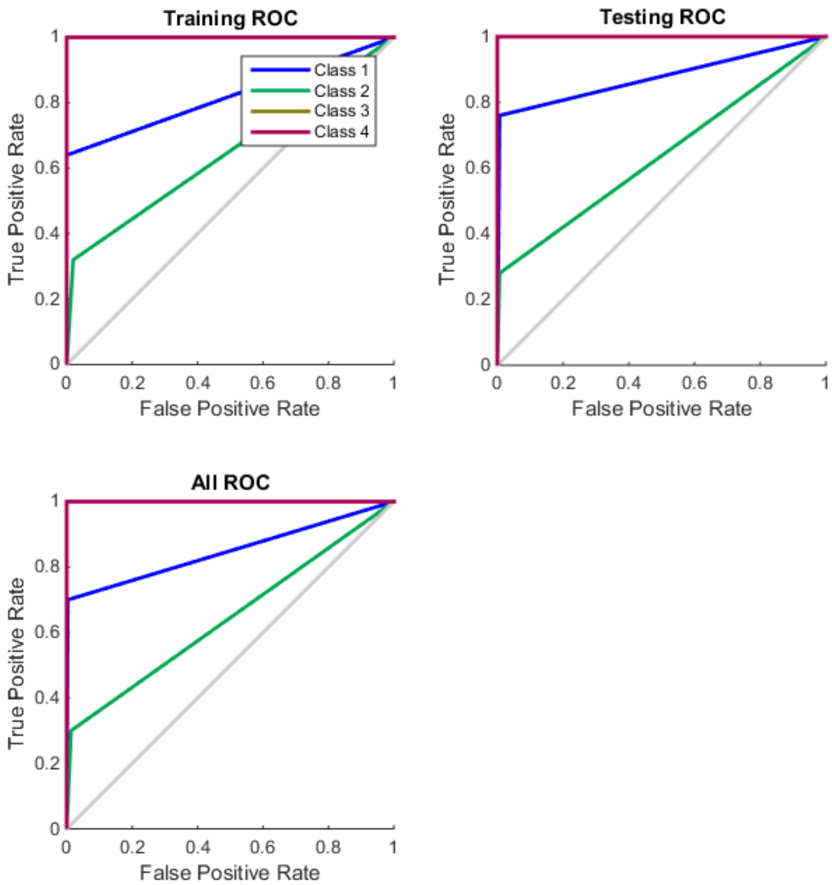


Figure. 4 Training phase and testing phases of receiver operating characteristic (ROC)

The support vector machine model (one vs all strategy)'s receiver operating characteristic curves for the training phases and testing phases are shown in Figure 5. From the receiver operating characteristic curves, the support vector machine model shows to have almost the same performances in the training phases and testing phases. Further to that, it was noticed the overlapping situation for the Class 3 curve and Class 4 curve. Besides, in both training phases and testing phases, the Class 3 curve and Class 4 curve are at the top and left edges of the plot. This means the classifiers performed good classification. According to the receiver operating characteristic plots, we can notice that during both training phases and testing phases, the one vs all strategy support vector machine model is more consistent than one vs one strategy support vector machine model in multi-bearing faults classification.

Table 7 shows the training and testing phase confusion matrix of the one vs all strategy support vector machine for multi-bearing faults classification. The accuracy of classification resulted in 70% in the training phase and 72% in the testing phase. This means the samples from the training phase and the sample from the testing phase have fitted well the support vector machine model's hyperplanes. However, we can notice quite a number of conflicting results generated from this support vector machine model. There is a total of 23% of conflicting results from testing data. From another perspective, this one vs all strategy can be improved further by having the conflicting results refined. Besides, the one vs all strategy could reduce the computing resources because it requires only a smaller number of support vector machine structures.

Table 7. Training and testing phase confusion matrix of one vs all strategy support vector machine

	Class	Actual result - training				Actual result - testing			
		I	II	III	IV	I	II	III	IV
Predicted Result	I	29	-	-	31	-	-	-	
	II	-	15	-	-	16	-	-	
	III	3	2	49	2	1	49	-	
	IV	1	-	-	1	-	-	48	
	Conflict	17	33	1	16	33	1	2	
Sensitivity in %		58	30	98	94	62	32	98	96
Accuracy in %		70.0				72.0			

6.4 Support Vector Machine Conflicting Results Analysis

The one vs all strategy in multi-bearing faults classification involves four support vector machine models. The desired class (for example, rolling element fault) will be separated from the other classes (such as outer raceway fault, inner raceway fault and healthy) from each hyperplane. In summary, a "yes" or a "no" is the decision from the support vector machine structure. For example, the healthy bearing conditions will represent and generate the result of "Yes". The faulty bearing condition represents and generates the result of "No" (could be any type of bearing fault). Next, the support vector machine structure will carry out the classification of each bearing fault against all the others. Therefore, in order to completely classify all four bearing conditions in this study, all four support vector machine models were designed for diagnosis. In similar conditions, the RBF kernel function performed better than other kernel functions in the one vs all strategy. Figure 6 shows the comparison of classification accuracy between one vs all strategy support vector machine models and one vs one strategy support vector machine models.

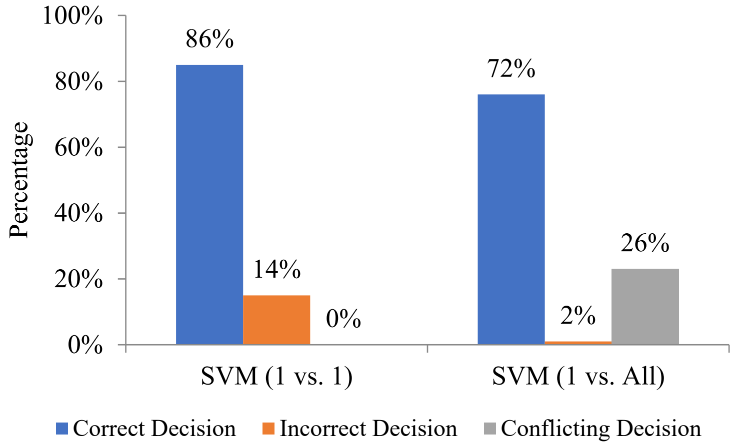


Figure 5. Testing phase - comparison of the classification accuracy for one vs all and one vs one strategy

The percentage of testing data misclassification with one vs one strategy support vector machine is 14%. The percentage of testing data misclassification with one vs all strategy support vector machine is 2%. The above comparison is conducted by excluding the indecisive results (conflicts). This means one vs one strategy has a higher misclassification percentage. This also means that we can improve the classification accuracy by eliminating conflicting results.

6.5 Fine-tuning the One vs All Classifier

The support vector machine model is trained and tested by using the "fitcsvm" and "predict" MATLAB functions. Adding to this, an RBF kernel function will be used too, because the RBF kernel function was found to be better classification accuracy. From MathWorks [31], Based on MATLAB User's Guide, when adjusting the "box constraint" and "Kernel scale" parameters, the support vector machine classifier can actually be tuned. The tuning process is conducted based on the classification error taken from cross-validation. Therefore, the one vs all classifiers in this study were tuned. Table 8 shows the values of "Box Constraint" and "Kernel Scale".

Table 8. The values of box constraint and kernel scale for support vector machine classifier

	S I	S II	S III	S IV
Kernel Scale	1.5580	1.4946	1.9362	4.5872
Box constraint	1.4798	4.3501	12.2898	6.0405

Figure 7 shows receiver operating characteristic curves of the tuned one vs all strategy support vector machine model for both the training phase and testing phase. Almost similar performances for training and testing were shown in receiver operating characteristic curves. During both training and testing phases, it showed the overlapping of the curves for Class 3 and 4. Both Class 3 curve and Class 4 curve are near to the top and left edges of the plot. This means the classification performance is better on these classifiers. This has proven that the tuned support vector machine model's performance is better than the support vector machine model without tuning for multi-bearing faults classification.

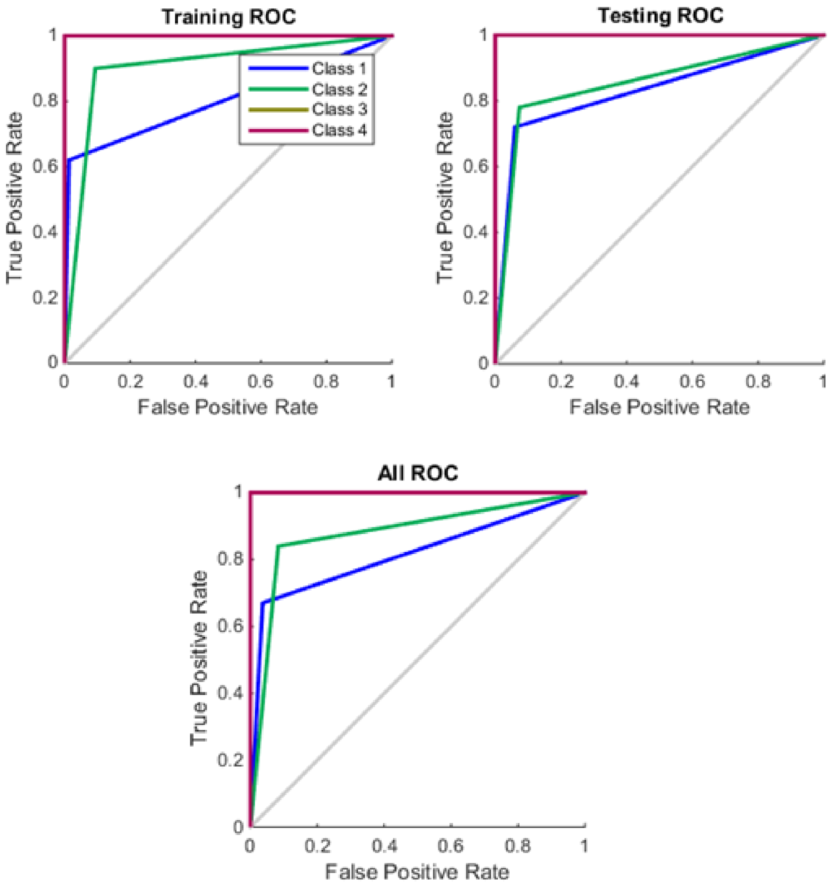


Figure 6. Training phase and testing phases of Receiver operating characteristic (ROC)

The accuracy of the tuned support vector machine classifiers is analyzed by using a confusion matrix. The confusion matrix for the training phase showed 90%, and the confusion matrix for the testing phase showed 89.5%. This indicates the training and testing samples fit the hyperplanes of the support vector machine model. Conflicting results generated by the support vector machine model were found to be a significant reduction. In the overall result, classification accuracy was improved in the tuned support vector machine model, and the conflicting result was decreased. Table 9 shows the training and testing phase confusion matrix for the tuned support vector machine model for multi-bearing faults classification.

Table 9. Training and testing phase confusion matrix for the tuned support vector machine model

Class		Actual result - training				Actual result - testing			
		I	II	III	IV	I	II	III	IV
Predicted Result	I	42	8	-	-	39	5	-	-
	II	-	40	-	-	-	41	-	-
	III	4	-	50	-	10	4	50	-
	IV	-	-	-	48	-	-	-	49
	Conflict	4	2	-	2	1	-	-	1
Sensitivity in %		84	80	100	96	78	82	100	98
Accuracy in %		90.0				89.5			

6.6 *Support Vector Machine – Bayes' Theorem for Multi-Bearing Faults Classification*

In this paper, we recommend a reliable method to eliminate conflicting results in order to increase multi-bearing faults classification accuracy. With the effort of combining the Support Vector Machine model and Bayes' Theorem, the automated bearing fault diagnosis model was developed. In this study, the support vector machine produces fault classification results. Then, Bayes' Theorem will further refine the results for the purpose of ultimate decision-making. Due to one vs all strategy support vector machine model multi-bearing faults classifications being proven to be better than other strategies, this strategy will continue to be used to have the bearing conditions classified. Four results were produced from four trained support vector machine models. In the beginning, the results from four support vector machine models were found contradicting to each other. The Bayes' Theorem was then triggered to refine the result, and one decisive conclusion arrived. Figure 8 shows the flowchart for the diagnosis of the automated bearing fault.

In-II

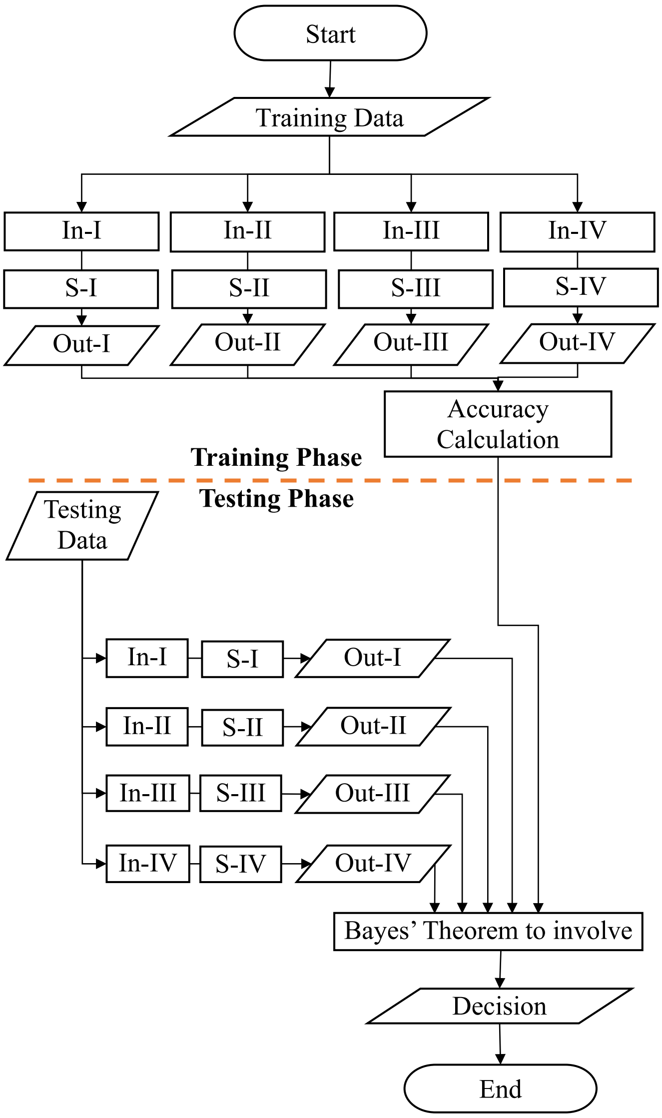


Figure 7. Flowchart - Diagnosis of automated multi-bearing fault

Table 10 shows example of calculation for Bayes' Theorem in handling conflicting results. When the Support Vector Machine results have conflicting between predicted results and actual results, the Bayes Theorem can calculate the probability of Class I, Class II, Class III or Class IV, in order to made the final decision. The probability of each class can be calculated as the following.

Table 10. The illustration of calculation for Bayes' Theorem

Class		Actual result				
		I	II	III	IV	Total
Predicted result	I	9	-	-	-	9
	II	-	15	-	-	15
	III	-	-	10	1	11
	IV	-	-	-	4	4
	Conflict	6	-	5	10	21
	Total	15	15	15	15	60

$$P\ (\text{Class I} \mid \text{Symptoms}) = \frac{P\ (\text{Symptoms} \mid \text{Class I})P\ (\text{Class I})}{P\ (\text{Symptoms})}$$

$$= \frac{P(\text{Symptoms} \mid \text{Class I})P(\text{Class I})}{P(\text{Symptoms} \mid \text{Class I})P(\text{Class I}) + P(\text{Symptoms} \mid \text{NonClass I})P(\text{NonClass I})}$$

$$= \frac{0.6 \times 0.25}{0.6 \times 0.25 + 1.133 \times 0.75}$$

$$= 0.15$$

$$P\ (\text{Class II l Symptoms}) = \frac{P\ (\text{Symptoms l Class II})P\ (\text{Class II})}{P\ (\text{Symptoms})}$$

$$= \frac{P(\text{Symptoms l Class II})P(\text{Class II})}{P(\text{Symptoms l Class II})P(\text{Class II}) + P(\text{Symptoms l NonClass II})P(\text{NonClass II})}$$

$$= \frac{1 \times 0.25}{1 \times 0.25 + 1.133 \times 0.75}$$

$$= 0.25$$

$$P\ (\text{Class III l Symptoms}) = \frac{P\ (\text{Symptoms l Class III})P\ (\text{Class III})}{P\ (\text{Symptoms})}$$

$$= \frac{P(\text{Symptoms l Class III})P(\text{Class III})}{P(\text{Symptoms l Class III})P(\text{Class III}) + P(\text{Symptoms l NonClass III})P(\text{NonClass III})}$$

$$= \frac{0.667 \times 0.25}{0.667 \times 0.25 + 1.111 \times 0.75}$$

$$= 0.1667$$

$$P\ (\text{Class IV l Symptoms}) = \frac{P\ (\text{Symptoms l Class IV})P\ (\text{Class IV})}{P\ (\text{Symptoms})}$$

$$= \frac{P(\text{Symptoms l Class IV})P(\text{Class IV})}{P(\text{Symptoms l Class IV})P(\text{Class IV}) + P(\text{Symptoms l NonClass IV})P(\text{NonClass IV})}$$

$$= \frac{0.333 \times 0.25}{0.333 \times 0.25 + 1 \times 0.75}$$

$$= 0.1$$

According to the calculation, the Class I probability was 15%, Class II probability was 25%, Class III probability was 16.67%, and Class IV probability was 10%. Based on the slightly higher value of probability, the final decision was Class II. The accuracy of support vector machine classification results was found to be increased from 72% to 95% (a total increase of 23%) by implementing the proposed method.

Bayes' Theorem was used in the proposed automatic bearing fault diagnosis model to differentiate the classification accuracy of each support vector machine model during the training phase. When classification accuracies are combined in the phase of training, the ultimate decision is made by refining the conflicting results in the testing phase. Table 11 shows elimination of conflicting results were implemented. It was recorded that the accuracy of classification was improved from 72% to 95%.

Table 11. Confusion matrix with Bayes Theorem

	Class	Actual Result			
		I	II	III	IV
Predicted result	I	42	-	-	-
	II	5	48	-	-
	III	3	2	50	-
	IV	-	-	-	50
	Conflict	-	-	-	-
Sensitivity in %		84.0	96.0	100	100
Accuracy in %		95.0			

Figure 9 shows the comparison between the three bearing fault diagnosis model, which are the Support Vector Machine Model (one vs all strategy), Tuned Support Vector Machine Model (one vs all strategy) and Support Vector Machine Model (one vs all strategy) with Bayes' Theorem. From the comparison, it is evident that the proposed Support Vector Machine (one vs all strategy) - Bayes' Theorem Model has demonstrated the highest accuracy in multi-bearing faults classifications.

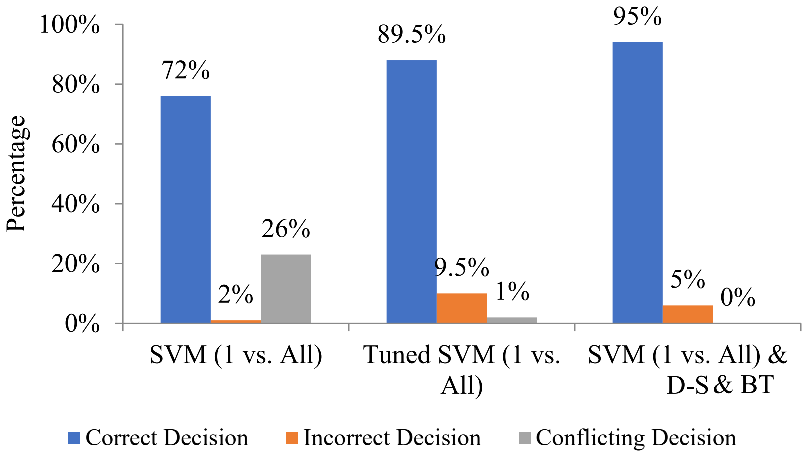


Figure 8. Comparison – Final decisions of support vector machine models

From all the data collected, half of them were used as training data (from each condition), and the other half of the data was used as testing data. Next, we triggered the validation process to obtain the standard deviation and the mean of accuracy. Each of the model implement the validation for 20 cycles of repetition. Table 12 shows the mean classification accuracy of the one vs all strategy Support Vector Machine Model is 81.3±5.2%. The mean classification accuracy of one vs all strategy Support Vector Machine Model with Bayes' Theorem is 91.3±1.4%.

Table 12. Cyclical assessment results

Cycle	Support Vector Machine Model (1 vs All)		Support Vector Machine Model (1 vs All) with Bayes' Theorem	
	Accuracy	Number of conflicting results	Accuracy	Number of conflicting results
1	0.780	39	0.915	-
2	0.760	42	0.930	-
3	0.775	43	0.940	-
4	0.855	17	0.900	-
5	0.735	51	0.905	-
6	0.855	16	0.900	-
7	0.825	28	0.905	-
8	0.780	39	0.925	-
9	0.835	27	0.905	-
10	0.825	19	0.910	-
11	0.880	11	0.915	-
12	0.745	47	0.920	-
13	0.725	54	0.910	-
14	0.740	49	0.900	-
15	0.865	11	0.905	-
16	0.870	11	0.895	-
17	0.860	21	0.910	-
18	0.855	16	0.910	-
19	0.825	26	0.945	-
20	0.860	18	0.905	-
Mean ± standard deviation	0.813±0.052	29±15	0.913±0.014	-

Further validation was needed to test the Support Vector Machine – Bayes' Theorem Model with a new set of testing data. These data include experimental data from a few bearing conditions, such as severities of faults, motor speeds and motor loads. Table 13 shows the classification accuracy of the new data managed by Support Vector Machine – Bayes' Theorem Model is 89%. This proves that training data was not overfitted because Support Vector Machine – Bayes' Theorem Model was not being trained to do so. Bearing conditions from several severities of faults, motor speeds and motor loads are being recognized.

Table 13. New data of testing phase confusion matrix for the support vector machine – Bayes' Theorem Model

	Class	Actual result			
		I	II	III	IV
Predicted result	I	630	98	2	22
	II	340	1835	3	0
	III	30	67	1976	189
	IV	0	0	19	1789
	Conflict	0	0	0	0
Sensitivity in %		63.0	91.8	98.8	89.5
Accuracy in %		89.0			

7.0 CONCLUSION

During the study, Case Western Reserve University Bearing Data Center simulated the experimental data that was used as artificial intelligence model input. Because of one vs all strategy for support vector machine multi-bearing faults classification has the lowest misclassification accuracy (only at 2%), it has proven to be better than others. Studies were conducted regarding employing Support Vector Machine and Bayes' Theorem in the classification when the feasibility of the proposed diagnosis of automated fault bearing. The introduction of the Support Vector Machine – Bayes' Theorem Model increase the classification accuracy from 72% to 95%. This has proven that Bayes Theorem can eliminate all the indecisive results of the support vector machine model and increase the classification accuracy. Further to that, new data sets are triggered to validate Support Vector Machine – Bayes' Theorem Model. From the new set of data, a classification accuracy of 89% was obtained. In summary, for the diagnosis of automated bearing fault conditions, the Support Vector Machine – Bayes' Theorem model in classification accuracy is found to be more accurate and effective.

8.0 CONFLICT OF INTERESTS

All authors declare that they have no conflicts of interest.

9.0 ACKNOWLEDGEMENT

A special thanks of gratitude to my supervisor, Associate Professor Ir. Dr. Lim Meng Hee, who gave me the opportunity to work on this wonderful research topic "Bayes Theorem for Multi Bearing Fault Diagnosis", which also encouraged me to explore the applications of artificial intelligence in the field of machinery fault diagnosis.

10. REFERENCES

[1] L. Gelman, B. Murray, T.H. Patel, A. Thomson, "Vibration diagnostics of rolling bearings by novel nonlinear non-stationary wavelet bicoherence technology," *Engineering Structures*, vol. 80, pp. 514–520, 2014.

[2] O. Janssens, R. Schulz, V. Slavkovikj, K. Stockman, M. Loccufier, et al., "Thermal image-based fault diagnosis for rotating machinery," *Infrared Physics & Technology*, vol. 73, pp. 78–87, 2015.

[3] D.P. Jena, and S.N. Panigrahi, "Automatic gear and bearing fault localization using vibration and acoustic signals," *Applied Acoustics,* vol. 98, pp. 20–33, 2015.

[4] J. Chen, Y. Zi, Z. He, and X. Wang, "Adaptive redundant multiwavelet denoising with improved neighbouring coefficients for gearbox fault detection," *Mechanical Systems and Signal Processing,* vol. 38, pp. 549–568, 2013.

[5] K.H. Hui, L.M. Hee, M.S. Leong, and A.M. Abdelrhman, "Time-frequency signal analysis in machinery fault diagnosis: review," *Advanced Materials Research,* vol. 845, pp. 41–45, 2013.

[6] L. Cui, N. Wu, C. Ma, and H. Wang, "Quantitative fault analysis of roller bearings based on a novel matching pursuit method with a new step-impulse dictionary," *Mechanical Systems and Signal Processing,* vol. 68-69, pp. 34–43, 2016.

[7] X. Yan, Y. Liu, Y. Xu, and M. Jia, "Multistep forecasting for diurnal wind speed based on hybrid deep learning model with improved singular spectrum decomposition," *Energy Conversion and Management,* vol. 225, pp. 1–22, 2020.

[8] X. Yan, Y. Liu, Y. Xu, and M. Jia, "Health condition identification for rolling bearing using a multi-domain indicator-based optimized stacked denoising autoencoder," *Structural Health Monitoring,* vol. 19, pp. 1602–1626, 2020.

[9] X. Yan, Y. Liu, Y. Xu, and M. Jia, "Multiscale cascading deep belief network for fault identification of rotating machinery under various working conditions," *Knowledge-Based Systems,* vol. 193, pp. 1–20, 2020.

[10] Y. Kaisi, Z. Lianyu, and W. Chenglin, "A new intelligent bearing fault diagnosis model based on triplet network and SVM," S*cientific Reports*, vol. 12, p. 5234, 2022.

[11] X. Zhang, D. Qiu, and F. Chen, "Support vector machine with parameter optimization by a novel hybrid method and its application to fault diagnosis," *Neurocomputing*, vol. 149, pp. 641–651, 2015.

[12] Ł. Jedliński, and J. Jonak, "Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform," *Applied Soft Computing,* vol. 30, pp. 636–641, 2015.

[13] P.K. Kankar, S.C. Sharma, and S.P. Harsha, "Fault diagnosis of rolling element bearing using cyclic autocorrelation and wavelet transform," *Neurocomputing*, vol. 110, pp. 9–17, 2013.

[14] R. Jegadeeshwaran, and V. Sugumaran, "Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines," *Mechanical Systems and Signal Processing,* vol. 52-53, pp. 436–446, 2015.

[15] H. Keskes, A. Braham, and Z. Lachiri, "Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet SVM," *Electric Power Systems Research,* vol. 97, pp. 151–157, 2013.
[16] S. Cheong, S. Oh, and Lee, "Support vector machines with binary tree architecture for multi-class classification," *Neural Information Processing - Letters and Reviews,* vol. 2, pp. 47–51, 2004.
[17] Y.S. Wang, Q.H. Ma, Q. Zhu, X.T. Liu, and L.H. Zhao, "An intelligent approach for engine fault diagnosis based on Hilbert–Huang transform and support vector machine," *Applied Acoustics,* vol. 75, pp. 1–9, 2014.
[18] Z. Liu, H. Cao, X. Chen, Z. He, and Z. Shen, "Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings," *Neurocomputing*, vol. 99, pp. 399–410, 2013.
[19] L.M.R. Baccarini, V.V. Rocha e Silva, B.R. de Menezes, and W.M. Caminhas, "SVM practical industrial application for mechanical faults diagnostic," *Expert Systems With Applications,* vol. 38, pp. 6980–6984, 2011.
[20] M. Cui, Y. Wang, X. Lin, and M. Zhong, "Fault diagnosis of rolling bearings based on an improved stack autoencoder and support vector machine," *IEEE Sensors Journal,* vol. 21, no. 4, pp. 4927–4937, 2021.
[21] L. Wenbo, L. Shengnan, and Q. Xiwen, "Weighted p-norm distance t kernel SVM classification algorithm based on improved polarization," *Scientific Reports*, vol. 12, p. 6197, 2022.
[22] P. Konar, and P. Chattopadhyay, "Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs)," *Applied Soft Computing Journal*, vol. 11, pp. 4203–4211, 2011.
[23] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," 2010 [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf. [Accessed: April, 2022].
[24] F. Chen, B. Tang, T. Song, and L. Li, "Multi-fault diagnosis study on roller bearing based on multi-kernel support vector machine with chaotic particle swarm optimization," *Measurement*, vol. 47, pp. 576–590, 2014.
[25] C. Chang, and C. Lin, "LIBSVM," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 1–27, 2011.
[26] C. Bishop, *Pattern recognition and machine learning*. New York: Springer, 2006.
[27] Z. Dienes, *Understanding psychology as a science: An introduction to scietific and statistical inference*. Palgrave Macmillan, 2008.
[28] M.C Frank, and N.D. Goodman, "Predicting pragmatic resoning in language games," *Science*, vol. 336, pp. 998, 2012.
[29] M. Jones, and B. Love, "Bayesian fundamentalism or enlightenment? status and theoretical contributions of Bayesian models of cognition. Behavioral and Brain Sciences," vol. 34, pp. 192-193, 2011.
[30] E. Parent, and E. Rivot, *Introduction to hierarchical bayesian modeling for ecological data*. Boca Raton: Chapman and Hall, 2012.
[31] MathWorks, Statistics and Machine Learning ToolboxTM User's Guide R2016a, 10.2 ed. The MathWorks Inc., Natick, 2016.