Journal of Engineering Science and Technology
Vol. 11, No. 6 (2016) 826 - 838

EFFECT OF KERNEL FUNCTION IN SUPPORT VECTOR MACHINE FOR THE FAULT DIAGNOSIS OF PUMP

N. R. SAKTHIVEL[1,*], S. SARAVANAMURUGAN[1], BINOY B.NAIR[2], M. ELANGOVAN[1], V. SUGUMARAN[3]

[1]Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Ettimadai, Coimbatore, Tamilnadu-641105, India
[2]Department Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Ettimadai, Coimbatore, Tamilnadu-641105, India
[3]School of Mechanical and Building Sciences, VIT University, Chennai campus, Chennai, Tamilnadu, India
*Corresponding Author: m_elangovan@cb.amrita.edu

Abstract

Pumps are widely used in a variety of applications. Defects and breakdown of these pumps will result in significant economic loss. Therefore, these must be under continuous observation. In various applications, the role of pump is decisive and condition monitoring is crucial. A completely automated on-line pump condition monitoring system which can automatically inform the operator of any faults, promising reduction in maintenance cost with a greater productivity saving both time and money.This paper presents the application of support vector machine for classification using statistical features extracted from vibration signals under good and faulty conditions of a pump. Effectiveness of various kernel functions of C-SVC and γ-SVC models are compared. The study gives some empirical guidelines for selecting an appropriate kernel in a classification problem.

Keywords: Pump, Support vector machine, Fault diagnosis, Statistical features, Kernel function.

1. Introduction

Pump plays vital role in various industries. Pumps are the important part of food industry, natural gas, steam, water, gasoline, solar, construction, distillery, and automotive companies etc. Keeping the pump well-maintained will extend its

Nomenclatures

BF	Bearing fault
CAV	Cavitation
SF	Seal fault
BFIF	Faulty bearing and faulty impeller

Abbreviations

BEP	Best Efficiency Point
NPSH	Net Positive Suction Head
PSVM	Proximal Support Vector Machine
RBF	Radial Basis Function
SVM	Support Vector Machine

existence and add quality to its functionality. Pump faults such as seal defect impeller defect, bearing defect and cavitation can cause adverse effects such as pitting of pump, erosion and structural vibration, drop in efficiency etc. Fault diagnosis involves the collection and evaluation of data from pump. A primary element of this method is vibration monitoring. Assessing vibration is the most common approach to evaluate changes in the operating condition of rotating machinery, including pumps, fans, turbines, rotary engines and propeller, and also best suited for detecting pump imbalances, misalignments of shaft. The faulty components in the pump directly influence the desired pump characteristics. Pump fault diagnosis is a three stage process. It consist of, selection of the parameters to be captured, extraction of features from the selected parameters, selection and classification of features. After acquiring the signal, to obtain meaningful information from the acquired signal, useful and significant features need to be extracted. Fast Fourier Transform does not give satisfactory results hence, sophisticated signal processing method is necessitate of the hour. Sophisticated methods are slow and it requires high speed and high capacity processor. This fact encouraged us to find techniques that require less computation efforts. In the same fortitude Sugumaran et al., [1] used C4.5 decision tree algorithm and Proximal Support Vector Machine for fault diagnostics of roller bearing. Jegadeeshwaran et al., [2] showed the use of statistical feature extracted from time domain signals for fault diagnosis of automobile hydraulic brake system effectively. Hence statistical features have been extracted from the acquired vibration signals for the fault diagnosis of pump. The use of proper diagnostic tool to classify the acquired and processed signals is the third stage in fault diagnosis. Currently there are a huge number of classification algorithms, each having their case history of success and failure.

Alfayez et al. [3] used acoustic emission for detecting cavitation and determining the best efficiency point (BEP) of a centrifugal pump based on net positive suction head (NPSH) However, the use of acoustic emission method is to detect only cavitation and is not useful in distinguish other faults. Wang and Hu [4] used fuzzy logic for fault classification of pump from the features extracted from the vibration signals. According to his work fuzzy logic was more sensitive to the quantity and the quality of inputs. Sugumaran et al., [5] were successful in establishing a quantitative relationship using entropy values and vibration measurements for feature selection. Hence entropy can be useful in selecting good features giving a motivation to use decision tree for feature selection in this work. Widodo et al., [6] attempted to summarize and assess the latest research and developments of SVM in fault diagnosis. Saravanan et al., [7] used fast single-shot multiclass Proximal Support Vector Machine (PSVM) for the fault diagnostics of inaccessible gear and was reported that multiclass proximal support vector machine yielded good results for large classes of data in less time. On the bearing fault detection, Samanta et al., [8] used artificial neural networks (ANNs) and support vector machines (SVM). They reported that the performance of SVM has better classification capability than the ANN. Yuan & Chu [9] brought out the use of artificial immunization algorithm (AIA) to optimize the parameters of SVM for the fault diagnosis of turbo rotor pump. The results of AIA-SVM and standalone SVM were compared and was reported that AIA-SVM outperforms the standalone SVM. Xiang et al., [10] proposed a method for the shaft orbit based on Walsh transform and support vector machine. Muralidharan et al., [11] used wavelet features and SVM for the fault diagnosis of monoblock centrifugal pump. Elangovan et al., [12] discussed condition monitoring of carbide tipped tool using SVM and compared the classification efficiency of c-SVC and $\gamma - SVC$. Such studies provide an impulsion to conduct work in condition monitoring of pump using c-SVC and $\gamma - SVC$ with different kernel function and find out the appropriate kernel function for this study.

In this paper the vibration signal from an accelerometer is captured for various conditions such bearing, seal, impeller, bearing and impeller faults together and cavitation etc,. The statistical features were extracted and best features were selected by using decision tree C4.5 algorithm and it was classified successfully using four different kernel functions such as linear function, polynomial function, radial basis function (RBF) and sigmoid function for c-SVC and γ -SVC model of SVM.

The rest of the paper is organised as follows. In Section 2, experimental setup and experimental procedure is described. Section 3 presents feature extraction from the time domain signal. In Section 4, feature selection using decision tree C4.5 algorithm is discussed. Support vector machine is detailed in Section 5 which is engaged here to classify the various faults of pump. Consequently Section 6 presents results of the experiment. Conclusions are presented in the final section.

2. Experimental setup [13]

Experimental setup consists of a pump with 2HP motor. To simulate cavitation a control valve is connected at the inlet of the pump. The control valve is used to make the pressure drop between the suction and at the eye of the impeller. Acrylic pipes are connected on the inlet and at the outlet of the impeller, to visualize the cavitation. A computer stored the signals from a piezo-electric accelerometer. Adhesive was used to hold the accelerometer on the casing near the eye of the impeller. The output of accelerometer was feed to DACTRAN for signal conditioning. A USB Port was used to feed digital signal directly to computer. Then the signals are processed from the memory to extract different features. The experimental setup is shown in Fig.1.

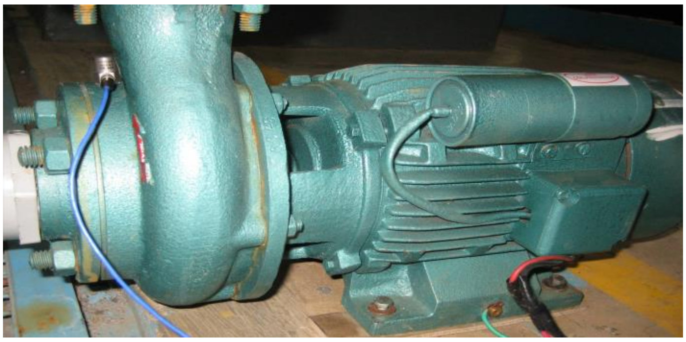


Fig. 1. Experimental test rig [14].

Bearing fault

In this work, KBC 6203 roller bearings were used. One was a defect free bearing and the other bearing, defect was created using wire cut electric discharge machining in order to keep the size of the defect under control. Figure 2 shows the faulty bearing.

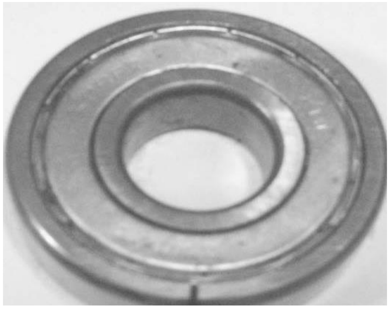


Fig. 2. Faulty bearing.

Seal defect

A seal will fail or seep out when the pump running under dry condition for a long period of time or at the time of installation due to extreme installation pressure. In this study seal defect was created by hammering the seal during installation. (refer Fig. 3).

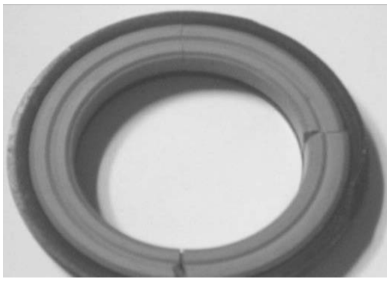


Fig. 3. Faulty seal.

Impeller fault

In the impeller defect was created by removing a tiny portion of metal through a machining process (see Fig. 4).

Fig. 4. Faulty impeller.

Experimental procedure

The vibration signals are acquired from the pump working at a constant speed of 2880 rpm. Pump details are shown in Table 1. The vibration signal from accelerometer mounted at the eye of the impeller was taken. The parameters for signal acquisition such as sampling frequency, sampling length, type of signal, etc., were set. According to Nyquest sampling theorem, frequency of sample acquirement needs to be twice the monitored frequency. The highest frequency monitored was 12 KHz; hence 24 KHz was set as sampling frequency and sampling length was set to 2^{13} = 8192 for all conditions of the pump. Data acquirement was started and don't taken into account the initial signals. After attaining the full speed, two hundred and fifty signals were recorded in each conditions of the pump.

Table 1. Pump specifications.

Speed	**:** 2880 rpm	**Pump size**	**:** 50×50mm
Current	**:** 11.5A	**Discharge**	**:** 392 L/s
Head	**:** 20 m	**Power**	**:** 2 HP

3. Feature Extraction

In this work Microsoft Excel was used to extract different descriptive statistical features. The code used to extract statistical feature is given in Fig. 5. The time domain signals can be used to perform fault diagnosis by analysing vibration signals obtained from the experiment. Time domain refers to variation of amplitude of signal with respect to time Statistical methods have been widely used, and can provide the physical characteristics of time domain data. The eleven statistical features have been extracted from the vibration signals. The extracted statistical features were classified using different kernels like linear, polynomial, radial basis function and sigmoid functions of c-SVC and γ -SVC.

```
Sub test()                                                           Procedure
'
' test Macro
'

'
   For i = 1 To 250
    ChDir "F:\sakthitest\good"
    Workbooks.OpenText Filename:="F:\sakthitest\good\input1(t) " & i & ".txt", Origin:= _
        437, StartRow:=1, DataType:=xlDelimited, TextQualifier:=xlDoubleQuote, _
        ConsecutiveDelimiter:=False, Tab:=True, Semicolon:=False, Comma:=False _
        , Space:=False, Other:=False, FieldInfo:=Array(1, 1), _
        TrailingMinusNumbers:=True
    Rows("1:29").Select
    Selection.Delete Shift:=xlUp
     Application.Run "ATPVBAEN.XLAM!Descr", ActiveSheet.Range("$A:$A"), "", "C" _
        , False, True
    Range("B3:B15").Select
    Selection.Copy
    Windows("temp.xls").Activate
    ActiveWindow.WindowState = xlNormal
    ActiveWindow.WindowState = xlNormal
    Range("A1" & i & "").Select
    Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
        :=False, Transpose:=True
    Windows("input1(t) " & i & ".txt").Activate
    ActiveWindow.WindowState = xlNormal
    ActiveWindow.WindowState = xlNormal
    ActiveWindow.Close
     Next i
```

Fig. 5. Code for extracting statistical features.

A brief outline of code

(1) Store the signal data in .txt format. Label files at the end with file number
(2) Generate a excel file and save it as temp.xlsx.
(3) Generate a macro enable file and create a new macro using this as reference.

4. Decision Tree

The tree which is used to represent decisions and decision making is called decision tree. It is a classifier. Classification or regression models are built by decision tree in the form of tree structure. In data mining the decision tree represents data. It breaks down data set into smaller and smaller subsets while at the same time a related decision tree is incrementally developed. The decision tree consists of two nodes. They are decision and leaf nodes. Decision node has two or more branches. Leaf node represents a decisions or classification. It classifies sample by starting at the root of the tree and moving through it until a leaf node. The target variable in the decision tree models can be predicted based on a number of input variables. Larger trees are having less accuracy than smaller trees. J48 algorithm is a widely used to construct decision trees [15]. At each decision node in the decision tree, the most important features can be selected for classification using appropriate estimation criteria. The criterion used to identify the best feature invokes the concept of information gain and entropy reduction. Information gain is derived from information theory and it is based on the concept of entropy. Information gain measures how well a given feature separates the training samples according to their target classification. This measure is used to select among the candidate features at each step while growing the tree. Entropy is a measure of homogeneity of the set of samples.

Given a set S of positive and negative examples of some target concept (a 2-class problem), the entropy of set S relative to this binary classification is

$$E(S) = -p(P)\log 2\, p(P) - p(N)\log 2\, p(N) \tag{1}$$

If the sample is entirely homogeneous the entropy is zero and if the sample is an equally divided it has entropy of one.

Information gain measures the expected reduction in entropy, or uncertainty.

$$Gain(S, A) = Entropy(S) - \sum_{v \in Value(A)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{2}$$

where Values (A) is the set of all possible values for attribute A, and S_v is the subset of S for which feature A has value *v (i.e.,* S_v is the subset for *S* which feature A has value *v*(i.e., $S_v = \{s\ S|A(s) = v\}$).

The first term in the Eq. (2) for Gain is the entropy of the original collection S and the expected value of the entropy after S is partitioned using feature A is the second term. The expected entropy described by the second term is the sum of entropies of each subset S_v, weighed by the fraction of samples $|S_v|/(S)$ that belong to S_v. Gain *(S, A)* is therefore the expected reduction in entropy caused by knowing the value of feature A.

In this work the decision tree is used for feature selection because the decision tree provide a clear indication of which fields are most important for prediction.

4.1. Application of decision tree for feature selection

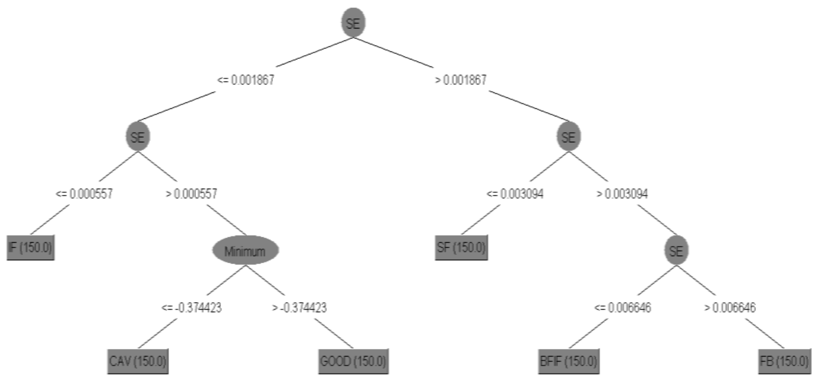
The algorithm has been applied to the problem under discussion for feature selection. Eleven statistical features have been given as input to the algorithm; the output of the algorithm is decision tree, which is shown in Fig. 3. It is clear there from that the top node is the best node for classification. The level of contribution is not same and all eleven features are not equally important. The level of contribution by each individual feature is given by a statistical measure within the parenthesis in the decision tree. The other features appear in the nodes of decision tree in descending order of importance. It is to be stressed here that only features that contribute to the classification appear in the decision tree and others do not. Features that have less discriminating capability can be consciously discarded by deciding on the threshold. This concept is made use of in selecting good features. A feature is ‘a good feature’, when its discriminating ability is high among the classes. It is characterised by the following.

(a) The feature values do not vary much within a class.
(b) It varies much among the classes.

The features which satisfy the above conditions will have more information gain while splitting and thus they appear in the order of importance in decision tree.

4.2. Features suggested by decision tree

The features that dominate generally represent the pump condition descriptors. Referring to Fig. 6, one can identify two such most dominant features, (a) minimum value (b) standard error.



IF- Impeller fault	GOOD-Good condition	BF-Fault bearing
CAV-Cavitation	BFIF-Faulty bearing and faulty impeller	SF-Seal Fault

Fig. 6. Decision tree [16].

5. Support vector Machines [16]

A Support Vector Machine (SVM) performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two categories. SVM models are closely related to neural networks. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, feed-forward neural network. SVM models are a close cousin to classical neural networks. Using a kernel function, SVM‘s are an alternative training method for polynomial function, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training.

In the parlance of SVM literature, a predictor variable is called an attribute, and a transformed attribute that is used to define the hyper plane is called a feature. The task of choosing the most suitable representation is known as feature selection. A set of features that describes one case (i.e., a row of predictor values) is called a vector. Hence the goal of SVM modeling is to find the optimal hyper plane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyper plane are the support vectors.

The basic idea of applying SVM to pattern classification can be stated as follows: first, map the inputs vectors into one features space, possible in higher space, either linearly or nonlinearly, which is relevant with the kernel function. Then, within the feature space from the first step, seek an optimized linear division, that is, construct a hyper-plane which separates two classes. However, this technique can also be extended to multi-class classification. SVM training seeks a global optimized solution and avoid over-fitting, hence it has the ability to deal with a large number of features. SVM has the potential to handle very large feature spaces, because the training of SVM is carried out so that the dimension of classified vectors does not has as distinct an influence on the performance of SVM as it has on the performance of conventional classifier. That is why it is noticed to be especially efficient in large classification problem. This will also benefit in faults classification, because the number of features to be the basis of fault diagnosis may not have to be limited. Also, SVM-based classifier is claimed to have good generalization properties compared to conventional classifiers, because in training SVM classifier the so-called structural misclassification risk is to be minimized, whereas traditional classifiers are usually trained so that the empirical risk is minimized.

The next logical step is classification using a classifier. Support Vector Machines is used as the classifier here. It is a new generation learning system based on statistical learning theory. SVM belongs to the class of supervised learning algorithms in which the learning machine is given a set of features (or inputs) with the associated labels (or output values). Each of these features can be looked upon as a dimension of a hyper-plane. SVMs construct a hyper-plane that separates the hyper-space into two classes (this can be extended to multi-class problems). While doing so, SVM algorithm tries to achieve maximum separation between the classes (Fig. 7). Separating the classes with a large margin minimizes the expected generalization error. By ‘minimum generalization error’, we mean that when a new set of features (that is data points with unknown class values) arrive for classification, the chance of making an error in the prediction (of the class to which it belongs) based on the learned classifier (hyper-plane) should be minimum. Intuitively, such a classifier is one, which achieves maximum separation-margin between the classes.

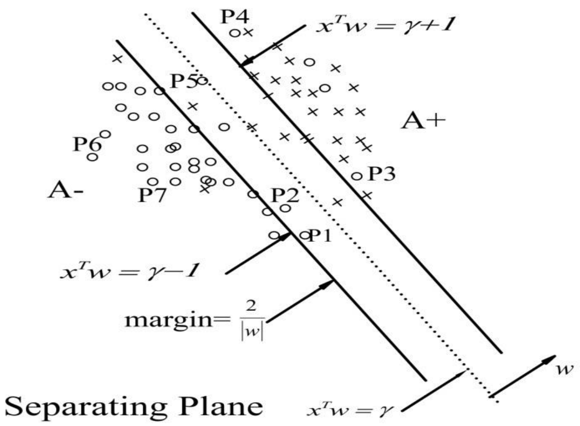


Fig. 7. Standard SVM classifier.

The above process of maximizing separation leads to two hyper-planes parallel to the separating plane, on either side of it. These two can have one or more points on them. The planes are known as ‘bounding planes’ and the distance between them is called as ‘margin’. By SVM ‘learning’, we mean, finding a hyper-plane, which maximizes the margin and minimizes the misclassification error. The points lying beyond the bounding planes are called support vectors. The data points P1, P2, P3, P4, and P5 belonging to A- are support vectors, but P6, P7 are not. Same facts hold good for class A+. These points play a crucial role in the theory and hence the name Support Vector Machines. Here, by 'machine', we mean an algorithm. In the formulation, 'A' is a $m \times n$ matrix whose elements belong to real space, 'D" is $m\times1$ matrix representing class label (+1 and -1), 'e' is a vector of ones and 'ν' is a control parameter that defines the weight of error minimization and bounding plane separation in the objective function. 'w' is orientation parameter and 'γ' is location parameter (location relative to origin) of separating hyper plane.

$$\left.\begin{array}{l} \min\limits_{(w,\gamma,y)\in R^{n+1+m}} \nu e' y + \frac{1}{2} w' w \\ \\ s.t. D(Aw - e\gamma) + y \geq e \\ \quad y \geq 0 \end{array}\right\} \tag{3}$$

where $A \in R^{m\times n}, D \in \{-1,1\}^{m\times 1}, e = 1^{m\times 1}$.

Vapnik [17] has shown that if the training features are separated without errors by an optimal hyper-plane, the expected error rate on a test sample is bounded by the ratio of the expectation of the support vectors to the number of training vectors. The smaller the size of the support vector set, more general the above result. Further, the generalization is independent of the dimension of the problem. In case such a hyper-plane is not possible, the next best is to minimize the number of misclassifications whilst maximizing the margin with respect to the correctly classified features.

After training, for any new set of features prediction of its class is possible using the decision function as given below, which is a function of 'w' and 'γ'. It is called testing.

$$f(x) = sign(w^T x - \gamma) \tag{4}$$

If the value of *f(x)* is positive then new set of features belongs to class A+; otherwise it belongs to class A-.

Application of SVM for problem under study

For each class (bearing fault, seal fault, impeller fault, bearing and impeller fault together and cavitation), features consisting of 250 feature value sets were collected from the experiment. 150 samples in each class were used for training the SVM and 100 samples were reserved for testing. Here two types of SVM models were (C-SVC and $\gamma - SVC$) tested for the following four basic kernels.

Linear: $K(X_i, X_j) = X_i^T, X_j$

Polynomial: $K(X_i, X_j) = \left(\gamma X_i^T X_j + \nu\right)^d, \gamma > 0$

Radial basis function (RBF): $K\left(X_i, X_j\right) = \exp\left(-\gamma \,||X_i - X_j||^2\right), \gamma > 0$

Sigmoid: $K(X_i, X_j) = \tanh\left(\gamma X_i^T X_j + \nu\right)$

Here γ, ν and d are kernel parameters. The results are discussed in the following section.

6. Results and discussion

The experimental studies have been carried out for good condition and various defective conditions of the pump as discussed in Section. 2.

Classification is a two phase process: training and testing. Training is the process of learning to label from the examples. Training can be supervised mode or unsupervised mode. Here, supervised mode is used for training. Testing is the process of checking how well the classifier has learnt to label the unseen examples. The four different SVM kernel functions of C-SVC and $\gamma - SVC$ models such as linear function, polynomial function, radial basis function (RBF) and sigmoid function, were used for classification. The performance of -SVC and $\gamma - SVC$ models for various kernel functions are shown in table 2. Fig. 8 shows the plot of various kernel functions and their percentage classification accuracy with C-SVC and $\gamma - SVC$. From the above table for both C-SVC and $\gamma - SVC$ models, the RBF kernel function gives higher classification accuracy. In general RBF is a reasonable first choice. The RBF kernel nonlinearly maps samples into a higher dimensional space, it can handle nonlinear relationships between target categories and predictor attributes.

Table 2. Performance of C-SVC and $\gamma - SVC$ models for various kernel functions.

Type of SVM model	% classification accuracy of various kernel functions			
	Radial basis function(RBF)	**Linear function**	**Polynomial function**	**Sigmoid function**
C-SVC	99.93	99.73	99.53	98.53
$\gamma - SVC$	99.93	99.83	99.53	99.83

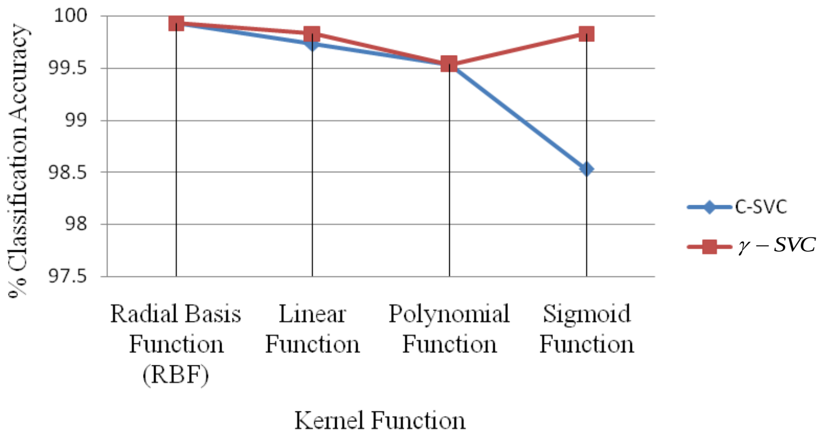


Fig. 8. classification accuracy of C-SVC and $\gamma - SVC$ models for different kernel functions.

7. Conclusion

This paper deals with vibration based fault diagnosis of pump. Six classical states viz., normal, bearing fault, impeller fault, seal fault, impeller and bearing fault together, cavitation are simulated on pump. Set of features have been extracted and classified using two models of SVM with four kernel functions. Both the SVM models yielded a classification accuracy in excess of 98.5% of all the kernel function tested, the RBF provided an accuracy of 99.93% in fault classification with both C-SVC and $\gamma - SVC$. From the results it can be confidently said that SVM classifier with RBF kernel function is a good candidate for fault diagnosis of pump.

References

1. Sugumaran, V.; Muralidharan, V.; and Ramachandran, K.I. (2007). Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing. *Mechanical Systems and Signal Processing*, 21, 930-942.
2. Jegadeeshwaran, R.; and Sugumaran, V.(2013). Comparative study of decision tree classifier and best first tree classifier for fault diagnosis of automobile hydraulic brake system using statistical features. *Measurement*, 46(9), 3247-3260.
3. Alfayez, L.; Mba, D.; and Dyson, G. (2005) The application of acoustic emission for detecting incipient cavitation and the best efficiency point of a 60 kW mono block centrifugal pump. *NDT&E International*,38, 354-358.
4. Wang, J.; and Hu, H, (2006).Vibration-based fault diagnosis of pump using fuzzy technique. *Measurement*, 39, 176-185.
5. Sugumaran, V.; and Ramachandran, K.I. (2007). Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing. *Mechanical Systems and Signal Processing*, 21, 2237-2247.
6. Widodo, A.; and Yang, B.S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 21, 2560-2574.
7. Saravanan, N.; Kumar Siddabattuni, V.N.S.; and Ramachandran, K.I. (2008). A comparative study on classification of features by SVM and PSVM extracted using Morlet wavelet for fault diagnosis of spur bevel gear box. *Expert Systems with Applications*, 35, 1351- 1366.
8. Samanta, B.; Al-Balushi, K.R.; and Al-Araim, S.A. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence*, 16, 657-665.
9. Yuan, S.; and Chu, F.(2007). Fault diagnosis based on support vector machines with parameter optimization by artificial immunization algorithm. *Mechanical Systems and Signal Processing*, 21, 1318-1330.
10. Xiang, X.; Zhou, J.; An, X.; Peng, B.; and Yang, J. (2008). Fault diagnosis based on Walsh transform and support vector machine. *Mechanical Systems and Signal Processing*, 22, 1685-1693.
11. Sugumaran, V.; Muralidharan, V.; and Sakthivel, N.R.(2011) Wavelet decomposition and support vector machine for fault diagnosis of monoblock centrifugal pump. *International Journal of Data Analysis Techniques and Strategies*, 3(2), 159-177
12. Elangovan, M.; Sugumaran, V.; Ramachandran, K.I.; and Ravikumar, S. (2011). Effect of SVM kernel functions on classification of vibration signals of a single point cutting tool. *Expert Systems with Applications*, 38, 15202-15207.

13. Sakthivel, N.R.; Sugumaran, V.; and Nair, B.B. (2010). Comparison of decision tree-fuzzy and rough set-fuzzy methods for fault categorization of mono-block centrifugal pump. *Mechanical Systems and Signal Processing*. 24, 1887-1906.
14. Sakthivel, N.R.; Sugumaran, V.; Nair, B.B. (2012). Automatic rule learning using roughset for fuzzy classifier in fault categorization of mono-block centrifugal pump. *Applied Soft Computing,* 12 (1), 196-203
15. Quinlan, J.R. (1996). Improved use of continuous attributes in C4.5. *Journal of Artificial Research*, 4, 77-90.
16. Sakthivel, N.R.; Sugumaran, V.; and Nair, B.B. (2010). Application of support vector machine (SVM) and proximal support vector machine (PSVM) for fault classification of monoblock centrifugal pump. *International Journal of Data Analysis Techniques and Strategies*, 2(1), 38-61.
17. Vapnik, V.N. (1999). An overview of statistical learning theory. *IEEE Transactions Neural Network*, 10, 988-999.