www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

# AN APPROPRIATE PROCEDURE FOR DETECTION OF JOURNAL-BEARING FAULT USING POWER SPECTRAL DENSITY, K-NEAREST NEIGHBOR AND SUPPORT VECTOR MACHINE

A. Moosavian[1], H. Ahmadi[1], A. Tabatabaeefar[1], B. Sakhaei[2]

[1]Department of Mechanical Engineering of Agricultural Machinery, University of Tehran

[2]Department of Mechanical Engineering, Sharif University of Technology

Tel: +98-937-546-8339; fax: +98-21-665-93099

Tehran, Iran

Email: a.moosavian@ut.ac.ir

***Submitted: June 20, 2012*** ***Accepted: Aug. 5, 2012*** ***Published: Sep. 1, 2012***

***Abstract- Journal-bearings play a significant role in industrial applications and the necessity of condition monitoring with nondestructive tests is increasing. This paper deals a proper fault detection technique based on power spectral density (PSD) of vibration signals in combination with K-Nearest Neighbor and Support Vector Machine (SVM). The frequency domain vibration signals of an internal combustion engine with three journal-bearing conditions were gained, corresponding to, (i) normal, (ii) corrosion and (iii) excessive wear. The features of the PSD values of vibration signals were extracted using statistical and vibration parameters. The extracted features were used as inputs to the KNN and SVM for three-class identification. The roles of PSD technique and the KNN and SVM classifiers were***

***investigated. Results showed that the accuracy rate of fault diagnosis was 100%. Also, the results demonstrated that the combined PSD-SVM model had the potential for fault diagnosis of engine journal-bearing.***

**Index terms*:* condition monitoring, power spectral density, k-nearest neighbor, support vector machine, vibration signal, fault diagnosis**

## I. INTRODUCTION

Condition monitoring of machines is gaining importance in industry because of the need to increase reliability and to decrease possible loss of production due to machine breakdown [1]. Reliability has always been an important aspect in the assessment of industrial products [25]. With the increase in production capabilities of modern manufacturing systems, plants are expected to run continuously for extended hours. Therefore, the condition monitoring of machines, especially early fault diagnosis, is proved to be necessary and has been received wide attentions in this decade [26]. Most of machinery used in the modern world operates by means of rotary parts which can develop faults. The monitoring of the operative conditions of a rotary machine provides a great economic improvement by reducing maintenance costs, as well as improving the safety level. [2-6-19].

The technique of early fault diagnosis is used to prevent serious damages in a mechanical system. Rotating machinery such as internal combustion engines, gearboxes, electromotor, pumps, and air compressors can have their vibration and acoustic emission signals monitored for early fault diagnosis [3-4-5].

Vibration analysis has been used in rotating machines fault diagnosis for decades. By measuring and analyzing the vibration of a machine, it is possible to determine both the nature and severity of the defect, and hence predict the machine's useful life or failure point [7].

It has been known for many years that the mechanical integrity of a machine can be evaluated by detailed analysis of the vibratory motion [8]. Vibration signals carry information about exciting forces and the structural path through which they propagate to vibration transducers. A machine generates specific vibrations when in a healthy state and the degradation of a component within it may result in a change in the character of the vibration signals.

In fact, each fault in a rotating machine produces vibrations with distinctive characteristics that can be measured and compared with reference ones in order to perform the fault detection and diagnosis. [9].

Journal-Bearings are multifunctional devices. In order to operate efficiently and provide long service life, journal-bearings often have to satisfy several requirements simultaneously. Journal bearings are used to carry radial loads, for example, to support a crankshaft. So, the detection and understanding of condition degradation in the journal-bearing is important for engine performance. Slavic et al. [22] have applied force measurements for identify journal-bearing faults. Saridakis et al. [23] have performed journal-bearing fault diagnosis based on Artificial Neural Network.

The analysis of vibration signals was often based on the Fast Fourier Transform (FFT). This approach suffers from some limitations. Among these limitations, the FFT is not efficient to describe the non-stationarities introduced by faults in the vibration signal. In order to overcome these performance limitations inherent to the FFT approach, many modern spectral estimation techniques have been proposed during the last two decades [10]. Power spectral density (PSD) is one of those methods that are reported by several research works [5] such as [11]. Mollazade et al. [5] have showed that PSD of frequency diagram is an effective method for fault diagnosis of external gear hydraulic pumps. In [20], PSD was applied for detection of faulty gearbox. Cusido et al. [21] have implemented combination of wavelet decomposition using PSD in induction machine diagnosis.

When full knowledge of the under lying probabilities of a class of samples is available Bayesian theory gives optimal new sample classification rates. In cases where this information is not present, many algorithms make use of the similarity among samples as a means of classification. The Nearest Neighbor decision rule has often been applied in these patters recognition problems. K-nearest neighbor (KNN) decision rule has been a ubiquitous classification method with good scalability.

SVM performs machine condition monitoring and diagnosis using its unique ability in classification process [17]. Several methods have been proposed for multi-class classification by combining several binary classifiers.

In the present work, a procedure is implemented to the detection of fault condition using vibration signals. Three conditions the journal-bearing is studied, namely, normal, corrosion and excessive wear conditions. The role of the PSD and feature extraction technique, the KNN and SVM classifiers is surveyed. The procedure is illustrated using the vibration data of an IC engine. The KNN and multi-class SVM classify the three situations of journal-bearing.

## II. MATERIALS AND METHODS

### Vibration Data

The case study for this work was a 4 cylinder internal combustion engine with power of 125 hp. Vibration signals were collected for the normal, corrosion and excessive wear conditions. Faulty journal-bearings were selected from the IC engine that was worked for a long time periods and their faults led to reduce of their efficiencies. The working speed of shaft of the engine was set at approximately 1500 rpm.

The vibration signals in frequency-domain were measured for three situations of journal-bearing by an accelerometer. Then, Root mean square (RMS) of vibration acceleration (g) was calculated for these signals. The accelerometer (VMI-102 model) was mounted horizontally on the crankcase of engine ahead main journal-bearing exactly. The sensor was connected to the signal-conditioning unit (X-Viber FFT analyzer). The software SpectraPro-4 that accompanies the signal-conditioning unit was used for recording the signals directly in the computer. The sampling rate was 8192 Hz and the number of data in each sample was 12800.

### Power Spectral Density (PSD)

Power spectral density (PSD) function indicates the vigor of the variations (energy) as a function of frequency. In other words, it shows at which frequencies variations are powerful and at which frequencies variations are weak. The PSD curve is the normal method used to describe random vibration specifications. Since the PSD curve is a plot of acceleration density, the overall rms acceleration can be found by summation of the density over frequency [12, 24].

The complex spectrum of a vibration x(t) in the time range ($t_1$, $t_2$) for any frequency f in the two-sided frequency domain (-F, +F) can be expressed as [5]:

$$X(f) = \int_{f_1}^{f_2} x(t) e^{-2\pi i f t} dt \qquad (1)$$

Generally, if FFT of vibration signal be applied, PSD can be computed directly in the frequency domain by following formula:

$$PSD = \frac{g_{rms}^{\ 2}}{f} \qquad (2)$$

where $g_{rms}$ is the Root Mean Square of acceleration in a certain frequency f [13].

**Feature Extraction**

The measured PSD values of signal were calculated to obtain the most significant features by feature extraction. The accuracy of feature extraction is of great importance since it directly affects the final diagnosis results [5]. In this work, 30 features extracted from the PSD values of vibration signals using statistical and vibration parameters. Some of used parameters are: Maximum, Minimum, Average, Root Mean Square (RMS), Standard Deviation (Stdv), Variance (Var), 5th Momentum (5th M), sixth momentum (6thM), Crest Factor, Skewness, Kurtosis, etc.

**K-Nearest Neighbor**

Some training samples are used for train the KNN rule. K nearest neighbour rule holds position of training samples and their class. When decision about new incoming data is needed, distance between query data and training samples is being calculated. Based on the defined threshold for the rule (it is the K number), K samples with least distances is selected and the class with more samples inbound is the result. In the other word, for example if there is 2 or 3 features for a classification situation, position of training samples and input sample can be visualized on 2D and 3D Cartesian coordinates. Process to find result is like to draw a circle (Sphere) centred on input location and increase radius until k samples are embed inside the circle (sphere) and then a class with more samples inbound is the result.

Without prior knowledge, the KNN classifier usually applies Euclidean distances as the distance metric. However, this simple and easy-to-implement method can still yield competitive results even compared to the most sophisticated machine learning methods [16]. The Euclidean distance between point p and q is the length of the line between them. In Cartesian coordinates, if $p_i$ and $q_i$ are two points in Euclidean $n$-space, then the distance from p to q is given by:

$$d_E = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \qquad (3)$$

## Support Vector Machine

The Support Vector Machines (SVM) have been developed by Vapnik and are gaining popularity due to many appealing features, and promising empirical performance. SVM, based on statistical learning theory, is a proper technique for solving a variety of learning and function assessment problems. Support vector machines (SVM) were originally designed for binary (2-class) classification [14]. In binary classification, the class labels can take only two values: 1 and -1.
A simple introduction of SVM is presented here for completeness. "The SVM can be considered to construct a line or hyperplane between two sets of data for classification. In case of two-dimensional situation, the action of the SVM can be explained simply without any loss of generality. The SVM tries to place a linear hyperplane between the two different classes." [18], with maximum margin (distance between closest vectors in each class and hyperplane). In other words, the SVM attempts to orient the hyperplane such that the distance between the hyperplane and the nearest data point in each class is maximal. Then the hyperplane is placed in the middle of this margin. The nearest vectors are used to define the margin and are known as support vectors. Once the support vectors are selected, the rest of the feature set can be removed, since the SVs contain all the essential information for the classifier.

## SVM Theory

Suppose label the training data $\{x_i, y_i\}, i = 1,...,l \;\; , y_i \in \{-1,1\}, x_i \in R^d$. There are some hyperplane that separates the positive (class +1) from the negative (class -1) examples. The vector $x$ which lie on the separating hyperplane satisfy $w.x + b = 0$, where $w$ is normal to the hyperplane. In the separable case, all data satisfy the following constraints:

$$w.x_i + b \geq +1,\; y_i = +1 \qquad (4)$$

$$w.x_i + b \leq -1,\; y_i = -1 \qquad (5)$$

These can be combined into the following inequalities:

$$y_i(w.x_i + b) - 1 \geq 0 \quad \forall i \qquad (6)$$

$d_+(d_-)$ is the shortest distance from the separating hyperplane to the closest positive (negative) training data. The margin of a separating hyperplane is defined to be $d_+ + d_-$. By constraints Eq.(4) and Eq.(5) , $d_+ = d_- = 1/\|w\|^2$ and the margin is simply $2/\|w\|^2$. Thus we can find the separating hyperplane which gives the maximum margin by minimizing $\|w\|^2$, subject to constraints Eq.(6). Using the Lagrange multiplier technique, a positive Lagrange multipliers $\alpha_i, i=1,...,l$, one for each of the inequality constraints Eq.(7) is determined. This gives Lagrangian:

$$\min \ell_p = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} \alpha_i y_i (x_i.w + b) + \sum_{i=1}^{l} \alpha_i \quad , \quad \alpha_i \geq 0 \qquad (7)$$

In order to deal properly with nonlinear SVM, $\ell_p$ is transformed into dual problem:

$$\max \ell_p = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i.x_j)\,,\ \alpha_i \geq 0\,,\ \sum_{i=1}^{l} \alpha_i y_i \geq 0 \qquad (8)$$

In the case where the data cannot be separated by hyperplane without errors, Vapnik propose that introducing positive slack variables $\xi_i, i=1,...,l$, the constraints become:

$$w.x_i + b \geq +1 - \xi_i \quad \text{for} \quad y_i = +1 \qquad (9)$$

$$w.x_i + b \leq -1 + \xi_i \quad \text{for} \quad y_i = -1 \qquad (10)$$

$$\xi_i \geq 0 \qquad (11)$$

The goal is to build hyperplane that makes the smallest number of errors. Hence the objection function becomes:

$$\text{minimize: } \frac{\|w\|^2}{2} + C(\textstyle\sum_i \xi_i) \qquad (12)$$

where $C$ is penalty parameter, a larger $C$ corresponding to assigning a higher penalty to errors. The $C$ must be chosen by the user. The optimization problem becomes:

$$\max \ell_D = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i.x_j)\,,\ 0 \leq \alpha_i \leq C\,,\ \sum_i \alpha_i y_i \geq 0 \qquad (13)$$

Suppose that the data is mapped to a higher dimension space (feature space), using a mapping which is called $\phi$

$$\phi : R^d \rightarrow F \qquad (14)$$

Then the training algorithm would only depend on the data through dot products in $F$, i.e. on functions of the from $\phi(x_i).\phi(x_j)$. Kernel function is the significant concept of SVM, The definition of kernel is:

$$k(x_i.x_j) = \phi(x_i).\phi(x_j) \qquad (15)$$

In below, the formulation of three kernels is given:

$$\text{Linear: } k(x_i.x_j) = x_i.x_j \qquad (16)$$

$$\text{Polynomial: } k(x_i.x_j) = (\gamma x_i.x_j + 1)^d, \gamma > 0 \qquad (17)$$

$$\text{Gaussian RBF: } k(x_i.x_j) = \exp\left(\frac{-\left\|x_i - x_j\right\|^2}{2\sigma^2}\right) \qquad (18)$$

So the optimization problem of nonlinear SVM is:

$$\max \ell_D = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j k(x_i.x_j), 0 \le \alpha_i \le C \ , \ \sum_i \alpha_i y_i \ge 0 \qquad (19)$$

After solving this optimization problem, those points which $\alpha_i > 0$ are called 'support vectors'. Then they determine $w$ by Eq.(20). And $b$ can be found by KKT 'complementarily' condition Eq.(21), where $s_i$ are support vectors and $N_s$ is the number of support vectors.

$$w = \sum_j^{N_s} \alpha_j y_j \phi(s_j) \qquad (20)$$

$$\alpha_i (y_j (w.\phi(s_j) + b) - 1) = 0 \qquad (21)$$

Finally, the class of $x$ is:

$$sign(w.x + b) = sign\left(\sum_j^{N_s} \alpha_j y_j k(s_j, x) + b\right) \qquad (22)$$

**Multi-class Support Vector Machine**

In the real world, we deal more than two classes for examples: in condition monitoring of rotating machineries there are several classes such as journal-bearing faults, mechanical unbalance, misalignment [17], different load conditions, gear faults etc. There are different methods for multi-class SVM such as one-against-all (OAA), one-against-one (OAO), etc.

Among various multi-class SVM, the OAA method is earliest implementation for SVM multi-class classification. It builds $k$ SVM models where $k$ is the number of classes. The $i$th SVM is trained with all of examples in the $i$th class with positive labels, and all the other examples with negative labels [17]. Thus given $l$ training set $(x_1, y_1), ..., (x_l, y_l)$ , where $x_i \in R^n, i = 1,...,l$ and $y_i \in \{1,...,k\}$ is the class of $x_i$, the $i$th SVM solve the following problem:

$$\text{minimize: } \frac{1}{2}\left\|w^i\right\|^2 + C\sum_{i=1}^{1}\xi_j^{\,i}(w^i)^T \tag{23}$$

$$\text{subject to: } (w^i)^T\phi(x_j) + b^i \geq 1 - \xi_j^{\,i} \quad \text{if} \quad y = i \tag{24}$$

$$(w^i)^T\phi(x_j) + b^i \leq -1 + \xi_j^{\,i} \quad \text{if} \quad y \neq i \tag{25}$$

$$\xi_j^{\,i} \geq 0 \quad , \quad j = 1,\ldots,l \tag{26}$$

where the training data $x_i$ is mapped to a higher-dimensional space by function $\phi$ and $C$, is the penalty parameter and $\xi$, is the slack variable.

Minimizing Equation (23) means to maximize $2/\|w_i\|$. When data is not separable, there is a penalty term $C\sum_{i=1}^{l}\xi_{i,i}$, which can reduce the number of training errors.

Finally, $x$ is in the class which has the largest value of the decision function $w^T\phi(x)+b$ [15].

We have implemented these methods by using SVM Toolbox. The software is available at http://asi.insa-rouen.fr/~gloosli/simpleSVM.html

## III. RESULTS AND DISCUSSION

The aim of this paper is the main journal-bearing fault detection using signal processing and classification techniques. The paper consists of three stages: signal processing, feature recognition and extraction, and decision-making. Always, the signal analysis must be done first, and then the information gathered in that process will be used towards decision making. In order to make an appropriate decision regarding fault diagnosis, it is essential to have a priori knowledge of the journal-bearing type of engine, operating conditions such as crankshaft rotating speed, defect types and the severity of each defect type.

Figure 1 shows the samples of the PSD diagram of vibration signals acquired for different conditions of the journal-bearing. By attention to this Figure, it can be seen that the maximum value of the PSD is lower in healthy situation. Also, the PSD values of vibration signal for excessive wear condition are more than other conditions.

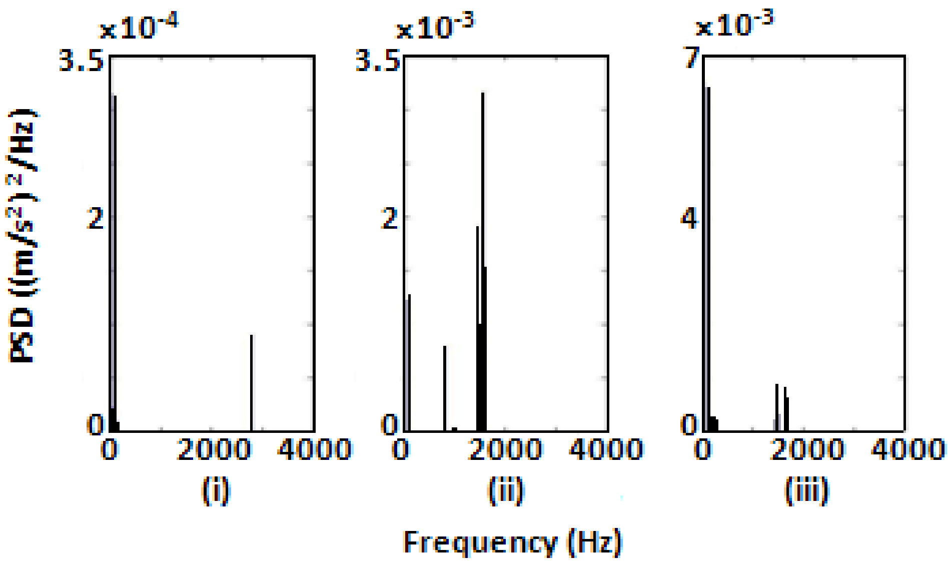


Figure 1. PSD- Frequency diagrams of journal-bearing in three conditions: (i) normal (ii) corrosion (iii) excessive wear

Feature extraction directly affects final diagnosis results. 30 features extracted were used as inputs to the classifiers. Constructed data set for each class divided into two parts, training and testing data. In the present work, the training process is using 60 data to train the KNN and SVM parameter, and the classification process is used to classify with 60 data for each journal-bearing condition in the constant rotating station [27, 28]. The detailed description of the data set is shown in Table 1.

Table 1. Description of journal-bearing data set

| Journal-Bearing condition | Number of training samples | Number of testing samples | Label of classification |
|---|---|---|---|
| Normal | 30 | 30 | 1 |
| Corrosion | 30 | 30 | 2 |
| Excessive wear | 30 | 30 | 3 |

In this work, we studied the accuracy rate of the KNN and SVM. The classifiers were trained by training data set, and then their performance was exactly estimated by testing data set [30]. Also, their accuracy rate was compared. It must be stressed that the most significant criterion for evaluating the performance of these classifiers is their classification success rate.

**Performance Comparison of the KNN and SVM**

The KNN is a classifier that its accuracy is always 100% on training data set, because the KNN holds the position of training data set and their class during the classification process. The test success of the KNN just depends on the K value. The variable K value was used between 1 and 10 for the KNN.

The performance of the SVM depends on the choice of the kernel function to map a data from original space to a higher dimensional feature space. There are no certain rules governing its choice. The RBF kernel is the one of the best kernel for constructing SVM and provides excellent results for fault diagnosis applications [1, 29]. We considered the penalty parameter $C$ of $10^3$, condition parameter for QP method (lambda=$10^{-7}$), RBF kernel and OAA multi-class method for the SVM. The values of the RBF kernel width (σ) was selected in the range of 0.1 to 1.

Sensitivity, specificity and total classification accuracy are three criteria to determine the test performance of classifiers. These criteria are defined as:

Sensitivity: number of true positive decisions/number of actually positive cases.

Specificity: number of true negative decisions/number of actually negative cases.

Total classification accuracy: number of correct decisions/ total number of cases.

Among these criteria, total classification accuracy is commonly applied in fault diagnosis applications [30, 31]. Table 2 shows the total classification results for different journal-bearing conditions.

Table 2. Performance of the KNN and SVM

| No | KNN (K), accuracy | SVM (σ), accuracy |
|---|---|---|
| 1 | (1), 85.7% | (0.1), 100% |
| 2 | (2), 82.51% | (0.2), 98.64% |
| 3 | (3), 79.03% | (0.3), 95.33% |
| 4 | (4), 76.56% | (0.4), 91.78% |
| 5 | (5), 76.06% | (0.5), 89.51% |
| 6 | (6), 74.08% | (0.6), 86.22% |
| 7 | (7), 70.64% | (0.7), 85.33% |
| 8 | (8), 53.74% | (0.8), 85.01% |
| 9 | (9), 49.26% | (0.9), 84.66% |
| 10 | (10), 33.33% | (1), 83.33% |

The performance of the KNN was in the range of 33.33% to 85.7%. The accuracy rate of OAA method was 100% on both training and testing data set. This accuracy was occurred in the RBF kernel width of 0.1 (σ=0.1). Figure 2 and 3 shows the performance variations of the KNN and SVM, respectively.

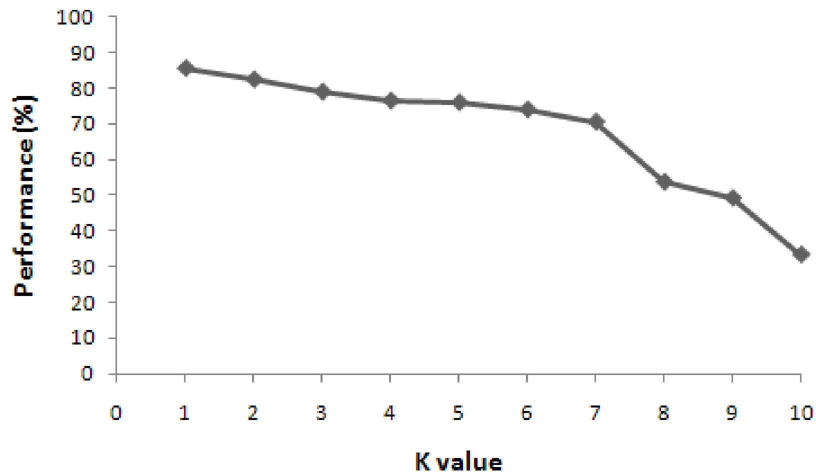


Figure 2. Accuracy rate of the KNN under different K values

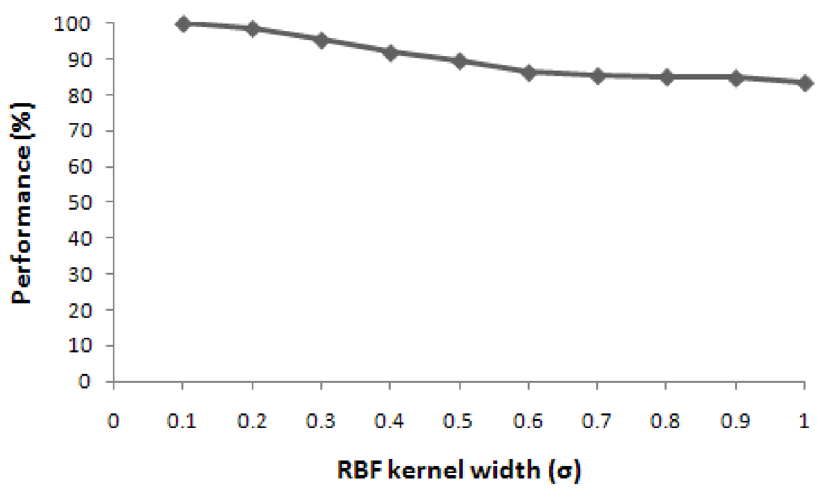


Figure 3. Accuracy rate of the SVM under different values of RBF kernel width

By attention to this result, it can be found that the PSD is the one of powerful signal processing technique. The performance of the SVM was substantially better than the KNN, but the computation time of the KNN was relatively lower than the SVM. Finally, it be seen that the

SVM classifies the different faults of journal-bearings perfectly. It's due to the unique ability in classification problems. It is to be stressed here that the SVM becomes famous and popular in machine learning community due to the excellence of generalization ability than the traditional method such as neural network [17]. The perfect accuracy rate was due to apply the proper diagnostic method. Therefore, it can be concluded that it is essential to apply the proposed intelligent system for the fault detection of main engine journal-bearing in order to increase accuracy and reduce errors caused by subjective human judgment [31].

## IV. CONCLUSION

A combined Power Spectral Density (PSD), K-Nearest Neighbor (KNN) and Support vector Machine (SVM) have been presented to perform fault diagnosis of a journal-bearing. Three states of journal-bearing was detected, namely, normal, corrosion and excessive wear conditions. Firstly, the PSD values of the vibration signal of journal-bearing were calculated from obtained spectrums. Secondly, 30 features were extracted as inputs to classifier. Finally, the structure of KNN and SVM classifiers was built by feeding the training set and then its performance was estimated by test set. Euclidean distance and one-against-all (OAA) multi-class method were used for the KNN and SVM. The best classification accuracy of three conditions of journal-bearing was 85.7% and 100% on test set for the KNN and SVM, respectively. The results show that the SVM is a stronger technique for fault diagnosis of rotating machinery. Also, the results demonstrate the ability and reliability of proposed the PSD-SVM model in diagnosing main engine journal-bearing faults.

## ACKNOWLEDGEMENT

Acknowledgment is made to the especially thanks for University of Tehran and Irankhodro Powertrain Co. for their concentration for this research.

## REFERENCES

[1] B.Samanta, K.R.Al-Balushi and S.A.Al-Araimi, "Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection'', Engineering Applications of Artificial Intelligence 16, 2003, pp. 657– 665.

[2] O.Castro, C.Sisamon and J.Prada, "Bearing fault diagnosis based neural network classification and wavelet transform", Proc. of the 6th WSEAS International Conference on Wavelet Analysis & Multirate Systems, Bucharest, Romania, 2006, pp. 16-18.

[3] Z.K. Peng, and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography'', Mechanical Systems and Signal Processing, Vol. 18, 2004, pp. 199–221.

[4] H.Zheng, Z.Li and X.Chen, "Gear fault diagnosis based on continuous wavelet transform'', Mechanical systems and Signal Processing, Vol. 16 (2–3), 2002, pp. 447–457.

[5] K.Mollazade, et.al, "An Intelligent Combined Method Based on Power Spectral Density, Decision Trees and Fuzzy Logic for Hydraulic Pumps Fault Diagnosis'', International Journal of Electrical and Computer Engineering 3:8, 2008, pp. 551-563.

[6] H.Ahmadi and A.Moosavian, "Fault Diagnosis of Journal-Bearing of Generator Using Power Spectral Density and Fault Probability Distribution Function'', Communications in Computer and Information Science, 2011, pp. 30–36.

[7] P.A.Laggan, "Vibration monitoring", IEE Proc proceeding Colloquium on Understanding your Condition Monitoring, 1999, pp. 1-11.

[8] R.C.Sr.Eisenmann, "Machinery Malfunction Diagnosis and Correction", Prentice Hall, 1998.

[9] S.Pöyhönen, P.Jover and H.Hyötyniemi, "Independent component analysis of vibration for fault diagnosis of an induction motor", Proc. on the IASTED International Conference on Circuits, Signals, and Systems (CSS), Mexico, Vol. 1, 2003, pp. 203-208.

[10] L.B.Jack and A.K.Nandi, "Fault detection using support vector machines and artificial neural network augmented by genetic algorithms'', Mechanical Systems and Signal Processing, Vol 16, 2002, pp. 373- 390.

[11] R.B.Gibson, "Power Spectral Density: a Fast, Simple Method with Low Core Storage Requirement'', M.I.T. Charles Stark Draper Laboratory Press, 1972, 57 pages.

[12] T.Irvine, "An Introduction to Spectral Functions'', Vibration Data Press, 1998.

[13] T.Irvine, “Power Spectral Density Units: [G2/Hz]’’, Vibration Data Press, 2000.

[14] C.Cortes and V.Vapnik, “Support-vector network’’, Machine Learning, 20: 1995, pp. 273-297.

[15] A.Widodo and B.S.Yang, “Review support vector machine in machine condition monitoring and fault diagnosis’’, Mechanical systems and signal processing, 2007, pp. 2560–2574.

[16] Y.Song, et.al, "IKNN: Informative K-nearest neighbor pattern classification", Proc. Springer-Verlag Berlin Heidelberg, LNAI 4702, 2007, pp. 248–264.

[17] A.Widodo and B.S.Yang, “Support vector machine in machine condition monitoring and fault diagnosis’’, Mechanical Systems and Signal Processing, Vol. 21, Issue 6, 2007, pp. 2560-2574.

[18] Z.Wu, et.al, “Automatic Digital Modulation Recognition Based on Support Vector Machines’’, Neural Networks and Brain, ICNN&B '05, Vol. 2, 2005, pp. 1025-1028.

[19] M, Khazaee, et.al, “Fault diagnosis and classification of planetary gearbox of MF285 tractor final drive using Fast Fourier Transform (FFT), Stepwise Backward Selection and support vector machine (SVM) classifier, Elixir Control Engg. 43, 2012, pp. 6974-6977.

[20] K.Heidarbeigi, et.al, "Fault diagnosis of Massey Ferguson gearbox using power spectral density", Journal of Agriculture Technology 5(1), 2009, pp. 1-6.

[21] J.Cusido, et.al., “Detection in induction machines by using power spectral density on the wavelet decompositions,” MCIA Research Group, Universitaty Politecnica de Catalunya. C. Colom 1, 08222 Terrassa, Spain: Catalunya, 2008, pp. 1-7.

[22] J.Slavic, et.al., “Typical bearing-fault rating using force measurements: application to real data,” Journal of Vibration and Control, 17(14), 2011, pp. 2164-2174.

[23] K.M.Saridakis, "Fault Diagnosis of Journal Bearings Based on Artificial Neural Networks and Measurements of Bearing Performance Characteristics", Proc. of the 9th International Conference on Computational Structures Technology, Civil-Comp Press, Stirlingshire, UK, 2008, Paper 118.

[24] H.Ahmadi and P.Salami, “Using of Power Spectral Density for Condition Monitoring of Fan”, Modern Applied Science, Vol. 4, No. 6, 2010, pp. 54-59.

[25] G.Niu, et.al, “Decision-level fusion based on wavelet decomposition for induction motor fault diagnosis using transient current signal”, Expert Systems with Applications 35, 2008, pp. 918–928.

[26] B.Bagheri, H.Ahmadi and R.Labbafi, “Implementing discrete wavelet transform and artificial neural networks for acoustic condition monitoring of gearbox”, Elixir Mech. Engg. 35, 2011, pp. 2909-2911.

[27] L.Zhang, G.Xiong, H.Liu, H.Zou and W.Guo, “Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference”, Expert Systems with Applications 37, 2010, pp. 6077-6085.

[28] Y.Lei, Z.He, Y.Zi and Q.Hu, “Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs”, Mechanical Systems and Signal Processing, Vol. 21, Issue. 5, 2007, pp. 2280-2294.

[29] B.S.Yang, T.Han and W.W.Hwang, “Fault diagnosis of rotating machinery based on multi-class support vector machine”, Journal of Mechanical Science and Technology 19(3), 2005, pp. 846-859.

[30] E.Ebrahimi and K.Mollazade, “Intelligent fault classification of a tractor starter motor using vibration monitoring and adaptive neuro-fuzzy inference system”, Insight, Vol. 52, No. 10, 2010, pp. 561-566.

[31] S.G.Jolandan, H.Mobli, H.Ahmadi, M.Omid and S.S.Mohtasebi, “Fuzzy-Rule-Based faults classification of gearbox tractor”, WSEAS Transactions on Applied and Theoretical Mechanics, Vol. 7, Issue. 2, 2012, pp. 69-82.