FAILURE DIAGNOSTICS WITH SVM IN MACHINE MAINTENANCE ENGINEERING

Krisztián DEÁK[1], Imre KOCSIS[2], Attila VÁMOSI[3], Zoltán KEVICZKI[4]
[1]deak.krisztian@eng.unideb.hu
[2]kocsisi@eng.unideb.hu
[3]vamosi.attila@eng.unideb.hu
[4]zoltan.keviczki@schaeffler.com

Abstract— Failure diagnostics as a part of condition monitoring (CM) technique is inevitable in modern industrial practice. Condition Based Maintenance (CBM) identifies all problems that cause further failures and suggests maintenance periods. Reducing maintenance costs and enhancing system availability are largely depends on information provided by precise and accurate failure diagnostics. The approach can be used widely in the several field of the industry. Data acquisition is related to measurement then data processing, feature extraction is needed, finally failure identification. In this paper Support Vector Machine (SVM) is discussed how to be used for diagnosing machines and machine elements. The aim of using SVM is to diagnose the system at a certain moment or predict its actual state in the future. SVM is progressing rapidly several new advances are revealed as the part of machine learning techniques. Due to experiments SVM efficiency could be approximately 90% or even higher.

***Keywords*—**Bearing test-rig, failure diagnostics, machine fault, machine learning, support vector machine

I. INTRODUCTION

It is critical to maintain machine health and extending the lifetime of its critical processes, minimizing emerging costs during operation. Diagnosing the impending failures is one of the significant trend, prognosis of the remaining useful lifetime (RUL) of the equipment is another major challenge in schedule maintenance operations so that maximize reliability. Scheduled maintenance practices tend to reduce machine lifetime and increase down-time, resulting in loss of productivity. Intelligent monitoring system can reduce maintenance costs, avoid catastrophic failures and increase machine availability. To develop an effective diagnostic and prognostic system deep understanding of the bearing behavior is well advised. Condition based maintenance (CBM) as a leading maintenance technique involves monitoring machine condition and predicting machine failure. It is critical as a part of preventive maintenance to reduce costs and increase reliability. Condition monitoring (CM) is an effective method of monitoring machine state parameters like vibration, temperature, wear debris analysis and noise monitoring.

Lot of approaches have been already defined in maintenance engineering to determine and monitor the actual state of rotary machines, reciprocating machines, and electrical machines. Bearings, gears as basic rotary machinery are successfully monitored by vibration signal analysis, cylinder pressure signal can be monitored in internal combustion engines and reciprocating compressors, measuring of power consumption and heat distribution are typical for monitoring electrical machines. Further methods include noise measurement, oil wear debris analysis, infrared spectrum analysis, current analysis, power analysis.

II. FAILURE DIAGNOSTIC

All failure diagnostics consist of three main steps, data acquisition and collection, data processing, and failure pattern recognition. (Fig. 1.)

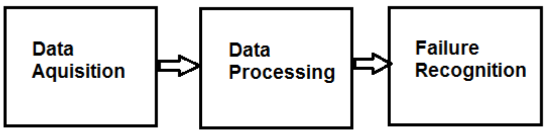

Fig. 1. Failure diagnostic stages

After the measurement process the raw signal frequently contains lot of noise, its signal-to-noise (SNR) ratio is low. In this form it is impossible to obtain the necessary information about the signal first noise should be eliminated or reduced. Statistical based features for example rms value, peak value, mean value, variance, kurtosis, crest factor are typical features to be extracted from the signal. Time domain features in machine fault diagnosis is published by several researchers. [8] Mostly, they are applied for investigating non-periodic signals and initial faults of both bearings and gears. For frequency domain analysis Fast Fourier Transform (FFT) is commonly applied that extract features, used for periodic signals, it can determine defects of bearings and gears. The main drawback of Fast Fourier Transform (FFT) is that it is only applicable for transforming a stationary signal. For non-stationary signal the Short-time Fourier Transform [14] or the Wavelet transform [4] can be applied. Feature extraction is domain specific and signal specific. There are some ways available to determine the cross-correlation between features, classical statistics such as F-test, Chi-squared test, variety of methods available to evaluate feature performance, are applied to test the performance of each individual feature [10] and the relief algorithm is another classical method [6]. Feature extraction is necessary in machine fault diagnosis and general approach is to obtain as many features as possible because more features give more reliable information. But lot of additional features make the computational time and cost more and the ability of the diagnostics system to make efficient diagnosis is decreased, moreover irrelevant features complicate the whole diagnostics system reducing its capability to make effective analysis.

Fig. 2. shows the so-called peaking phenomenon [10] increasing the number of features can only improve the performance initially, but after a critical number of features, the performance decreases. Theoretically, infinite data sets would enhance the capability of diagnostics model but having so large models are nearly impossible in real conditions.

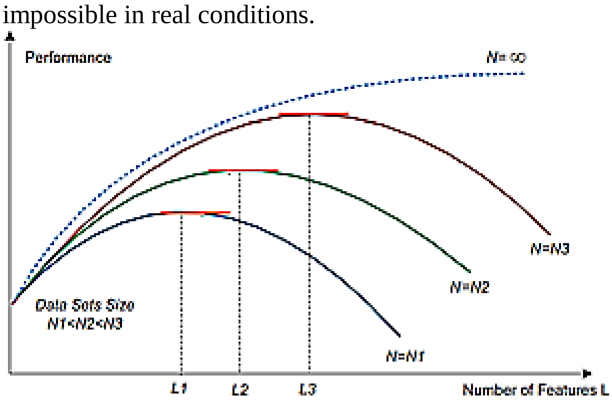

Fig. 2. Peaking phenomenon [18]

In order to increase the performance of the diagnostics model feature selection is necessary. In this process the number of selected features is decreasing to a value that even acceptable to accomplish proper further diagnosis. Removing the irrelevant features is one of the most accepted method in this field. Mainly subset feature selection and individual feature selection are used as two basic approaches for feature selection. Researchers [12] defined that each features should be ranked on priority and less important one to be removed the others are applied. SVM weight all features in the input space, and these weights can be evaluated during the training process [9]. More important features get higher score of weights, less important ones get smaller values or nearly zero. SVM algorithm was invented by Vladimir N. Vapnik and soft margin was proposed by Corinna Cortes and Vapnik in 1993 and published 1995. SVM has been successfully applied to a number of applications ranging from particle identification, face identification, and text categorization, to engine knock detection, bioinformatics, and database marketing [2].

As for further possible applications there are researches about optimization e.g. the Optimization of the Shape of Axi-Symmetric Rubber Bumpers [20].

As it was mentioned that appropriate number of features is well advised to optimize calculation time and machine fault diagnosis system. Classifier error rate measures; distance measures; information measures; dependence measures; and consistency measures [3] are commonly used for measuring the performance of the model, classifier error rate measures are mostly debated. Searching for minimum subset of features is exhaustively searched field. Actual condition of the machine can be represented by the features extracted from the data sets. Determining the predefined threshold is a major task feature’s value above a predefined threshold may imply a possible failures, severity of the failure is indicated by the deviation from the predefined threshold.

Generally, threshold value is evaluated by experience, the value that differ the intact status of the machine from the faulty condition. The Kurtosis, the fourth normalized statistical moment, corresponds to the peak value of the data. For an undamaged bearing, the value is equal to three in frequencies. Kurtosis value is more useful, when it is compared with the RMS, crest factor, and peak value. The Crest Factor is the ratio of the peak acceleration to the RMS value. Crest factor is a good indicator of small size defects; although, when localized damage grows, the value of the crest factor decreases significantly because of the increasing RMS. So, kurtosis of 3 can be as threshold to make difference between health and faulty bearings. In real industrial circumstances there is no enough data to determine this value in all working conditions referring to all machines in different operating environments. K-nearest neighbour [10], Bayesian classifier, learning algorithms as classical pattern recognition techniques are frequently applied to determine threshold or boundary from available data. Learning algorithm finds the optimal decision function automatically from the available data and discriminate different patterns.

Non-Linear classifiers such as Artificial Neural Networks (ANN) and SVM can be used in supervised and unsupervised learning. Classifying data is a common task in machine learning. Given data points belong to one of two classes, and the goal is to decide which class a new data point is in. To define support vector machines and first linear classifiers, a data point is viewed as a p-dimensional vector a list of p numbers, and the purpose to decide whether it can be separated such with a $(p-1)$-dimensional hyperplane. This is called a linear classifier. Best hyperplane should be chosen that represents the largest separation or margin in other word, between the two classes. In this case the distance from the hyperplanes to the nearest data point on each side is maximized. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier it defines is known as a maximum margin classifier.

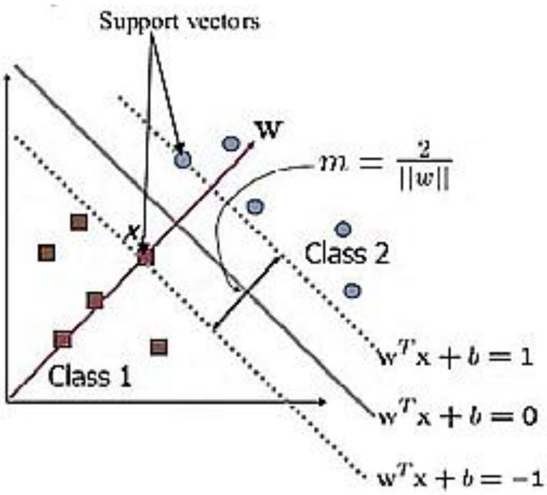

Fig. 3. SVM classification principle [18]

Supervised learning pairs the input and output data. Feature vectors extracted from the signal can be used as input data. SVM is sensitive to noise in the signal especially when raw data is used to the input of the SVM system. The performance of the trained SVM can be evaluated by using a set of test data. Internal coefficients are chosen by the supervising learning. Decreasing the real output and the predicted output is important to get the training result as the optimal decision.

III. ONE CLASS SUPPORT VECTOR MACHINES

Correct maintenance actions is only possible if diagnostics of the machinery is satisfy the requirements. The purpose of this action is to reduce maintenance cost and enhance reliability and system availability. The Support Vector Machine (SVM) method can be interpreted as a transformation to put the lower dimensional data to a higher dimension space. Support vector machine constructs a hyperplane or set of hyperplanes in a high or infinite-dimensional space, which can be used for classification, regression. The hyperplanes in the higher dimensional space are defined as the set of points whose dot product with a vector in that space is constant. SVM gets non separable patterns to separated patterns the existing failure or incipient failure is getting more identifiable because failure diagnostics is in the higher dimensional space. The core method of SVM is to use the maximal margin method to defeat the overfitting problem that makes the model fit special data sets. Small sample size problems are solved by using the maximal margin approach.

Support Vector Machine (SVM) is a state-of-the-art method, frequently used as nonlinear classifier or learning algorithm which is able to evaluate automatically dependency between data and defined as a regression problem. SVM estimate the connection between predictive variables and explanatory variables. Maximal margin approach and kernel method are combined in SVM to make prediction as Fig. 4. presents. Support Vector Machine (SVM) is a classification and regression method. Support Vector machines uses hypothesis space of a linear functions in a high dimensional feature space. SVM can be trained with a learning algorithm from optimization theory.

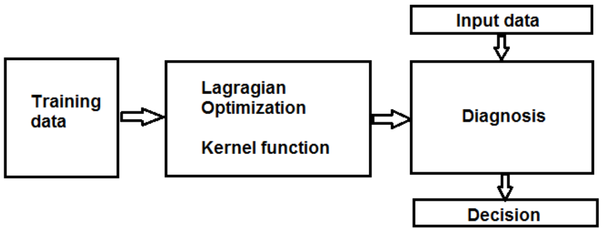

Fig. 4. SVM based diagnostics classifier architecture

Good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class, the functional margin. The SVM decision function is an application of the kernel function and Lagrangian optimization method is used to obtain the optimal decision function from the training data [7]. SVM is generally suitable for two-class tasks.

The decision function is used to predict the output for a given input. The maximal margin method is applied to improve the accuracy of the prediction.

For machine learning algorithms, the kernel trick is a way of mapping observations from a general set S into an inner product space V (equipped with its natural norm), without having to compute the mapping explicitly. It can transform the problem from a lower dimension to a higher dimension, while the computation complexity does not change. Transforming the problem from a lower to a higher dimension makes the approximation function more flexible with its data, reducing the risk of empirical error.

With fewer SVs (support vectors, data taking effect) the generalization ability is improved. Furthermore, as the decision function is comprised of SVs, having fewer SVs can reduce the computation complexity.

The optimal solution of the SVM is achieved by the use of a quadratic optimization problem. The convex property of the formulation makes the solution unique. The SVM utilizes the Lagrangian optimization method to solve this problem.

The original optimal hyperplane algorithm proposed by Vapnik in 1963 was a linear classifier. The original problem is defined in finite dimensional space. Sometimes the sets to discriminate are not linearly separable in that space. However, in 1992, Bernhard E. Boser, Guyon and Vapnik suggested a way to create nonlinear classifiers by applying the kernel trick to maximum-margin hyperplanes. The resulting algorithm is formally similar, except that every dot product is replaced by a nonlinear kernel function. Polynomal kernels, Gaussian radial basis function are used frequently.

In statistical learning theory the problem of supervised learning is formulated as follows. Taking a set of training data $\{(x_i, y_i)\}$ in $R_n \times R$ sampled according to unknown probability distribution P(x, y), and a loss function V(y, f(x)) that measures the error, for a given x, f(x) is predicted instead of the actual value y. The problem consists in finding a function f that minimizes the expectation of the error on new data that is, finding a function f that minimizes the expected error [16]:

$$\int \mathbf{V(y.f(x))P(x,y)dx\,dy}$$

Maximum margin is given as [15]-[17]:

$$\mathbf{margin = arg\,min\,d(x) = arg\,min \frac{|\ x \cdot w + b\ |}{\sqrt{\sum_{i=1}^{d} w_i^2}}}$$

For calculating the SVM we see that the goal is to correctly classify all the data. For mathematical calculations we have,

If $\mathbf{y_i = +1;}$ $\quad \mathbf{wx_i + b \geq 1}$
If $\mathbf{y_i = -1;}$ $\quad \mathbf{wx_i + b \leq 1}$
For all i; $\quad \mathbf{y_i\,(w_i + b) \geq 1}$

In this we present the QP formulation for SVM classification [15]-[17]-[18]-[19]. This is a simple representation only.

SV classification:

$$\mathbf{\min_{f,\xi_i} \|f\|_K^2 + C\sum_{i=1}^{l} \xi_i}$$

$\mathbf{y\ if\ (x_i) \geq 1 - \xi_i}$**, for all i** $\quad \mathbf{\xi_i \geq 0}$

SVM classification, Dual formulation:

$$\mathbf{\min_{\alpha_i} \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j)}$$

$\mathbf{0 \leq \alpha_i \leq C}$**, for all i;** $\quad \mathbf{\sum_{i=1}^{l} \alpha_i y_i = 0}$

Variable ξ_i is called slack variable and it measures the error made at point (x_i, y_i). Training SVM becomes quite challenging when the number of training points is large. A number of methods for fast SVM training have been proposed [15]-[17]-[19].

IV. Experimental Study

In general, SVM applications can be classified into two categories: failure diagnostics such as novelty detection and multi-failure discrimination; and secondly reliability data analysis, such as reliability prediction and system reliability assessment. There are some examples where SVM can be used in maintenance engineering.

Planetary gearboxes are widely used in machine industry and other heavy-duty industry such as helicopters, heavy-duty trucks, and other large-scale machinery. Planetary gearboxes are able to undertake heavy-duty tasks with high torque need.

There are two approach for investigating machine element, bearings or gear failures, life-tests and short-term tests by creating artificial faults on the bearing elements. Manual pitting damage experiments were designed and implemented to provide vibration data corresponding to different levels of gear pitting damage. In our researches, spark erosion was created with EDM or laser machine. To validate the fuzzy approach test- rig had been built which has 5 individual roller bearings. One of them was monitored by measuring vibration acceleration level. The list of measuring instruments that are used in the measurement:

- PCB 603C01 accelerometer
- Bearing test-rig (Fig. 5.)
- NI 9234 DAQ
- NI Labview Sound and Vibration Measurement
- TESTO 476 stroboscope
- Soundbook measurement system

Fig. 5. Bearing test rig for validation

For the measurement written in this paper, PCB603C01 acceleration sensor was used. It was attached strongly on the top of the bearing house with magnet providing enough force to keep the sensor and produce ideal circumstances for further operations. Both Soundbook vibration measurement system, Samurai Software, and National Instrument (NI) devices were used for the measurement, namely NI DAQ 9234. NI DAQ 9234 has a 4-Channel, ±5 V, 51.2 kS/s per Channel, 24-Bit IEPE device with 51.2 kS/s per-channel maximum sampling rate; ±5 V input, 24-bit resolution; 102 dB dynamic range; antialiasing filters. Software-selectable AC/DC coupling; AC-coupled (0.5 Hz), software-selectable IEPE signal conditioning (0 or 2 mA), smart TEDS sensor compatibility, NIST-traceable calibration are the typical features of this device used in the bearing vibration measurements. Different virtual instruments were built in Labview software to make vibration and spectrum measurement. The simplest applied is showed in Fig. 6.

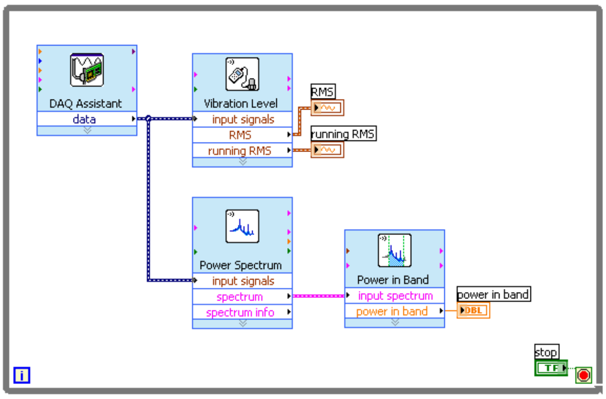

Fig. 6. Virtual Instrument (VI) for vibration measurement

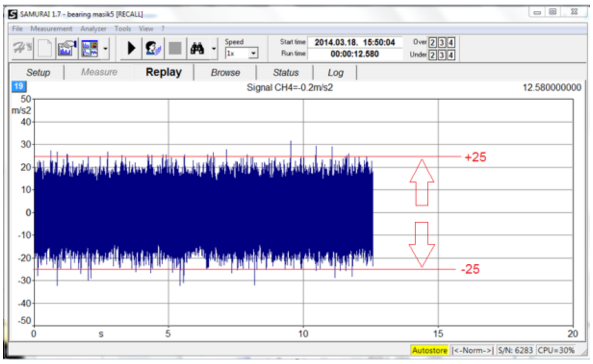

Fig. 8. Time domain spectrum of the measured bearing

In this research, spark erosion was created with laser machine (Fig. 7.) then geometrical measured under Olympus BX61 optical microscope. Artificial spall was 658 um in diameter and 257 um deep on the inner ring of the 6206 type bearing. Being a plastic cage bearing balls could be removed easily before creating the fault. BPFI was calculated, basically by 2880 1/ min engine speed, exact rotation speed was measured and BPFIs were recalculated.

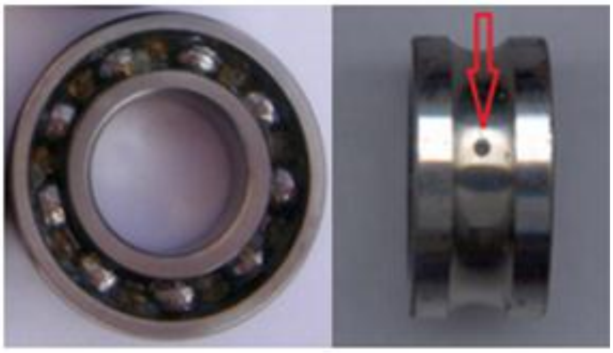

Fig. 7. Artificial bearing fault on the inner race

BPFO (ball passing frequency outer race), BPFI (ball passing frequency inner race), FTF (fundamental train frequency), BSF (ball spin frequency) can be calculated in Hz with the following formulas by knowing the contact angle (Φ), rotation speed (f), number of rollers, outer ring diameter (D), and inner ring diameter (d) of the bearing. According to calculations engine speed is 48 Hz, BPFO 171,84 Hz, BPFI 260,16 Hz, FTF 19,2 Hz, BSF 112,32 Hz by using the ball bearing in the validation test. Fig. 8. indicates the time domain spectrum of the bearing, vibration value is normal.

Frequency-domain (FFT) analysis of the vibration signal is the most widely used approach of detecting roller-bearing defects. The interaction of defects in rolling element bearings produces impulses. Impulses generate the natural frequencies of bearing elements and housing structures, that causes peaks in the vibration energy spectrum. Comparing the calculated fault frequencies with the measured values the exact faults can be revealed. Fig. 9. presents the vibration spectrum of the bearing.

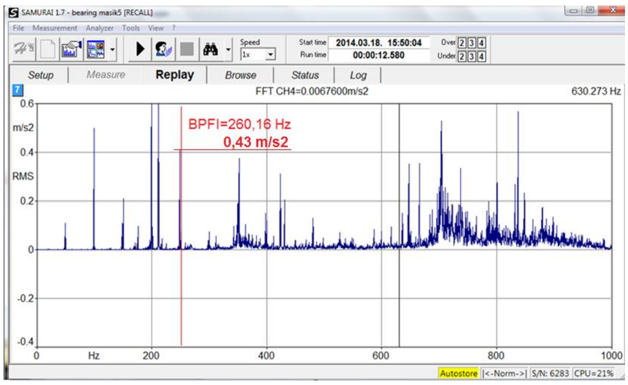

Fig. 9. Frequency domain spectrum of the measured bearing

The general algorithm of SVM based diagnosis can be seen on Fig. 10. Test and training data fed the system and make performance evaluation by using features.

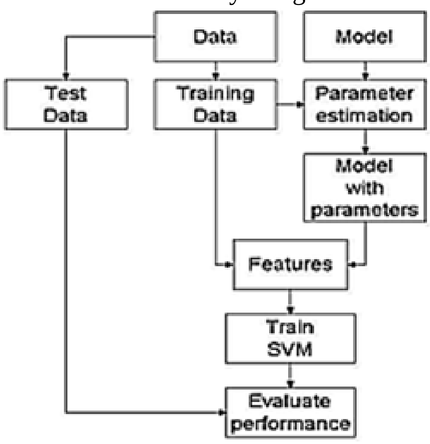

Fig. 10. Algorithm for bearing fault detection with SVM [21]

Diagnosis rate approximately 90% was reached in the experiments thus SVM method could be applied efficiently in machine fault diagnostics. Several measurements were performed by applying bearings with different faults. SVM diagnostics system is able to distinguish healthy and faulty bearings. Further experiments are planned to investigate outer race faults, cage faults and ball faults. Motor Current Signature Analysis, MCSA tests are planned as well, a method for monitoring machinery in the industry by means of its electrical currents. Motor current signature analysis (MCSA) has proven to be a highly valuable predictive maintenance tool.

V. CONCLUSION

Condition Monitoring (CM) is vital part in maintenance engineering. Proper maintenance can reduce cost and improve reliability. Lot of methods are used to diagnose the machines or their components such as rotary machineries e.g. bearings, gears. These machine elements tend to suffer serious damages that are measured by mostly vibration measurement techniques because of its simplicity. Additional measurements can be used such as temperature measurement, wear debris analysis, oil analysis to create a complex diagnostics condition monitoring system. SVM has been successfully applied to a number of applications in industry. Classifying data is a common task in machine learning, linear and non-linear classification is applied used hyperplane to separate points belongs to two classes. Kernel trick is a way of mapping observations from a general set S into an inner product space V. It can transform the problem from a lower dimension to a higher dimension. SVM has lot of application in maintenance engineering. The SVM has been used to diagnose failure in rolling element bearings, induction motors, machine tools, pumps, compressors, valves, turbines, and various other machines. Engineering experiments show the feasibility and effectiveness of this method, the diagnosis rate nearly 90% or even more so SVM method can be applied efficiently in many investigations. Test rigs measurements are proved the efficiency of SVM tools in machine fault diagnosis. Later as part of further experiments bearing with outer race faults, cage faults, ball faults will be examined. Moreover, Motor current signature analysis (MCSA) could be applied as a modern tool of machine diagnostics. Multiplied measurements using both vibration and motor current signature analysis is possible to enhance the efficiency of the system, even temperature measurement could be used to make more accurate measurements.

REFERENCES

[1] B.SREEJITH, A.K.VERMA & A.SRIVIDYA 2008. Fault diagnosis of rolling element bearing using time-domain features and neural networks. Proceedings of Third IEEE International Conference on Industrial and Information Systems ICIIS 2008. Kharagpur,India.

[2] BENNETT, K. P. & CAMPBELL, C. 2000. Support Vector Machines: Hype or Hallelujah? SIGKDD 2.

[3] DASH, M. & LIU, H. A. 2003. Consistency-based search in feature selection. Artificial Intelligence, 151, 155-176.

[4] DAUBECHIES, I. 1990. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. Ieee Transactions on Information Theory, 36, 961-1005.

[5] GRIFFIN, D. W. & LIM, J. S. 1984. Signal Estimation from Modified Short-Time Fourier-Transform. Ieee Transactions on Acoustics Speech and Signal Processing, 32, 236-243.

[6] KIRA, K. & RENDELL, L. A. Year. A practical approach to feature selection. In: ML92 Proceedings of the ninth international workshop on Machine learning, 1992 San Francisco, CA, USA. Morgan Kaufmann

[7] LUENBERGER, D. G. & YE, Y. 2008. Linear and nonlinear programming, New York, Springer.

[8] MATHEW, J. & ALFREDSON, R. J. 1984. The Condition Monitoring of Rolling Element Bearings Using Vibration Analysis. Journal of Vibration Acoustics Stress and Reliability in Design-Transactions of the Asme, 106, 447-453.

[9] NGUYEN, M. H. & DE LA TORRE, F. 2010. Optimal feature selection for support vector machines. Pattern Recognition, 43, 584-591.

[10] THEODORIDIS, S. & KOUTROUMBAS, K. 2006. Pattern recognition, San Diego, CA, Academic Press.

[11] Y.KIM, E., C.C.TAN, A., YANG, B.-S. & KOSSE, V. 2007. Experimental Study on Condition Monitoring of Low Speed Bearings:Time domain Analysis. 5th Australasian Congress on Applied Mechanics, ACAM2007. Brisbane,Australia.

[12] YU, L. & LIU, H. 2004. Efficient feature selection via analysis of relevance and redundancy. Journal of Machine Learning Research, 5, 1205-1224.

[13] ZHANG, Y. X. & RANDALL, R. B. 2009. Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram. Mechanical Systems and Signal Processing, 23, 1509-1517.

[14] ZHU, X. L., BEAUREGARD, G. T. & WYSE, L. L. 2007. Real-time signal estimation from modified short-time Fourier transform magnitude spectra. Ieee Transactions on Audio Speech and Language Processing, 15, 1645-1653

[15] Burges C., “A tutorial on support vector machines for pattern recognition”, In “Data Mining and Knowledge Discovery”. Kluwer Academic Publishers, Boston, 1998, (Volume 2).

[16] Theodoros Evgenuiu and Massimilliano Pontil, Statistical Learning Theory: a Primer 1998.

[17] Nello Cristianini and John Shawe-Taylor, “An Introduction to Support Vector Machines and Other Kernel-based Learning Methods”, Cambridge University Press, 2000.

[18] J.P.Lewis, Tutorial on SVM, CGIT Lab, USC, 2004.

[19] Vapnik V.,”Statistical Learning Theory”, Wiley, New York, 1998.

[20] Mankovits Tamás, Szabó Tamás, Kocsis Imre, Páczelt István: Optimization of the Shape of Axi-Symmetric Rubber Bumpers, STROJNISKI VESTNIK-JOURNAL OF MECHANICAL ENGINEERING 60:(1) pp. 61-71. (2014)

[21] Karthik, Nataraj, Biswanath: Model based bearing fault detection using support vector machines. Annual Conference of the Prognostics and Heath Management Society, 2009.