NUEJ
NARESUAN UNIVERSITY
ENGINEERING JOURNAL

MONITORING SCREW FASTENING PROCESS BASED SUPPORT VECTOR MACHINE CLASSIFICATION

Supawan Ponpitakchai

Department of Electrical and Computer Engineering, Naresuan University, Thailand
supawanpo@nu.ac.th

ABSTRACT

In industrial manufacturing, costs of production can be reduced by detecting the failures of running process. Automated systems are widely applied for detecting the failures instead of human inspection. This research develops a Support Vector Machine (SVM) classifier, which can be integrated in an automated system, for detecting the faults of screw fastening on the cover of hard disk. The data of screw fastening can be distinguished in 2 patterns; complete and incomplete fastening. These data are necessary to be pre-processed for developing the classifier, which 4 methods of the pre-processing data are proposed. The SVM classifier is also investigated through three popular kernel functions;linear, polynomial and radial basis (RBF), with the proper values of their corresponding parameters, and the results in terms of missed class percentage are compared. The experiments show that SVM performs as an efficient classifier for all pre-processing data methods and the proposed kernel functions. The SVM classifier using RBF kernel with appropriate method of data processing has the best results; therefore, it is strongly recommended for the implementation of self-automated screw monitoring machine.

KEY WORDS: SVM, kernel methods, machine learning, screw fastening and classification

1. INTRODUCTION

In industrial manufacturing, there is a need to reduce costs of production. One major possibility is to detect failures from running processes by using automated systems, and reports the errors to quality control systems. Human operators are typically used for inspection the processes. However, errors of human inspection can be occurred due to the difficulty and/or complexity of running processes, and fatigue from long working hours. Therefore, automated monitoring strategies are desirably applied for detecting the failures instead of human as the following examples.

Firstly, three computer-aided systems, which are neural networks, fuzzy logic systems and fuzzy c-means, are studied in [1]and [2]. These algorithms are applied to control automated screw fastening process where the torsion signal is measured in real time. The proposed analytical models are compared to obtain thir advantages and disadvantages under the fastening process. Secondly, the idea of weightless neural networks is employed to detect the screw fastening assembly in automated process [3].This research develops fault detection strategies using in automated self-tapping screw insertions through weightless neural network, which has the advantageous characteristics comparing to conventional neural networks. Next, the screw insertion process is automately monitored with on-line identification technique; Newton-Raphson [4]. In this study, the completeness of a threaded insertion is predicted through the relation of torque signal and insertion depth [5]. This prediction technique estimates two important parameters (friction and screw properties) online. This technique can be used to predict insertion signal and develop automated monitoring algorithms. Finally, the screw tightening process in a 3.5 inch hard disk drive assembly is investigated via a three dimensional finite element analysis in [6] and [7]. This research focuses on the effect of a screw tightening sequences on the torque loosening and the hard disk top cover deformation. The proper screw fastening sequences as well as the applied torques can reduce the looseness and prevent the top cover slip.

Therefore, this research applies Support Vector Machines (SVMs) to detect a running process (fastening top covers of hard disks with tiny screws) of hard disk manufacturing. In the process, incomplete fastenings, which the screws are not fully inserted in the holes, can be occurred and cause floating screws on the top covers. At present, the floating screws are inspected and corrected by human operators by looking at each screw straightforward. However, the human investigation possibly has errors because of the tiny size of screw and fatigue from long period of working hours. This study applies SVM as the classifier to monitor the errors of screw fastening process. This classifier can be integrated in screw fastening machines as a part of self automated monitoring system, which can use to detect incomplete fastenings instead of human.

SVM is one of the most well known learning systems. It is based on kernel methods which is increasingly popular and widely used in many applications such as handwritten digit recognition, text categorisation, time serie prediction, and investigation of DNA microarray data [8]. SVM algorithm are different from other learning techniques as their frameworks are based on a strong mathematical theory rather than on loose analogies (such as neural networks) [9]. The data is mapped into a high dimensional space where it is easier to separate the data into two classes by a linear classifier [10]. Moreover, the ease of use, theoretical appeal, and remarkable performance have made SVM to be a widely used technique for many learning problems [8].

To the author knowledge, there is no application of SVM for detecting screw fastening processes. Therefore, the main contribution of this research is to develop a classifier with SVM for automated monitoring the screw fastenings from driver motor current data. This data has two types (complete and incomplete fastenings) which have significantly difference appearance. This study applies three popular kernel functions, which are linear, polynomial and radial basis functions (RBF), in SVM algorithm to classify these two types of screw fastenings. The kernel functions with different values of their corresponding parameters are investigated in the experiments and the results in terms of missed class percentage are compared. This study also applies four methods for preparing data to classify in the experiments. Moreover,the best SVM classifier with pre-processing data method is recommended for the implementation of screw monitoring machine.

This research is organised as follows; the framework of SVM and theoretical support are given in the next section. Then, the screw fastening process and patterns of complete and incomplete fastening data are presented in Sections 3.. Finally, Section 4. and 5. describe the experimental design along with the results and conclusions respectively.

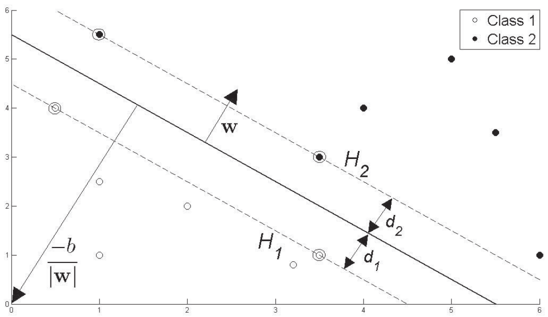


Figure 1: Hyperplane for two linearly separable classes. [12]

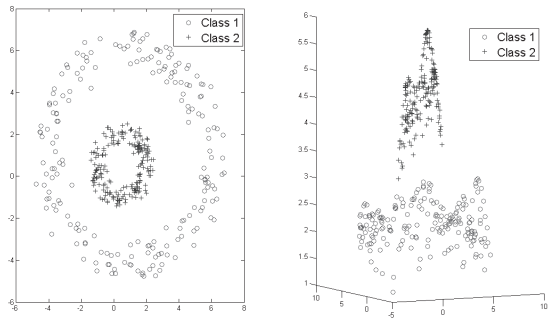


Figure 2: Data mapping to a feature space using radial basis kernel. [12]

2. SUPPORT VECTOR MACHINES

SVM is employed as a monitoring tool or classifier in this research. The screw fastening data is classified into 2 categories; complete and incomplete fastening at each end of fastening process. Suppose that the set of given fastening data consists of two classes of objects. Then, a new object is presented and we have to assign it to one of the two classes. This is a classification problem and can be stated as follows: we are given empirical data $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_i, y_i) \in \mathcal{X} \times \{\pm 1\}$, and we want to estimate a decision function $f : \mathcal{X} \longrightarrow \{\pm 1\}$ [11]. Note that, $\mathcal{X}$ is a nonempty set from which the patterns $\mathbf{x}_i$ are taken, the y_i are called labels or classes. According to our data, the screw fastening is devided to 2 types; complete and incomplete fastening. The classifier framework for 2 classifier or binary classifier needs to be explained in the next section.

2.1 Binary classifiers

Assume that the given data has D attributes (or D dimensions) and is linearly separable, and then we can draw a line to separate that two classes. The line is called hyperplane when the data are more than 2 dimensions. The hyperplane can be presented by

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \tag{1}$$

where $\mathbf{w}$ is a weight vector normal to hyperplane and b is a threshold. The optimal hyperplane is shown as a solid line which is oriented as far as possible from the closest members of both classes. The data closest to the hyperplane is called support vectors. Figure 1 shows a hyperplane which separates 2 classes of data. The optimum hyperplane can be constructed by selecting $\mathbf{w}$ and b such that the following conditions are satisfied.

$$\begin{aligned} \mathbf{x}_i \cdot \mathbf{w} + b &\geq +1 \quad \text{for } y_i = +1 \\ \mathbf{x}_i \cdot \mathbf{w} + b &\leq -1 \quad \text{for } y_i = -1 \end{aligned} \tag{2}$$

Then, we have $y_i(\mathbf{x} \cdot \mathbf{w} + b) - 1 \geq 0 \quad \forall_i$. The geometry of hyperplane shows that the margins (d1 and d2) are equal to $\frac{1}{\|\mathbf{w}\|}$ and we need to maximise this distance subject to the constrain in (2). So it is ideal to minimise the following problem:

$$\min \|\mathbf{w}\| \quad \text{such that} \quad y_i(\mathbf{x} \cdot \mathbf{w} + b) - 1 \geq 0 \quad \forall_i. \tag{3}$$

This constrained optimisation problem (3) is dealt by introducing Lagrange multipliers $\alpha_i \leq 0$ ($\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_m)$) and a Lagragian [11]

$$L(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{m} \alpha_i (y_i(\langle \mathbf{x}_i, \mathbf{w} \rangle + b) - 1). \tag{4}$$

Then we have dual optimisation problem:

$$\max_{\boldsymbol{\alpha} \in \mathbb{R}^m} \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j K_{ij} \tag{5}$$

$$\begin{aligned} \max_{\boldsymbol{\alpha} \in \mathbb{R}^m} \quad & \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j K_{ij} \\ \text{subject to} \quad & \alpha_i \geq 0 \ \forall_i \ i = 1, \cdots, m \\ \text{and} \quad & \sum_{i=1}^{m} \alpha_i y_i = 0 \end{aligned} \tag{6}$$

where $K_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$ is a dot product and the decision function can be written as

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{m} y_i \alpha_i \langle \mathbf{x}_i, \mathbf{x} \rangle + b \right). \tag{7}$$

Note that, b is computed through (7). Therefore, the new fastening data can be predicted its class by using (7). For the optimal hyperplane, the data points which lie closest to the hyperplane have $\alpha_i > 0$, and these points are the support vectors whereas all other points have $\alpha_i = 0$. Therefore the representation of the hyperplane is given by these points. The data which are not support vectors do not influence the position and orientation of the separating hyperplane [13].

2.2 Nonlinear SVM

The optimal hyperplane given in the previous section is formulated in a dot product space. However, this is not enough to deal with many interesting problems such as nonlinear problems. Then, we need to represent our input as vectors in space $\mathcal{H}$ called the feature space using a map $\Phi : \mathcal{X} \rightarrow \mathcal{H}$ where $x \mapsto \mathbf{x} = \Phi(x)$.

Transforming the data into $\mathcal{H}$ has a benefit in studying the learning algorithms by using linear algebra. It means that a nonlinear separable data can be recast into a higher dimensional space (feature space), which linear separator can be performed. This is illustrated in Figure 2, where the left shows input data and the right shows data mapping in the feature space. Kernel functions $k : \mathbf{X} \times \mathbf{X} \longrightarrow \mathbf{R}$ have to have the property called positive definite function such that [14]

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j k(x_j, x_i) \geq 0 \tag{8}$$

for all $n \in \mathbb{N}$, $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ and all $x_1, \ldots, x_n \in \mathbf{X}$. Some examples of kernel functions are presented below [14]:

1. Linear kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$

2. Radial basis kernel $K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2}$
3. Polynomial kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^b$
4. Taylor kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \sum_{n=0}^{\infty} a_n \langle \mathbf{x}_i, \mathbf{x}_j \rangle$ for $a_n \geq 0$
5. Exponential kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(\langle \mathbf{x}_i, \mathbf{x}_j \rangle)$
6. Binomial kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (1 - \langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle)^{-b}$
7. Fourier kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \prod_{i=1}^{d} f(\mathbf{x}_i - \mathbf{x}_j)$ for $f(t) = \sum_{n=0}^{\infty} a_n \cos(nt)$ and $a_n \geq 0$

Note that, σ and b are parameters defining the kernel's behaviour.

To transform our input into feature space, we replace $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$ by $K(\mathbf{x}_i, \mathbf{x}_j)$. This substitution called the kernel trick that develops nonlinear SVM. Hence, the decision function (7) can be rewritten as

$$f(\mathbf{x}) = \mathrm{sgn}\left(\sum_{i=1}^{m} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_j) + b\right). \quad (9)$$

An example of SVM using in nonlinear separable data presents in Figure 3 where radial basis function is selected as a kernel function. Hyperplane is the solid line and support vectors are the data lies on the dot lines.

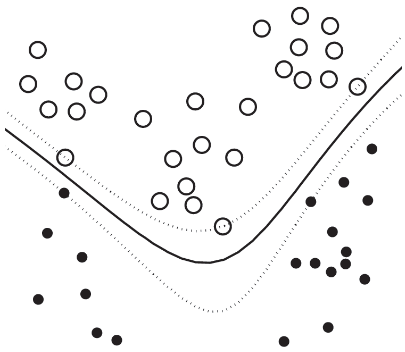


Figure 3: SVM classifier using radial basis function kernel. [11]

3. SCREW FASTENING PROCESS

In hard disk manufacturing, there is a process to fasten top covers of hard disks with tiny screws by fastening machine (PEDIII Controller) shown in Figure 5. Incomplete fastenings, which the screws are not fully inserted in the holes, can be occurred and causes floating screws on the top covers shown in Figure 4 (left) where the scientifically characteristic of the floating screws are presented in Figure 6(a). At present, human operators investigate and correct the floating screws by looking at each screw straightforward. However, the human investigation possibly has errors because of the tiny size of screw and fatigue from long period of working hours.

During the fastening process, the PEDIII Controller can record the driver motor current data that is proportional to the torque of

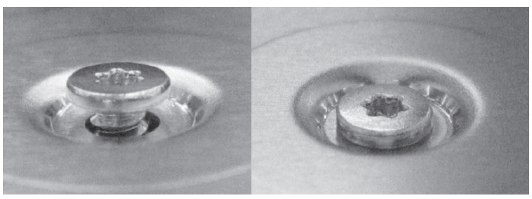
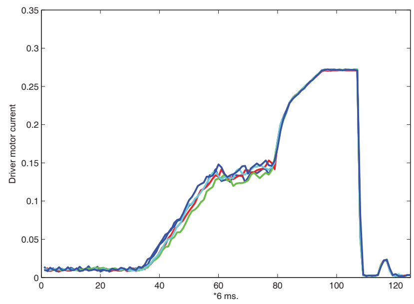
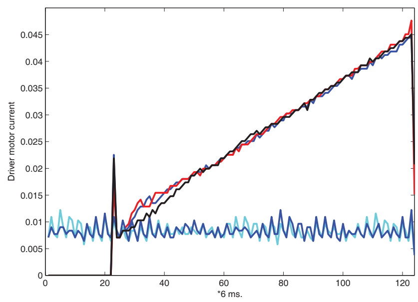


Figure 4: Incomplete (left) and complete (right) screws in fastening process.

Figure 5: Screw tightening controller



(a)



(b)

Figure 6: Time series data for complete (a) and incomplete (b) screw fastenings.

screw fastening. The data is saved to the memory of PEDIII Controller directly without additional filter or sensors, which can be opened with Microsoft Excel. SVM algorithm is implement in MatLab programming where we can import the torque data from Excel. The objective of MatLab simulation is to investigated the best type of pre-processing data and the best kernel function with an appropriate value of kernel parameters. These simulation results can be used to construct automated screw fastening machine in future.

The driver motor current data mensioned earlier is sampled at every 6 milliseconds for each screw fastening. These data can be categorised into 2 types; complete and incomplete fastenings which are presented in Figure 6(a) and 6(b) respectively.

Figure 6(a) is 4 selected torque data of 4 screw where complete fastening happens. Noise of each screw fastening is included in each data which can be seen from the figure. In Figure 6(b), the 4 different lines shows torque of fastening in case of incomplete fastening. The characteristic of incomplete fastening has two type; increase without any steady current and swing at low current. However, the figure shows that these 2 types of data have obviously different patterns that we can create a classifier to separate them.

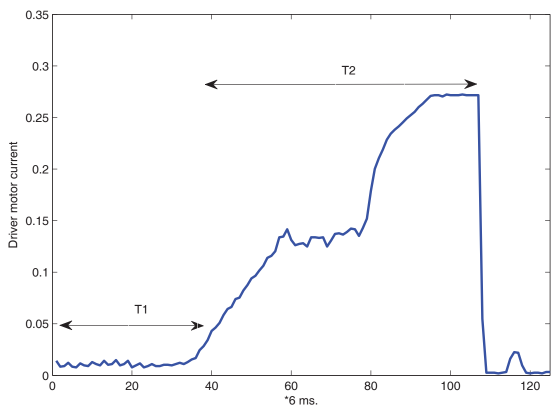


Figure 7: Features T1 and T2 of the driver motor current.

4. EXPERIMENTAL DESIGN AND RESULTS

The data taken in the previous section are used for training (400 screws) and validation (100 screws). In each screw fastening, 125 points of current data are recorded every 6 milliseconds which means each data has 125 attributes.Then, the data is in the form $(\mathbf{x}_{ij}, y)$ where $i = 1, \cdots, 400$ or 100, $j = 1, \cdots, 125$ and $y = \pm 1$. Y is a type or a class of the screw fastening assigned to be 1 for complete and -1 for incomplete screws. Four techniques of pre-processing data are applied as shown below:

1. Data 1: normalised (125 attributes),
2. Data 2: normalised and reduced by averaging to be 62 attributes,
3. Data 3: normalised and reduced by averaging to be 30 attributes,
4. Data 4: normalised and used 5 specific features which are T1, T2, area, A1, and maximum current (5 attributes).

The normalised process of Data 1-4 is to transform the current data to be in the range of [-1, 1]. After normalising, Data 2 and 3 are reduced the dimensions of current data by averaging to be 62 and 30 attributes respectively, while Data 4 uses the specific features of the current data as 5 attributes.

The features T1 of data 4 is the lower steady period of current data as shown in Figure 7 where T2 is the period from the end of T1 to the end of fastening. The feature area is the area under the current data curve. A1 is the rotating degrees from start to the screw seating level, and the maximum current is the maximum value of 125 attributes. T1, T2 and A1 are recorded data in PEDIII controller where area and maximum current are calculated in MatLab programming.

To classify the data with SVM, there are two steps to carry out which are training and validation. In training, we firstly need to define kernel functions and their corresponding parameters which are selected by trial and error method. In these experiments, three types of kernel functions; linear, polynimial and radial basis are selected due to their difference in complexity. Then the hyperplane or classifier is investigated from the optimisation problem (6) using training data. The optimum value of $\boldsymbol{\alpha}_i$ and b, which are used to create the hyperplane, are calculated via the quadratic programming of MatLab. The expression of optimum hyperplane is (9) also known as a decision function or classifier.The performance of decision function in previous step is validated with validation data and presented in term of missed class percentage. The lower percentage represents the better classifier, and the 0 percent means that SVM classifier can perfectly separate the complete and incomplete screw fastenings.

The missed class percentage of SVM using linear kernel is presented in Table 1. It reveals that the first three pre-process methods (Data 1, 2 and 3) have no different performance as they have equal percentage of miss class (1.28%). However, their performances are better than Data 4 as they have the lower percentage.

Table 1: Missed class percentage in SVM using linear kernel

Kernel function	Data 1	Data 2	Data 3	Data 4
Linear kernel	1.28	1.28	1.28	8.98

Table 2 shows the percentage of missed class in SVM using polynomial kernel. The results show the percentage with different values of parameter b, where increasing value of parameter b creates more complexity of the classifier. From the experiment, more complexity of classifier is not necessary to produce good result (low value of missed class percentage). Thus, the parameter b should be selected at proper value. From the table, b equals to 1 or 2 produces the good results for Data 1-3 while using b equals to 3 produces the good results only for Data 2-3. Data 4 performs good results for all b values except b equal to 1. Especially using b equals to 2 or 3 for Data 4 produces the best results (0% of missed class). This means that Data 4 with b values equals to 2 or 3 can perfectly separate the complete and incomplete screw fastenings using polynomial kernel.

Table 2: Missed class percentage in SVM using polynomial kernel

b	1	2	3	4	5
Data 1	1.28	1.28	10.29	10.29	10.29
Data 2	1.28	1.28	1.28	10.29	10.29
Data 3	1.28	1.28	1.28	10.29	10.29
Data 4	8.98	0	0	1.28	1.28

Using RBF kernel, the results are presented in Table 3 with different values of parameter σ. The complexity of classifier is affected by value of σ where too high and too low value can reduce the performance of classifier. For each data sets, the proper value of σ generate 0 percent of missed class. For instance, σ for zero percentage have to be 3 or 4 for Data 1 . The best result is achieved by using RBF kernel with Data 4 that the classifier generate zero percentage with several values of σ. Therefore, SVM classifier using RBF produces the best results for all sets of data, when the proper value of σ is selected. However, Data 4 with RBF kernel and their proper value of σ should be used in implementation of screw monitoring machine. The calculation time for Data 4 is shorter than the others as the data has the lowest number of attributes.

Table 3: Missed class percentage in SVM using RBF kernel

σ	0.01	0.1	0.6	1	2	3	4
Data 1	10.29	10.29	8.98	1.28	1.28	0	0
Data 2	10.29	10.29	1.28	1.28	0	1.28	1.28
Data 3	10.29	10.29	0	0	1.28	1.28	1.28
Data 4	2.56	0	0	0	0	0	8.98

Comparing among all studied functions, SVM classifier using RBF kernel produces the best results for all types of data especially for Data 4, when the proper value of σ is selected. From the result, using linear kernel and polynomial kernel have a lower performance of classification than RBF kernel. Because, linear kernel is too simple function where as polynomial kernel is too complex function. Therefore, Data 4 with RBF kernel and their proper value of σ (0.1, 0.6, 1, 2, or 3) should be used for the implementation of screw monitoring machine. This is because the calculation time for Data 4 is shorter than the other data types as it has the lowest number of attributes.

5. CONCLUSIONS

This research develops a SVM classifier for using in automated monitoring of the screw fastenings, which are performed at the cover of hard disk, instead of using human operators. The driver motor current data are collected for training and validation and are pre-processed into 4 types of data with different number of attributes. SVM is used to classify these data types with three kernel functions and the different values of their corresponding parameters. The experiments show that SVM performs as an efficient tool for monitoring the screw fastenings as most results have the percentage of missed class between 0-1.28% for the proposed kernel functions. Moreover, several results have 0% of missed class depending on the techniques of pre-processing data and the proper values of kernel parameters especially for Data 4 (normalised and used five specific features of current data) and RBF kernel.

From the experimental results, this study strongly recommends the SVM classifier using RBF kernel with Data 4 for the implementation of screw monitoring machine. Normally, fastening data in the form of driver motor current is sampled via data acquisition card (DAQ card). Then, the sampling current data can pass to a computer using program interfaces such as LabView. The program can transfer each current data with the selected preprocessed method (recommended method of Data 4) and then classify with the selected SVM classifier (recommended RBF kernel). After that, the program will shows the result at the end of each screw fastening that it is complete or incomplete.

In order to apply to another size of screw, we need to collect data of screw fastening firstly. These data will be used in training and test SVM algorithm to find the appropriate kernel function with appropriate value of kernel parameter. The best kernel function and value of kernel parameter are used in testing process to ensure performance of SVM by observed at missed class percentage. Finally, the automated screw monitoring system can be implemented using SVM classifier with the appropriate kernel function with appropriate value of its parameter.

6. ACKNOWLEDGMENT

The author would like to acknowledge National Electronics and Computer Technology Center, National Science and Technology Development Agency and Industry/University Cooperative Research Center (I/UCRC) in HDD Component, the Faculty of Engineering, Khon Kaen University for funding support. The author also would like to thank Hitachi Global Storage Technology for the research topic, experimental data and assistance.

7. REFERENCES

[1] A. Grauel, L. A. Ludwig, and g. Klene, "Comparison of different intelligent methods for process and quality monitoring," *International Journal of Approximate Reasoning 1997*, vol. 16, pp. 89–117, 1997.

[2] N. Dhayagude, Z. Gao, and F. Mrad, "Fuzzy logic control of automated screw fastening," *Robotics&Computer-Integrated Manufacturing*, vol. 12, no. 3, pp. 235–242, 1996.

[3] L. D. Seneviratne and P. Visuwan, "Weightless neural network based monitoring of screw fastening in automated assembly," in *Intelligent Robots and Systems, 1999.IROS' 99. Proceedings*, vol. 1, pp. 561–566, 1999.

[4] M. Klingkajay and P. Toncharoen, "The intelligence automated thread-fastening assembly using the newton raphson," *World Academy of Science, Engineering and Technology*, vol. 42, pp. 127–131, 2008.

[5] M. Klingajay and N. I. Giannoccaro, "The automated threaded fastening based on on-line identification," *International Journal of Advanced Robotic Systems*, vol. 1, no. 4, pp. 263–272, 2004.

[6] K. Suknikhom, T. Kamnerdtong, and P. Jongpradist, "Study of screw tightening sequence on the looseness of the top cover in the hard disk drive assembly," *The First TSME International Conference on Mechahical Engineering*, 2010.

[7] K. Suknikhom, T. Kamnerdtong, and S. Jongpradist, P.and Chutima, "Effects of screw fastening sequence to top cover loosening in hard disk drive assembly," *Applied Mechanics and Materials*, vol. 110-116, pp. 3534–3540, 2012.

[8] N. Cristianini and B. Schölkopf, "Support vectoc machines and kernel methods: The new generation of learning machines," *AI Magazine*, vol. 23, no. 3, 2002.

[9] S. Gunn, "Support vector machines for classification and regression," tech. rep., University of Southampton, 1998.

[10] C. Campbell, "Kernel methods: a survey of current techniques," *Neurocomputing*, vol. 48, pp. 63–84, 2002.

[11] B. Schölkopf and A. Smola, "Kernel methods and support vector machines," in *Encyclopedia of Biostatistics* (P. Armitage and T. Colton, eds.), John Wiley and Sons, 2003.

[12] T. Fletcher, "Support vector machines explained," 2008.

[13] C. Campbell, "An introduction to kernel methods," in *Radial Basis Function Networks: Design and Applications* (C. Campbell, R. Howlett, and L. Jain, eds.), pp. 155–192, Physica Verlag, 2002.

[14] I. Steinwart and A. Christmann, *Support Vecter Machines*. Information Science and Statistics, Springer, 2008.