International Journal of the Physical Sciences Vol. 6(27), pp. 6295-6302, 2 November, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.104
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

Applications of support vector machines in oil refineries: A survey

Mahmoud Reza Saybani[1]*, Teh Ying Wah[2], Amineh Amini[3], Saeed Reza Aghabozorgi Sahaf Yazdi[4] and Adel Lahsasna[5]

Department of Information Systems, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia.

Accepted 2 September, 2011

Support vector machine has been explored and many applications found within various research areas and application domains. Many support vector machine techniques have been specifically developed for certain application domains. This paper is an attempt to provide an overview on applications of support vector machines within the oil refineries to the professionals inside oil refineries, researchers and academicians. This paper has grouped and summarized applications of support vector machines within various units inside refineries. Application of support vector machines to a particular domain within refineries can be used as guidelines to assess the effectiveness of the support vector machines in that domain. This survey provides a better understanding of the different applications that have been developed for one area which allows finding of applications in other domains.

Key words: Support vector machines, data mining, machine learning, oil refinery, oil refining.

INTRODUCTION

Oil refineries are vital to daily life, industries and economies, and in general they are of importance to the maintenance of civilization itself.Refineries are large and complex and with millions of records of data, decision making is very crucial. Data mining techniques, such as support vector machine (SVM) are very useful for this major industry. Despite the fact that SVMs have found some interesting usage within oil refineries, there exists no survey on where, why and how SVMs have been used in the process of oil refining. To fill this gap, researchers decided to investigate this issue, and this paper is the result of the study. This survey is an attempt to provide an overall view of the usage of SVMs within oil refining process, authors of this paper not only present the techniques and algorithm used, but also bring the application domain within the oil refining process for which applications were developed. This survey is based on a search in the keyword index and publications and information for support vector machines and oil refineries on the following search engines: ACM digital Library, IEEE Xplore, Science Direct e-Books, Scopus, Springer Link, and Web of Science.

SUPPORT VECTOR MACHINES THEORY

There are many resources about definition of SVM, however, most of them are difficult to follow, therefore in order to understand SVM easier, this paper attempted to combine definitions and explanations, from these resources (Laura, 2008; Nationalbank, 2007; Statsoft, 2010; SVMS.org, 2010; Wikipedia, 2010). SVMs are a set of related supervised learning technique from the field of

*Corresponding author. E-mail: Saybani@gmail.com or Saybani@siswa.um.edu.my.

Abbreviations: ARIMA, Auto regressive integrated moving average; **ANN-MLP,** artificial neural network multilayer perceptron; **BPNN**, back-propagation neural network; **C-SVM,** classification **SVM; DCS**, distributed control system; **DEA,** data envelopment analysis; **DMUs**, decision making units; **KNN**, K-nearest neighbor; **LDA**, linear discriminant analysis; **LS-SVM,** least Squares support vector machine; **NIR,** near infrared; **PLS,** partial least squares; **PNN**, probabilistic neural network; **PSO,** particle swarm optimization**; QDA,** quadratic discriminant analysis; **QP,** quadratic programming**; RBF,** radial basis function; **RDA**, regularized discriminant analysis; **SIMCA,** soft independent modeling of class analogy; **SVM,** support vector machine.

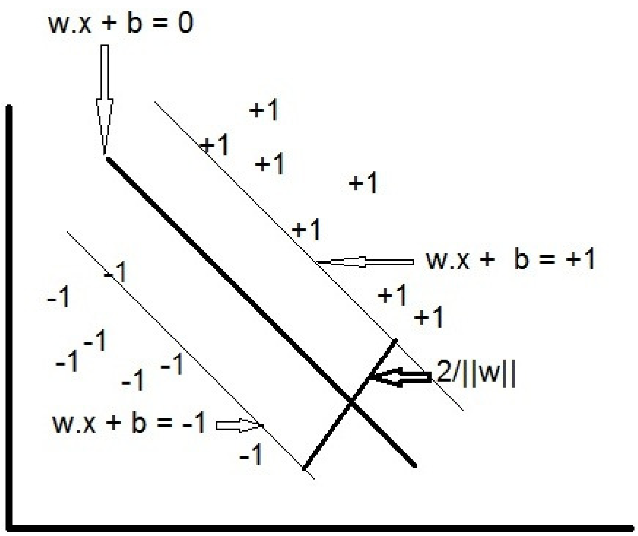

Figure 1. Margin maximization.

machine learning that analyze data and recognize patterns, which is used for classification and regression analysis. Its root can be found in the statistical learning theory (SVMS.org, 2010) which originally was invented by Vapnik and colleagues at AT&T Bell Laboratories in 1995, however the current standard "soft margin" was proposed by Cortes and Vapnik (Wikipedia, 2010).

SVM is a non-linear classifier, and the main concept behind it is to non-linearly map the input data to some high dimensional space, where the data can be linearly separated. This provides greater classification (or regression) performance. SVM is an excellent tool for data classification and regression (Chunming Wu, 2010). SVM is one of the data mining techniques, which is also mentioned in the connection with machine learning.

SVM learns the relationship between input and output data, let's consider a set of training dataT = $\{(x_1,y_1), (x_2,y_2),\ldots,(x_n,y_n)\}$, where $x_i \in \Re^m$ is m-dimensional input vector and $y_i \in$ {-1, +1} is the discrete output vector, each y_i can be either 1 or –1, indicating the class to which the point x_i belongs. SVM constructs hyperplanes in a multidimensional space which separates cases of different classes. SVM supports both regression and classification tasks. An optimal hyperplane can be produced through an iterative training algorithm which is used to minimize an error function.

The goal is to find the maximum-margin hyperplane that divides the points having $y_i = +1$ from those having $y_i = -1$. Any hyperplane can be written as a set of points x satisfying:

$$\hat{y} = f(x,w,b) = w^T.x + b = 0 \quad (1)$$

where $\hat{y}$ is the predicted output, x is the input pattern, w is weight vector which is normal to the hyperplane and b is bias. Weight and bias have to be set during the training process. Thinking of geometric, calculating the values of the parameters *w* and *b* means looking for a hyperplane that best separates class $y_i = +1$ from class $y_i = -1$ based on some criterion (Venkoparao et al., 2009). SVMs maximize the margin between the two classes, which is also known as margin-maximization. The margin is the distance between the hyperplanes bounding each class, where in the hypothetical perfectly separable case no observation may lay. Maximizing the margin allows us to select the two hyperplanes of the margin in a way that there are no points between them and then try to maximize their distance; in other words we search for the classification function that can most safely separate the classes. Figure 1 represents a binary space with two input variables. Here - 1 represent one class of the training sample and + 1 the other class. The threshold separating the classes is the line in the middle between the two margin boundaries, which are canonically represented as $w^T.x + b = 1$ and $w^T.x + b = -1$. Then, the margin is $2 / ||w||$, where $||w||$ is the norm of the vector *w* (Laura, 2008; Nationalbank, 2007).

In order to avoid data points falling into the margin, except for some classification error ξ, the following constraints are added:

$$w^T.x_i + b \geq +1 - \xi_i \quad (2)$$
for x_i in class $y_i = +1$

$$w^T.x_i + b \leq -1 + \xi_i \quad (3)$$
for x_i in class $y_i = -1$

This can be summarized in the form of:

$$y_i(w^T.x_i + b) \geq 1 - \xi_i \quad (4)$$
for all i = 1, …, n and $\xi_i \geq 0$

It is expressed as an optimization problem for w and b:

$$min_w \ \frac{1}{2}||w||^2 + C\sum_{i=1}^{n}\xi_i \quad (5)$$

subject to:

$$y_i(w^T.x_i + b) \geq 1 - \xi_i$$
for all i = 1, …, n
$$\xi_i \geq 0$$

where *C* is capacity, or a tuning parameter, it weighs in-sample classification errors and therefore controls the generalization ability of an SVM. If *C* increases, the weight given to in-sample misclassifications increases too, this leads to a lower generalization of the machine. On the other hand low generalization means that the machine may work well on the training set, but would perform miserably on a new sample. As a result of overfitting on the training sample, for example, we may get bad generalization, in the case that this sample shows some untypical and non-repeating data structure.

By choosing a low *C,* the risk of over fitting an SVM on the training sample is reduced. The smaller the *C*, the wider is the margin, and the more and larger in-sample classification errors are permitted (Laura, 2008; Nationalbank, 2007). There could be a temptation to express the previous optimization problem by means of non-negative Lagrange multipliers $\propto_i$ as:

$$L(w,b,\propto) = min_{w,b,\propto}\{\frac{1}{2}\left||w|\right|^2 - \sum_{i=1}^{n} \propto_i [y_i(w.x_i + b) - 1]\} \quad (6)$$

However, this would be wrong, because by choosing $\propto_i = +\infty$ we could reach minimum for all members, and not only for the best one which is chosen solving the original problem. This problem is written as follows:

$$L(w,b,\propto) = min_{w,b} max_{\propto}\{\frac{1}{2}\left||w|\right|^2 - \sum_{i=1}^{n} \propto_i [y_i(w.x_i + b) - 1]\} \quad (7)$$

The solution of this optimization problem is given by the saddle-point of the Lagrangian, minimized with respect to *w*, *b* and ξ, and maximized with respect to ∝. This will lead to the fact that all the points which is separated as $y_i(w.x_i + b) - 1 > 0$, do not matter since the corresponding $\propto_i$ must be set to zero. This problem is solved by standard quadratic programming (QP) techniques and programs. The solution can be expressed by terms of linear combination of the training vectors as:

$$w = \sum_{i=1}^{n} \propto_i y_i x_i \quad (8)$$

$$b = \frac{1}{2}(x_{+1}^T + x_{-1}^T).w \quad (9)$$

Only a few $\propto_i$ will be greater than zero. The corresponding x_i are exactly the *support vectors*, which lie on the margin and satisfy:

$$y_i(w.x_i + b) = 1 \quad (10)$$

(Wikipedia, 2010).

The points, whose $\propto_i \neq 0$, are called support vectors and are the relevant ones for the calculation of *w*. Support vectors lie on the margin boundaries or, for non-perfectly separable data, within the margin. This way, the complexity of calculations does not depend on the dimension of the input space, but on the number of support vectors. Here x_{+1} and x_{-1} are any two support vectors belonging to different classes, which lie on the margin boundaries.

The score function z_j can be written as follows:

$$z_j = \sum_{i=1}^{n} \propto_i y_i \langle x_i, x_j \rangle + b \quad (11)$$

The scalar product is the product of the class x_j which is to be classified and x_i the support vectors in the training sample, of $\propto_i$, and of y_i. By comparing z_j with a benchmark value, we are able to estimate if a company has to be classified as x_i or x_j (Laura, 2008; Nationalbank, 2007).

In the case of a non-linear SVM, the score function is computed by substituting the scalar product with a kernel function.

$$z_j = \sum_{i=1}^{n} \propto_i y_i K(x_i, x_j) + b \quad (12)$$

"Kernels are symmetric, semi-positive definite functions satisfying the Mercer theorem. If this theorem is satisfied, it is ensured that there exists a (possibly) non-linear map φ from the input space into some feature space, such that its inner product equals the kernel. The non-linear transformation φ is only implicitly defined through the use of a kernel, since it only appears as an inner product" (Laura, 2008; Nationalbank, 2007).

$$K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle \quad (13)$$

SVM models can be classified into four distinct groups, depending on the form of error function (Statsoft, 2010):

Classification SVM type I

Classification SVM type I also known as C-SVM classification. For this type of SVM, training involves the minimization of the error function:

$$\frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i \quad (14)$$

such that:

$y_i(w^T \varphi(x_i) + b) \geq 1 - \xi_i$ for i= 1,...,n

$\xi_i \geq 0$

where C is the capacity constant, w is the vector of coefficients, b is a constant and ξ_i are parameters for handling nonseparable data (inputs). The index i labels the n training cases. Note that $y_i \in$ {-1, +1} is the class labels and x_i is the independent variables. The kernel φ is used to transform data from the input (independent) to the feature space. It should be noted that the larger the C, the more the error is penalized. Thus, C should be chosen with care to avoid over fitting (Statsoft, 2010).

Classification SVM type II

Classification SVM type II, also known as nu-SVM Classification. In contrast to classification SVM type I, the classification SVM type II model minimizes the error function (Statsoft, 2010):

$\frac{1}{2} w^T w - \nu \rho + \frac{1}{n} \sum_{i=1}^{n} \xi_i$ (15)

such that:

$y_i (w^T \varphi(x_i) + b) \geq \rho - \xi_i$ for i= 1,...,n

$\xi_i \geq 0, \rho \geq 0$

In a regression SVM, you have to estimate the functional dependence of the dependent variable y on a set of independent variables x. Like other regression problems, the assumption is that the relationship between the independent and dependent variables is given by a deterministic function f plus the addition of some additive noise (Statsoft, 2010).

Regression SVM type I

Regression SVM type I also known as epsilon-SVM regression.

y = f(x) + noise (16)

The task is then to find a functional form for f that is able to correctly predict new cases that the SVM has not been presented with before. This can be achieved by training the SVM model on a sample set that is, training set, a process that involves, like classification case, the sequential optimization of an error function. Depending on the definition of this error function, two types of SVM models are recognized (Statsoft, 2010):

Regression SVM type 1

$\frac{1}{2} \mathrm{w^T w} + \mathrm{C} \sum_{\mathrm{i=1}}^{\mathrm{n}} \xi_{\mathrm{i}} + \mathrm{C} \sum_{\mathrm{i=1}}^{\mathrm{n}} \xi_{\mathrm{i}}^{*}$ (17)

such that:

$(w^T \varphi(x_i) + b - y_i) \leq \varepsilon + \xi_i^*$
$\boldsymbol{(y_i - w^T \varphi(x_i) - b_i) \leq \varepsilon + \xi_i}$
$\boldsymbol{\xi_i \geq 0; \xi_i^* \geq 0; i = 1, \ldots, n}$

Regression SVM type 2

Regression SVM type 2 also known as nu-SVM regression. For this SVM model, the error function is given (Statsoft, 2010):

$$\frac{1}{2} w^T w - C \left(\nu \varepsilon + \frac{1}{n} \sum_{i=1}^{n} (\xi_i + \xi_i^*) \right)$$

such that:

$\boldsymbol{(w^T \varphi(x_i) + b) - y_i) \leq \varepsilon + \xi_i}$
$\boldsymbol{(y_i - (w^T \varphi(x_i) + b_i)) \leq \varepsilon + \xi_i^*}$
$\boldsymbol{\xi_i \geq 0; \xi_i^* \geq 0; \varepsilon \geq 0; i = 1, \ldots, n}$

There are number of kernels that can be used in support vector machines models. These include linear, polynomial, radial basis function (RBF) and sigmoid (Statsoft, 2010):

Kernel functions:

$$= \begin{cases} x_i x_i \\ (\gamma x_i x_j + coefficient)^d \\ \exp(-\gamma |x_i - x_j|^2) \\ \tanh(\gamma x_i x_j + coefficient) \end{cases}$$

in the order of appearing:

$$= \begin{cases} Linear \\ Polynomial \\ RBF \\ Sigmoid \end{cases}$$

SVM enjoys a relative relevance, because like all other classification techniques, SVM has advantages and disadvantages, and its importance depends on the kind of data being analyzed (Laura, 2008; Nationalbank, 2007). Authors mention that SVMs are useful tools for analyzing of non-regularity in the data, for instance when the data are not regularly distributed or have an unknown distribution.

The flexibility of choosing threshold separating one class from the other one increase through kernel function, which does not have to be linear and even does not have to have the same functional form for all data, because it is a non-parametric function and operates locally (Nello, 2004). As a consequence these functions work with non-linearly dependent data coming from DCS sensors, which are used in the refineries. The use of kernels, the absence of local minima, the sparseness of the solution and the capacity control obtained by optimizing the margin or other ‘dimension independent’ quantities such as the number of support vectors are the key features of SVM (Nello, 2004).

There is no need to make assumptions about the functional form of the transformation, which makes data linearly separable, because kernel implicitly contains a non-linear transformation. In other words, the human expertise judgment in advance is not needed, since the transformation occurs implicitly on a robust theoretical basis (Laura, 2008; Nationalbank, 2007).

Another advantage of SVM is that training sample data is allowed to have some bias, because by choosing an appropriate generalization grade C and r (in the case of a Gaussian kernel), SVMs are robust (Laura, 2008; Nationalbank, 2007).In comparison with neural networks, which have multiple solutions associated with local minima and therefore, may not be robust over different samples, SVMs deliver a unique solution, because the optimality problem is convex (Laura, 2008; Nationalbank, 2007).

Burges (1998) sees the biggest limitation of the support vector approach in choice of the kernel, the second limitation he mentions is speed and size, both in training and testing. Discrete data presents another problem, and the optimal design for multiclass SVM classifiers is a further area for research.

The lack of transparency of results is seen as a common disadvantage of non-parametric techniques, such as SVMs. High dimensionality of data can represent another limitation for SVMs (Laura, 2008; Nationalbank, 2007).

LITERATURE REVIEW

There are good number of articles and surveys that show the usage of support vector machines, examples are (Byun and Lee, 2002): which looks at the applications of SVM for pattern recognition, or (Campbell, 2000) which covers the algorithmic approaches to training SVM, another survey (Guosheng, 2008) handles training algorithms for SVM classifiers (Samanta et al., 2003) concentrate on artificial neural networks and SVMs with genetic algorithm for bearing fault detection, and (Sapankevych and Sankar, 2009) deals with time series prediction using SVM. These surveys demonstrate the diversity of usage of SVM in many branches of the industry. Recent survey shows, that SVM has found most applications in the financial market prediction, followed by electrical utility forecasting, control process and signal processing, miscellaneous applications, business applications, environmental parameter estimation and machine reliability forecasting (Sapankevych and Sankar, 2009). In his survey, Campbell (2000) indicates the popularity of SVM for machine learning tasks, such as classification, regression or novelty detection, he also mentions that SVMs were successfully applied in particle identification, face recognition, text categorization, engine knock detection and bioinformatics. Guosheng (2008) reviewed more than thirty algorithms that were used in various SVMs applications. Pai (2005) used SVM to forecast the production values of the machinery industry in Taiwan. Yan et al. (2009) propose a forecasting approach on development of manufacturing industry using SVM combined with “Grey system”. Saman (2010) refers to the use of SVM for anomaly detection.

ATMOSPHERIC DISTILLATION AND SUPPORT VECTOR MACHINES APPLICATIONS

Yafen et al. (2006) used least squares support vector machine (LS-SVM), RBF neural network and squares SVM regression methods to control quality of dry point of aviation kerosene in the atmospheric distillation column, which is a very important factor. Authors adopted a method based on LS-SVM regression to implement on-line estimation of aviation kerosene dry point, and compared this method with RBF neural network and SVM regression. Authors claim that their simulation results show that the soft sensing based on LS-SVM regression has better abilities of model generalization and real-time character. Yan et al. (2004) introduced SVM into soft sensor modeling and proposed a soft sensing modeling method based on SVM. A model selection method within the Bayesian evidence framework was proposed to select an optimal model for a soft sensor based on SVM. In their case study, researchers applied soft sensors based on SVM to estimate the freezing point of light diesel oil in distillation column.

Authors showed that the estimated outputs of SVM soft sensors with the optimal model match the real values of the freezing point of light diesel oil which follows the varying trend of the freezing point of light diesel oil very well. Experimental results demonstrated that SVM provides a new and effective method for soft sensing modeling and has promising application in industrial process applications. (Petkovi et al. (2009) used support vector machines for predicting electrical energy consumption in the atmospheric distillation of oil refining at a particular oil refinery. During cross-validation process of the SVM training, they used particle swarm optimization (PSO) algorithm for selecting free SVM kernel parameters. Researchers showed that incorporation of PSO into SVM training process greatly enhanced the quality of prediction.

ENTERPRISE PERFORMANCE AND PREDICTION USING SUPPORT VECTOR MACHINES

Xie et al. (2006) proposed a method for predicting crude oil price; they used SVM and compared its performance with auto regressive integrated moving average (ARIMA) and back-propagation neural network (BPNN). They concluded that SVM outperformed the other two methods.

In order to evaluate and predict the performance of an enterprise such as an oil refinery, Jiekun and Zaixu (2009) proposed a model which uses data envelopment analysis (DEA) and SVM. He used DEA method to first, evaluate DEA efficiency of all the oil refining enterprises performance. Then, he took the input/output data and results of some decision making units (DMUs) as the learning examples to train the SVM network and test the network. He showed that his real example testifies the efficiency, practicability and intellectual ability of this method. Petkovi et al. (2009) used SVM to predict the level of power consumption.

FLARE ANALYSIS USING SUPPORT VECTOR MACHINES

Venkoparao et al. (2009) proposed an algorithm using

Table 1. Applications of SVM in oil refineries.

Number	Algorithm used	Type of SVM	Kernel	Usage of SVM within oil refining process	Purpose	Source
1	Not specified by the authors	LS-SVM regression, SVM regression method	RBF neural network	Atmospheric distillation column	To control quality of the dry point of the aviation kerosene	(Yafen et al., 2006)
2	SVM soft sensors within the Bayesian evidence framework	LS SVM	RBF	Distillation column, fluid catalytic cracking unit	To estimate freezing point of light diesel oil	(Yan et al., 2004, 2009)
3	Particle swarm optimization	Regression SVM	Linear, polynomial, RBF, Sigmoid	Atmospheric distillation	To predict electrical energy consumption	(Petkoviҫ et al., 2009)
4	Not specified by the authors	Not specified by the authors	RBF	Forecasting crude oil price	To predict crude oil price	(Xie et al., 2006)
5	Not specified by the authors	Not specified by the authors	Not specified by the authors	Enterprise performance evaluation and prediction	To evaluate and predict the performance of an oil refinery	(Jiekun and Zaixu, 2009)
6	Own algorithm	Not specified by the authors	Not specified by the authors	Flare analysis	To measure quantity of gases and identify gas components	(Venkoparao et al., 2009)
7	MATLAB SVM Tool Box. Standard programs of this tool box was modified by Balabin	MATLAB SVM Tool Box. Standard programs of this tool box was modified by Balabin	Linear, polynomial, Gaussian (RBF)	Comparison with other classifiers	Classification of gasoline	(Balabin et al., 2010)

SVM to analyze videos captured from the flares. Refineries flare up the exhaust gases prior to releasing the gases in to the atmosphere. The model proposed, was an attempt to reduce environment pollution. They argue that the area or volume of the flare and its color can be interpreted as the quantity of released gas during refining process, and these parameters of flare indirectly indicate the performance of refining process.

CLASSIFICATION OF GASOLINE USING SUPPORT VECTOR MACHINES

Balabin et al. (2010) did near infrared (NIR) spectroscopy for gasoline classification and compared the abilities of nine different multivariate classification methods: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), regularized discriminant analysis (RDA), soft independent modeling of class analogy (SIMCA), partial least squares (PLS) classification, K-nearest neighbour (KNN), support vector machines (SVM), probabilistic neural network (PNN) and multilayer perceptron (ANN-MLP). They found that KNN, SVM and PNN techniques for classification were the most effective ones.

Table 1 summarizes the articles, and gives an overview about the areas within the oil refineries, where SVM has found applications. The second column shows whether or not the authors used a particular known algorithm or developed their own algorithm, the third column shows the type of SVMs used by the authors. The fourth column shows what Kernel function were used. The fifth column shows the area within the oil refineries, where SVM were applied, the sixth column shows the purpose of using SVM and the last column gives the reference to the articles.

RESULTS AND DISCUSSION

SVMs, produce accurate and robust classification results based on a sound mathematical basis without the need for human interference to choose an optimal linearization function. Kernel functions are used to linearize the data. Many support vector machine techniques have been specifically developed for certain application domains. Support vector machines have gained in popularity, because of its powerful learning mechanisms. These mechanisms provide a method of classifications for numerous applications detailed in this survey paper. This survey tries to provide a comprehensive and structured overview of the research on application of support vector machines within oil refineries, since oil refineries are crucial for the civilization.

This paper has grouped the applications of support vector machines within various units inside refineries. For each unit this survey has identified the purpose for using support vector machines. Application of support vector machines to a particular domain within refineries can be used as guidelines to assess the effectiveness of the support vector machines in that domain. This survey also delivers an easier and more succinct understanding of the support vector machines, and identifies the advent-ages and disadvantages of support vector machines. This survey provides a better understanding of the different applications that have been developed for one area which may find applications in other domains at refineries.

This paper is an attempt to provide a broad overview on applications of support vector machines in the oil refineries to the professionals inside oil refineries, researchers and academicians. Authors also summarized the result in a table, and the information were classified as follows: kind of algorithm used, type of support vector machines used, kernel function used, usage of support vector machines within oil refining process, purpose of using support vector machines and source of information. Researchers found that support vector machines were used to analyze flare, classify gasoline, predict crude oil price, evaluate and predict the performance of an oil refinery, control quality of the dry point of the aviation kerosene, and to estimate freezing point of light diesel oil. Researchers also suggested other areas within oil refineries, where support vector machines could be used. Several refineries have their own weather stations which measure different parameters; these data could be used in connection with SVM to help reduce the pollution. SVM can also be used in the laboratory to predict the quality of various products of a refinery. It can also be used to predict viscosity of various fuels. These are just a few research areas and future work that the authors intend to carry out. Authors believe that SVMs can be applied in many other areas within the refineries.

CONCLUSION

This paper has grouped and summarized applications of support vector machines within various units inside refineries. Application of support vector machines to a particular domain within refineries can be used as guidelines to assess the effectiveness of the support vector machines in that domain. In conclusion, SVM research continues to be an important approach for classifications of data within oil refineries; however, researchers who applied SVM for condition monitoring, diagnosis and prognosis within refineries are relatively rare.

REFERENCES

Balabin RM, Safieva RZ, Lomakina EI (2010). Gasoline classification using near infrared (NIR) spectroscopy data: Comparison of multivariate techniques. Analytica Chimica Acta., 671: 27-35.

BurgeS CJC (1998). A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining Knowledge Discovery, 2: 121-167.

Byun H, Lee SW (2002). Applications of Support Vector Machines for Pattern Recognition: A Survey. Proceedings of the First International Workshop on Pattern Recognition with Support Vector Machines. Springer-Verlag, pp. 213-236.

Campbell C (2000). Algorithmic Approaches to Training Support Vector Machines: A Survey. ESANN2000, D-Facto Publications, pp. 27-36.

Chunming WU, Xiaofeng XL, Yalong CAO, Chao Chen MO (2010). Application of support vector regression to predict metallogenic favourability degree. Int. J. Phy. Sci., 5: 2523-2527.

Guosheng WA (2008). Survey on Training Algorithms for Support Vector Machine Classifiers. Networked Computing and Advanced Information Management, NCM '08. Fourth International Conference on, 2-4: 123-128.

Jiekun S, Zaixu Z (2009). Oil Refining Enterprise Performance Evaluation Based on DEA and SVM. Knowledge Discovery and Data Mining, 2009. WKDD 2009. Second International Workshop on, 23-25: 401-404.

Laura RA, Auria M (2008). Support Vector Machines (SVM) as a Technique for Solvency Analysis. Berlin: Deutsches Institut für Wirtschaftsforschung, pp. 3-8.

Nationalbank O (2007). Credit Risk Assessment Revisited Methodological Issues and Practical Implications 2007. Vienna: Oestrreichische Nationalbank, pp. 50-56.

Nello J, Cristianini ST (2004). An introduction to support vector machines: and other kernel-based learning methods, Cambridge University Press, p. 189.

Pai PF, Chih-Sheng LIN (2005). Using support vector machines to forecast the production values of the machinery industry in Taiwan. Int. J. Advan. Manufact. Technol., 27: 205-210.

PetkoviĊ MR, RapaiĊ MR, JakovljeviĊ BB (2009). Electrical Energy Consumption Forecasting in Oil Refining Industry Using Support Vector Machines and Particle Swarm Optimization. WSEAS Transactions on Inf. Sci. Appl., 6: 1761-1770.

Saman-MIRZA LM Abdulalla K, Omar Z (2010). A biological model to improve PE malware detection: Review Int. J. Phys. Sci., 5: 2236-2247.

Samanta B, Al-Balushi KR, Al-Araimi SA (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. Eng. Appl. Artificial. Intell., 16: 657-665.

Sapankevych N, Sankar R (2009). Time Series Prediction Using Support Vector Machines: A Survey. Computational Intelligence Magazine, IEEE, 4: 24-38.

STATSOFT (2010). Support Vector Machines, http://www.statsoft.com/textbook/support-vector-machines/[Online]. StatSoft.com. Available: http://www.statsoft.com/textbook/support-vector-machines/, p. 1.

SVMS.ORG. (2010). Introduction to Support Vector Machines, http://www.svms.org/introduction.html[Online]. SVMS.org. Available: http://www.svms.org/introduction.html [Accessed January 1st 2011], p. 1.

Venkoparao VG, Hota RN, Rao VS, Gellaboina MK (2009). Flare monitoring for petroleum refineries. Industrial Electronics and

Applications, ICIEA 2009. 4th IEEE Conference on, 25-27: 3022-3027.
Wikipedia (2010). Support Vector Machine, http://en.wikipedia.org/wiki/Support_vector_machine[Online]. Wikipedia. Available: http://en.wikipedia.org/wiki/Support_vector_machine [Accessed January 01 2011].
Xie W, Yu L, Xu S, Wang S (2006). A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. Computational Science–ICCS, 3994: 444-451.
Alexandrov V, Van Albada G, Sloot P, Dongarra J (2006). (eds.) Computational Science – ICCS Springer Berlin / Heidelberg.
Yafen L, Qi L, Huijuan W, Ningsheng M (2006). Soft Sensing Based on LS-SVM and Its Application to a Distillation Column. Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on, 16-18: 177-182.
Yan L, Xue H, Guo XA (2009). Forecasting Approach on Development of Manufacturing Industry Using Support Vector Machine Combined with Grey System.. IEEE, pp. 554-558.
Yan W, Shao H, Wang X (2004). Soft sensing modeling based on support vector machine and Bayesian model selection. Comput. Chem. Eng., 28: 1489-1498.