Research Journal of Applied Sciences 9 (4): 208-213, 2014
ISSN: 1815-932X

Double Linear Support Vector Machine for Dimensionality Reduction

Wipawan Buathong and Phayung Meesad
Faculty of Information Technology,
King Mongkut's University of Technology North Bangkok,
10800 Bangkok, Thailand

Abstract: This study proposes an alternative feature selection technique for dimensionality reduction namely Double Linear Support Vector Machine or "DLSVM" Weight. The efficiency of DLSVM was measured based on four performance evaluation criteria (i.e., accuracy, F-measure, precision and recall). The efficiency of well recognised feature selection techniques was also measured for comparative purposes. The Support Vector Machine (SVM), a prominent classifier was also used with DLSVM and these feature selection techniques. The Leukemia dataset from the University of California Irvine (UCI) machine learning repository was used for the experiments. Downsized data dimensions were classified into 60, 50, 40, 30, 20 and 10, respectively. The experimental results showed that the DLSVM was much more efficient than other feature selection techniques at almost all of the data dimensions. Particularly, all performance evaluation criteria of DLSVM could reach 100% when original data dimensions were downsized from 5,147-60, 50 and 40.

Key words: Data dimensions, dimensionality reduction, DLSVM, feature selection, Linear SVM Weight

INTRODUCTION

Data mining is simply defined as "the exploration of databases to find patterns in the data" (Downing *et al*., 2009). It is also defined as the analysis of large quantities of data to discover previously unknown patterns (Daintith and Wright, 2006; Tan *et al*., 2006). In order to find patterns in large quantities of data, a number of procedures, techniques and tools are required. Dimensionality reduction, for example is a widely known technique of data mining which is about converting data of very high dimensionality into data of much lower dimensionality. It is the process of reducing the number of random variables or 'data dimensions' under specified criteria which is considered as an effective downsizing data approach. The least possibility of important data loss and the accuracy of data classification are key objectives of dimensionality reduction (Tan *et al*., 2006).

Two classification approaches have been widely used for dimensionality reduction: feature selection (Sarrafzadeh *et al*., 2012; Sriurai *et al*., 2009; Sun and Wu, 2008; Chang and Lin, 2008) (a process for selecting an optimal subset of features or attributes in accordance with an objective function) and feature extraction (a process of creating new features from functions of the original features and use them as representatives for the original features, e.g., sampling and clustering) (Sriurai *et al*., 2009; Sun and Wu, 2008). Data dimensions downsized using the feature selection technique can be generalized for the majority of data. The data is separated into training set or the portion of data which is used to map a Data Mining Model. This process is an important part of evaluating Data Mining Models. The training set is then tested with different algorithms in order to find the most reliable model.

Since, different data dimensions are not equally important, it is necessary to select an efficient feature selection technique to ensure that the key objectives of dimensionality reduction will be minimally affected. In this case, a number of feature selection techniques have been extensively studied in the literature (Chang and Lin, 2008; Sun and Wu, 2008; Guyon and Elisseeff, 2003; Weston *et al*., 2001; Dash and Liu, 1997). Advantages and disadvantages of these feature selection techniques have been pointed out. Among these techniques, ReliefF and Linear SVM Weight are prominent feature selection algorithms that have been widely discussed in the literature (Buathong and Meesad, 2013). The ReliefF technique is well recognised in its noise resistance and capability in dealing with different data classes, whereas the Linear SVM Weight technique is acknowledged among research scholars in its flexibility which is applicable to all classifiers. In addition to feature selection, a classifier also plays a crucial role in

Corresponding Author: Wipawan Buathong, Faculty of Information Technology,
King Mongkut's University of Technology North Bangkok, 10800 Bangkok, Thailand

dimensionality reduction. It is a model that labels or classifies a test case into one of a finite number of output classes. The classifier's performance depends on its algorithm that performs classification (classification algorithm) which is generally measured by its ability to correctly classify test cases (accuracy) or the ratio of the number of incorrect predictions and the number of all predictions which can be both correct and incorrect (Error Rate or "ER") (Jin *et al*., 2007). While some classification algorithms may first construct a model that then can be used to classify test cases (e.g., decision tree, logistic regression), others may perform the classification directly (K-nearest-neighbor). This research proposes Double Linear SVM or "DLSVM" as an alternative feature selection technique for dimensionality reduction. The efficiency of DLSVM for dimensionality reduction will be measured and compared with other feature selection techniques that have been widely discussed in the literature, i.e., Linear SVM Weight, ReliefF, Gini Gain, Information Gain and Gain Ratio. The SVM classifier will be used together with these feature selection techniques for dimensionality reduction.

Literature review: According to Sriurai *et al*. (2009), the efficiency of the Information Gain and Chi-squared feature selection techniques was measured. These two feature selection techniques, together with naive bayesian, Support vector machine and decision tree classifiers were used for document classification. The results showed that the Information Gain was more effective than the Chi-squared. In addition, SVM was pointed out as the most efficient document classifier (Sriurai *et al*., 2009). In the ReliefF algorithm was used as a feature selection for content-based image retrieval (Sarrafzadeh *et al*., 2012). It was discovered that the data retrieval performance using the ReliefF algorithm was faster and more accurate than other algorithms. Jin *et al*. (2007) conducted the experiment on the hidden naive bayes classifier, originally proposed by Zhang *et al*. (2005) for web pages categorisation. In order to compare the efficiency of the hidden naive bayes, four feature selection techniques, i.e., ReliefF, Information Gain, Gain Ratio and Chi-squared were used with different classifiers such as naive bayes, SVM and decision tree, in addition to the hidden naive bayes. Error Rate (ER) and F1 (a criterion that combines recall and precision) were used as performance evaluation criteria for the efficiency measurement. The measurement results showed that using ReliefF with the hidden naive bayes classifier was the most efficient method for web pages categorisation: the lowest ER (5.7% with feature selection and 8.8% without feature selection) and the highest F1 (0.96 with feature selection and 0.94 without feature selection). The efficiency of feature selection techniques for ranking was measured by Buathong (2012). The Information Gain, Gain Ratio and Linear SVM Weight ranking techniques were measured using four classifiers including k-NN, naive bayes, SVM and classification tree. The Linear SVM Weight feature selection technique with the SVM classifier was discovered by the another as the most efficient method providing the best result for dimensionality reduction.

While a single feature selection is widely used for dimensionality reduction, it may not ensure the least possibility of important data loss from dimensionality reduction. A single feature selection may not be efficient enough in excluding redundant and irrelevant features. Some important dimensionality may not be selected when a subset is chosen. In this case, several combined methods of feature selection have been proposed as promising techniques that can be more efficient than a single feature selection technique for dimensionality reduction. According to Zhang *et al*. (2007), two feature selection techniques, including the reliefF and mRMR were combined for gene selection. The results showed that the classification accuracy rate based on this combined method was much higher than a single feature selection when original data dimensions were downsized to 30. According to Buathong and Meesad (2013), a combination of Linear SVM Weight and ReliefF algorithms was proposed as an alternative feature selection technique for improving the efficiency of dimensionality reduction. Two datasets, including leukemia and DLBCL from UCI machine learning repository were used in the experiments. When original data dimensions of the Leukemia dataset were downsized from 5,147-20, the combined Linear SVM Weight and ReliefF feature selection technique was much more efficient than using either Linear SVM Weight or ReliefF alone. All performance evaluation criteria of the combined method could reach 100% for the leukemia dataset. At the same data dimensions for the DLBCL dataset, the combined method also showed satisfactory results for all performance evaluation criteria which were higher than those in other feature selection techniques. However, Linear SVM was still summarised by Buathong and Meesad (2013) as an efficient feature selection technique for dimensionality reduction since, its accuracy performance was higher than that of the combined method when data dimensions were downsized to 10.

MATERIALS AND METHODS

The proposed techniques: In this study, the proposed technique is described. The flowchart of proposed technique is shown in Fig. 1.

Select benchmarking dataset: Selecting the dataset with high dimensional data for benchmarking is an initial stage for the efficiency measurement of feature selection

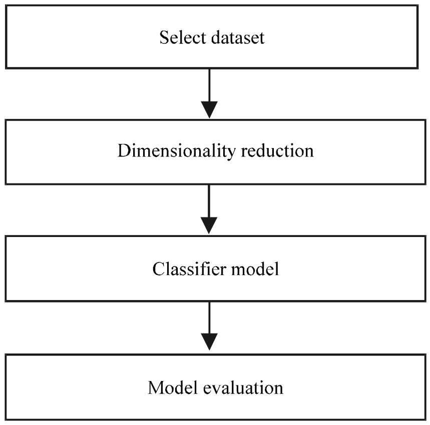

Fig. 1: Flowchart of proposed technique

techniques. The efficiency of data dimensions downsized from high dimensionality data is more credible than that of data dimensions reduced from lower dimensionality data. In this research, the Leukemia dataset from UCI database was selected for benchmarking because of its high dimensional data. There were 5,147 attributes in the selected dataset and there were no missing values. The Leukemia dataset was based on two data classes including Acute Lymphoblastic Leukemia (ALL) and Acute Myelogenous Leukemia (AML).

Perform dimensionality reduction using selected feature selection techniques: Linear SVM, ReliefF, Gini Gain, Information Gain, Gain Ratio and Double Linear SVM were feature selection techniques chosen for dimensionality reduction. Advantages and disadvantages of these feature selection techniques have long been discussed in the literature. For example, Linear SVM Weight is acknowledged among research scholars as a very efficient that can be applied to all types of Classifier (Chang and Lin, 2008). The algorithm of Linear SVM Weight is given:

Algorithm 1: Feature ranking based on Linear SVM Weights
Input: Training sets (X_i, Y_i), i = 1,...,m.
Output: Sorted feature ranking list.

1. Use grid search to find the best parameter C.
2. Train aL2-loss Linear SVM Model using the best C.
3. Sort the features according to the absolute values of weights in the model.

ReliefF, invented by Kira and Rendell is one of the most successful algorithms for evaluating the quality of features due to its simplicity and effectiveness (Sarrafzadeh *et al*., 2012; Sun and Wu, 2008; Zhang *et al*., 2007; Jin *et al*., 2007; Harris, 2002). ReliefF is well recognized among researchers in its robustness against noise. This algorithm is mainly used for calculating Weight (W) of data feature in reference to random data (R). Given that A, B and C are three data classes and R is allocated to A, a variable H will be assigned to for the nearest data to A. Mb and Mc will be assigned for the nearest data to B and C, respectively. The ReliefF feature selection technique can be applied to the data that contains more than two classes (Sun and Wu, 2008; Zhang *et al*., 2007).

Other widely studied feature selection techniques based on series of gain algorithms, i.e., Gain Gini, Information Gain and Gain Ratio were also selected for efficiency measurement. Although, feature selection techniques based on Gain algorithms are well known for their bias when applied to attributes with a considerable number of distinct values (Deng *et al*., 2011; Harris, 2002), they are worth considering because of their classical approaches for deciding the relevance of an attribute.

Apply the SVM classifier: Since, the emphasis of this study is to measure the efficiency of DLSVM in comparison with other feature selection techniques, an appropriate classifier is also required. Although, a number of classifiers have been proposed from time to time, the SVM classifier was chosen because of its simplicity widely recognised among research scholars. Support Vector Machine or "SVM" (Buathong *et al*., 2012; Campbell *et al*., 2003; Hsu and Lin, 2002) is a data classifier algorithm that has been applied in various disciplines. SVM has been considered by a number of research scholars as the most efficient classifier (Buathong, 2012; Zakaria *et al*., 2011; Pongpatharakan, 2009; Chang and Lin, 2008) which is considered as an open research area for further applications and improvements. Generally, SVM takes a set of input data and applies a simple linear method to the data but in a high dimensional feature space which is non-linearly related to the input space. The SVM algorithm is consisted of support vectors (training samples) which are the data points that are nearest to the decision surface or 'hyperplane' specified by a subset of support vectors. The SVM uses a kernel function that corresponds to a dot product of two feature vectors in some expanded feature space that aims to minimize errors and maximize the margin around the separating hyperplane (Vapnik, 1995). The original input space can be mapped to some higher dimensional feature space where the training set is separable. There are some occasions that two groups of data cannot be divided using general methods of the SVM classifier because data are clustered in different positions.

In this case, a multidimensional linear classifier considered as more efficient than general methods in data classification was also used. The Radial Basis Function Kernel Classifier (RBFKC) of SVM uses C as a variable for balancing point for measuring the best classification range and the least potential error rate. The SVM classifier was used together with each feature selection technique for dimensionality reduction.

Evaluate the efficiency of feature selection techniques: The dataset is classified into two parts: training set (60%) and test set (40%). These two parts of the dataset will be used as a model for measuring the accuracy of data classification. The ten folds cross validation will be applied to each set of data. The measurement metric is consisted of data dimensions and performance evaluation criteria. The dataset was classified into six dimensions based on the number of attributes (10, 20, 30, 40, 50 and 60). In the meantime, accuracy, F-measure, precision and recall are the performance evaluation criteria. The feature selections techniques are measured based on precision, recall, F-measure and accuracy of the classification. The Accuracy calculated after precision, recall and F-measure have already been calculated (Eq. 1-8):

$$\text{Precision (TP)} = \frac{\text{TP}}{\text{TP+FP}} \tag{1}$$

$$\text{Precision (TN)} = \frac{\text{TN}}{\text{TP+FP}} \tag{2}$$

$$\text{Precision} = \frac{\text{Precision (TP)+Precision (TN)}}{2} \tag{3}$$

$$\text{Recall (TP)} = \frac{\text{TP}}{\text{TP+FN}} \tag{4}$$

$$\text{Recall (TN)} = \frac{\text{TN}}{\text{TP+FN}} \tag{5}$$

$$\text{Recall} = \frac{\text{Recall (TP)+Recall (TN)}}{2} \tag{6}$$

$$\text{F-measure} = \frac{2\times(\text{Recall}\times\text{Precision})}{\text{Recall+Precision}} \tag{7}$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{8}$$

Where:
TP = Number of True Positives
TN = Number of True Negatives
FP = Number of False Positives
FN = Number of False Negatives

RESULTS AND DISCUSSION

Table 1 represents the efficiency of dimensionality reduction using Linear SVM Weights that has been pointed out by a number of research scholars (Buathong *et al*., 2012; Buathong, 2012; Chang and Lin, 2008; Zakaria *et al*., 2011; Pongpatharakan, 2009). The results confirm the strength of Linear SVM Weights that has been pointed out by a number of research scholars. All four key performance evaluation criteria for measuring the efficiency of Linear SVM Weights reached 100% when original data dimensions were downsized to 20. In addition, the 100% rate of performance could be found in the recall criterion when data dimensions were reduced to 40, 50 and 60. Please note that there is no single criterion where the performance is below 97%.

In Table 2, the efficiency of the ReliefF feature selection for dimensionality reduction is represented. Different from Linear SVM Weights, there is no single key evaluation criterion reaching 100% when data dimensions are downsized using ReliefF. All four performance evaluation criteria for measuring the efficiency of ReliefF are below 97% at all data dimensions. Particularly, the accuracy and F-measure of ReliefF are <94% at all data dimensions. Precision is the only one criterion where the performance is the least efficient which is <92% at most data dimensions.

The efficiency of using Information Gain for dimensionality reduction is represented in Table 3. When data dimensions were downsized to 20 using Information

Table 1: The performance of the Linear SVM Weights Method

Dimension	Accuracy	F-measure	Precision	Recall
60	99.31	99.48	98.96	100.00
50	99.31	99.48	98.96	100.00
40	98.97	99.22	98.45	100.00
30	98.62	98.95	98.95	98.95
20	100.00	100.00	100.00	100.00
10	97.24	97.91	97.40	98.42

Table 2: The performance of the ReliefF Ranking Method

Dimension	Accuracy	F-measure	Precision	Recall
60	93.41	93.43	91.00	96.0
50	92.93	92.91	90.91	95.0
40	93.41	93.40	91.39	95.5
30	93.41	93.37	91.79	95.0
20	93.17	93.14	91.35	95.0
10	92.44	92.27	92.04	92.5

Table 3: The performance of the Information Gain Ranking Method

Dimension	Accuracy	F-measure	Precision	Recall
60	96.55	97.33	98.91	95.79
50	94.48	95.81	95.31	96.32
40	94.48	95.81	95.31	96.32
30	94.48	95.81	95.31	96.32
20	95.52	96.69	93.60	100.00
10	92.07	94.15	91.13	97.37

Table 4: The performance of the Gain Ratio Ranking Method

Dimension	Accuracy	F-measure	Precision	Recall
60	94.48	95.81	95.31	96.32
50	96.21	97.08	97.86	96.32
40	94.48	95.81	95.31	96.32
30	94.48	95.81	95.31	96.32
20	95.52	96.69	93.60	100.00
10	92.07	94.15	91.13	97.37

Table 5: The performance of the Gini Gain Ranking Method

Dimension	Accuracy	F-measure	Precision	Recall
60	96.90	97.60	98.92	96.32
50	94.14	95.56	94.82	96.32
40	95.86	96.83	97.34	96.32
30	96.21	97.19	94.53	100.00
20	95.52	96.69	93.60	100.00
10	92.07	94.15	91.13	97.37

Gain, recall is the only criterion that reaches 100%. Although, the results of feature ranking using Information Gain cannot be comparable to those of linear SVM weights, they are slightly better than those of ReliefF. However, the results also show significant differences between the performance of precision when data dimensions were downsized to 10 and that of recall when data dimensions were reduced to 20. Although, the lowest performance could be found in precision criterion (91.13%) at 10 data dimensions, the performance of recall could reach 100% at a dimension of 20.

Table 4 shows the efficiency of Gain Ratio when used to downsize data dimensions. It is notable that the results are comparatively similar to those of Information Gain represented in Table 3. Particularly, the performance of all four criteria between Information Gain and Gain Ratio is exactly the same when the data dimension is downsized to 40 and below. The recall criterion reached 100% when the data is downsized to 20 dimensions. Please also note that the efficiency of Information Gain at 60 data dimension is slightly higher than that of Gain Ratio at the same dimension. Gain Ratio, on the other hand is more efficient than Information Gain at a dimension of 50 attributes.

The efficiency of Gini Gain for dimensionality reduction is represented in Table 5 which is slightly different from that of Information Gain and Gain Ratio. All four performance evaluation criteria are exactly the same as those of Information Gain and Gain Ratio at a dimension of 10 and 20. However, Gini Gain is more efficient than Information Gain and Gain Ratio when data dimensions are downsized to 40 and 30.

Table 6 shows the efficiency of Double Linear SVM. The results indicate that double Linear SVM Weights is the most efficient feature selection technique in comparison with others for reducing data dimensions. Particularly, all four key performance evaluation criteria for measuring the efficiency of this technique reached 100% when original data dimensions are downsized to 60, 50 and 40, respectively. Despite being the most efficient, the efficiency of Double Linear SVM Weights is less than that of Linear SVM at a dimension of 20 attributes where all four performance evaluation criteria of linear SVM reach 100%.

Table 6: The performance of the Double Linear SVM Weight Ranking Method

Dimension	Accuracy	F-measure	Precision	Recall
60	100.00	100.00	100.00	100
50	100.00	100.00	100.00	100
40	100.00	100.00	100.00	100
30	99.31	99.48	98.96	100
20	99.66	99.74	99.48	100
10	97.93	98.45	96.94	100

Table 7: The accuracy comparison of feature selection techniques for dimensionality reduction

Feature selection technique	Data demensions (accuracy)					
	60	50	40	30	20	10
Linear SVM Weight	99.31	99.31	98.97	98.62	100.00	97.24
ReliefF	93.41	92.93	93.41	93.41	93.17	92.44
Information Gain	96.55	94.48	94.48	94.48	95.52	92.07
Gain Ratio	94.48	96.21	94.48	94.48	95.52	92.07
Gini Gain	96.90	94.14	95.86	96.21	95.52	92.07
Double Linear SVM Weight	100.00	100.00	100.00	99.31	99.66	97.93

Table 7 compares the accuracy of feature selection techniques for dimensionality reduction. It is obvious that Double Linear SVM is the most accurate feature selection technique in comparison with others for dimensionality reduction. The result also shows that there is no significant difference among the three gain based feature selection techniques.

CONCLUSION

The efficiency of feature selection techniques for dimensionality reduction with the SVM classifier has been measured. The efficiency measurement was based on four performance evaluation criteria, including accuracy, F-measure, precision and recall. Based on this experimental findings, data dimensions could be downsized to 20 data dimensions with all performance evaluation criteria at 100% for the Leukemia dataset. Based on the measurement results, DLSVM is the most efficient feature selection technique for dimensionality reduction in comparison with others. All performance evaluation criteria of DLSVM could reach 100% which were superior to those of other techniques when data dimensions were downsized to 60, 40 and 30. However, LSVM is still more efficient than DLSVM when data dimensions were downsized to 20.

REFERENCES

Buathong, W. and P. Meesad, 2013. Enhancing the efficiency of dimensionality reduction using a combined linear svm weight with reliefF feature selection method. Proceedings of the 9th International Conference on Computing and Information Technology, May 9-10, 2013, Bangkok, pp: 125-134.

Buathong, W., 2012. A comparison of dimensionality reduction techniques using information gain, gain ratio and linear SVM weights ranking methods. Proceedings of the 5th ACTIS National Conference on Applied Computer Technology and Information Systems, September 2012, Songkhla, pp: 185-189.

Buathong, W., K. Suwannaraj and W. Channasap, 2012. Effect of data dimension reduction analysis and data classification performance of decision tree, support vector machine and naive bayes. Proceedings of the 8th National Conference on Computing and Information Technology, May 3-Jun 1, 2012, Bangkok, pp: 568-573.

Campbell, C., C. Lin, S. Keerthi and A.V. Sanchez, 2003. Special issue on support vector machines. Neurocomputing, 55: 1-3.

Chang, Y.W. and C.J. Lin, 2008. Feature ranking using linear SVM. J. Mach. Learn. Res. Proc., 3: 53-64.

Daintith, J. and E. Wright, 2006. The Facts on File Dictionary of Computer Science. 2nd Edn., Facts On File, New York, ISBN: 9780816059997, Pages: 273.

Dash, M. and H. Liu, 1997. Feature selection for classification. Intell. Data Anal., 1: 131-156.

Deng, H., G. Runger and E. Tuv, 2011. Bias of importance measures for multivalued attributes and solutions. Proceedings of the 21st International Conference on Artificial Neural Networks, June 14-17, 2011, Berlin, Heidelberg, pp: 293-300.

Downing, D.A., M.A. Covington, M.M. Covington and C.A. Covington, 2009. Dictionary of Computer and Internet Terms. 10th Edn., Barron's, New York.

Guyon, I. and A. Elisseeff, 2003. An introduction to variable and feature selection. J. Machine Learn. Res., 3: 1157-1182.

Harris, E., 2002. Information gain versus gain ratio: A study of split method biases. http://rutcor.rutgers.edu/~amai/aimath02/PAPERS/14.pdf.

Hsu, C.W. and C.J. Lin, 2002. A comparison of methods for multiclass support vector machines. IEEE Trans. Neural Networks, 13: 415-425.

Jin, X., R. Li, X. Shen and R. Bie, 2007. Automatic web pages categorization with relieff and hidden naive bayes. Proceedings of the ACM Symposium on Applied Computing, March 11-15, 2007, Korea, pp: 617-621.

Pongpatharakan, P., 2009. A comparative study of classification properties between cart, SVM, c5.0 and hybrid methods. Proceedings of the 5th National Conference on Computing and Information Technology, August 23-28, 2009, Atlanta GA., pp: 1102-1106.

Sarrafzadeh, A., H.A. Atabay, M.M. Pedram and J. Shanbehzadeh, 2012. ReliefF based feature selection in content-based image retrieval. Proceedings of the International Multi Conference on Engineering and Computer Science, Volume 1, March 14-16, 2012, Hong Kong, pp: 19-22.

Sriurai, W., P. Meesad and C. Haruchaiyasak, 2009. A topic-model based feature processing for text categorization. Proceedings of the 5th National Conference on Computing and Information Technology, August 23-28, 2009, Atlanta GA., pp: 146-151.

Sun, Y. and D. Wu, 2008. A relief based feature extraction algorithm. Proceedings of the SIAM International Conference on Data Mining, April 24-26, 2008, Atlanta, Georgia, pp: 188-195.

Tan, P.N., M. Steibach and V. Kumar, 2006. Introduction to Data Mining. Pearson Addison Wesley, Boston, MA., USA., ISBN-13: 9780321420527, Pages: 769.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. 1st Edn., Springer-Verlag, New York, USA.

Weston, J., S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio and V. Vapnik, 2001. Feature Selection for SVMs. In: Advances in Neural Information Processing Systems 13, Leen, T.K., T.G. Dietterich and V. Tresp (Eds.). MIT Press, USA., pp: 668-674.

Zakaria, M.A.M.S.O., N.W. Ahmad and M.A. Zaidi, 2011. A classification method for data mining using Svm-weight and euclidean distance. Australian J. Basic Applied Sci., 5: 2053-2059.

Zhang, H., L. Jiang and J. Su, 2005. Hidden Naive Bayes. Proceedings of the 20th National Conference on Artificial Intelligence, Volume 2, July 9-13, 2005, Pittsburgh, Pennsylvania, pp: 919-924.

Zhang, Y., C. Ding and T. Li, 2007. A two-stage gene selection algorithm by combining relieff and mrmr. Proceedings of the 7th IEEE International Symposium on Bioinformatics and Bioengineering, October 14-17, 2007, Boston, pp: 164-171.