I.J.Modern Education and Computer Science, **2013, 10, 64-71**
Published Online November 2013 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2013.10.08

Empirical Analysis of Bagged SVM Classifier for Data Mining Applications

M.Govindarajan
Assistant Professor, Department of Computer Science and Engineering,
Annamalai University, Annamalai Nagar – 608002, Tamil Nadu
Email: govind_aucse@yahoo.com

***Abstract*—** Data mining is the use of algorithms to extract the information and patterns derived by the knowledge discovery in databases process. Classification maps data into predefined groups or classes. It is often referred to as supervised learning because the classes are determined before examining the data. The feasibility and the benefits of the proposed approaches are demonstrated by the means of data mining applications like intrusion detection, direct marketing, and signature verification. A variety of techniques have been employed for analysis ranging from traditional statistical methods to data mining approaches. Bagging and boosting are two relatively new but popular methods for producing ensembles. In this work, bagging is evaluated on real and benchmark data sets of intrusion detection, direct marketing, and signature verification in conjunction with as the base learner. The proposed is superior to individual approach for data mining applications in terms of classification accuracy.

***Index Terms*—** Data Mining, Support Vector Machine, Intrusion Detection, Direct Marketing, Signature Verification, Classification Accuracy, Ensemble Method

I. INTRODUCTION

1.1 Intrusion Detection

Traditional protection techniques such as user authentication, data encryption, avoiding programming errors and firewalls are used as the first line of defense for computer security. If a password is weak and is compromised, user authentication cannot prevent unauthorized use; firewalls are vulnerable to errors in configuration and suspect to ambiguous or undefined security policies (Summers, 1997). They are generally unable to protect against malicious mobile code, insider attacks and unsecured modems. Programming errors cannot be avoided as the complexity of the system and application software is evolving rapidly leaving behind some exploitable weaknesses. Consequently, computer systems are likely to remain unsecured for the foreseeable future. Therefore, intrusion detection is required as an additional wall for protecting systems despite the prevention techniques. Intrusion detection is useful not only in detecting successful intrusions, but also in monitoring attempts to break security, which provides important information for timely countermeasures (Heady et al., 1990; Sundaram, 1996). Intrusion detection is classified into two types: misuse intrusion detection and anomaly intrusion detection. Several machine-learning paradigms including neural networks (Mukkamala et al.,2003), linear genetic programming (LGP) (Mukkamala et al., 2004a), support vector machines (SVM), Bayesian networks, multivariate adaptive regression splines (MARS) (Mukkamala et al., 2004b) fuzzy inference systems (FISs) (Shah et al.,2004), etc. have been investigated for the design of IDS.

1.2 Direct Marketing

In direct marketing, companies or organizations try to establish and maintain a direct relationship with their customers in order to target them individually for specific product offers or for fund raising. There are two main approaches for companies to promote their products / services: through mass campaigns, which target the general public population, and directed campaign, which targets only a specific group of people. Formal study shows that the efficiency of mass campaign is pretty low. Usually less than 1% of the whole population will have positive response to the mass campaign. In contrast, direct campaign focuses only on a small set of people who are believed to be interested in the product/service being marketed and thus would me much more efficient. This paper focuses only on the direct marketing data. The goal is to predict if a customer will subscribe the service provided by the bank, thereby improving the effect of direct marketing.

1.3 Signature Verification

Optical Character Recognition (OCR) is a branch of pattern recognition, and also a branch of computer vision. OCR has been extensively researched for more than four decades. With the advent of digital computers, many researchers and engineers have been engaged in this interesting topic. It is not only a newly developing topic due to many potential applications, such as bank check processing, postal mail sorting, automatic reading of tax forms and various handwritten and printed materials, but it is also a benchmark for testing and verifying new pattern recognition theories and algorithms. In recent

years, many new classifiers and feature extraction algorithms have been proposed and tested on various OCR databases and these techniques have been used in wide applications. Numerous scientific papers and inventions in OCR have been reported in the literature. It can be said that OCR is one of the most important and active research fields in pattern recognition. Today, OCR research is addressing a diversified number of sophisticated problems. Important research in OCR includes degraded (heavy noise) omni font text recognition, and analysis/recognition of complex documents (including texts, images, charts, tables and video documents). Handwritten numeral recognition, (as there are varieties of handwriting styles depending on an applicant's age, gender, education, ethnic background, etc., as well as the writer's mood while writing), is a relatively difficult research field in OCR.

In the area of character recognition, the concept of combining multiple classifiers is proposed as a new direction for the development of highly reliable character recognition systems (C.Y.Suen et al., 1990) and some preliminary results have indicated that the combination of several complementary classifiers will improve the performance of individual classifiers (C.Y.Suen et al., 1990 and T.K.Ho et al., 1990).

The rest of this paper is organized as follows: Section 2 describes the related work. Section 3 presents classification methods and Section 4 explains the performance evaluation measures. Section 5 focuses on the experimental results and discussion. Finally, results are summarized and concluded in section 6.

II. RELATED WORK

2.1 Intrusion Detection

The Internet and online procedures is an essential tool of our daily life today. They have been used as an important component of business operation (T. Shon and J. Moon, 2007). Therefore, network security needs to be carefully concerned to provide secure information channels. Intrusion detection (ID) is a major research problem in network security, where the concept of ID was proposed by Anderson in 1980 (J.P. Anderson, 1980). ID is based on the assumption that the behavior of intruders is different from a legal user (W. Stallings, 2006). The goal of intrusion detection systems (IDS) is to identify unusual access or attacks to secure internal networks (C. Tsai, et al., 2009) Network-based IDS is a valuable tool for the defense-in-depth of computer networks. It looks for known or potential malicious activities in network traffic and raises an alarm whenever a suspicious activity is detected. In general, IDSs can be divided into two techniques: misuse detection and anomaly detection (E. Biermannet al.2001; T. Verwoerd, et al., 2002) Misuse intrusion detection (signature-based detection) uses well-defined patterns of the malicious activity to identify intrusions (K. Ilgun et al., 1995; D. Marchette, 1999) However, it may not be able to alert the system administrator in case of a new attack. Anomaly detection attempts to model normal behavior profile. It identifies malicious traffic based on the deviations from the normal patterns, where the normal patterns are constructed from the statistical measures of the system features (S. Mukkamala, et al., 2002). The anomaly detection techniques have the advantage of detecting unknown attacks over the misuse detection technique (E. Lundin and E. Jonsson, 2002). Several machine learning techniques including neural networks, fuzzy logic (S. Wu and W. Banzhaf, 2010), support vector machines (SVM) (S. Mukkamala, et al., 2002; S. Wu and W. Banzhaf, 2010) have been studied for the design of IDS. In particular, these techniques are developed as classifiers, which are used to classify whether the incoming network traffics are normal or an attack. This paper focuses on the support vector machine (SVM) among various machine learning algorithms.

2.2 Direct Marketing

Various data mining techniques have been used to model customer response to catalogue advertising. Traditionally statistical methods such as discriminant analysis, least squares and logistic regression have been applied to response modeling.

Given the interest in this domain, there are several works that use DM to improve bank marketing campaigns (Ling and Li, 1998) (Hu, 2005) (Li et al, 2010). In particular, often these works use a classification DM approach, where the goal is to build a predictive model that can label a data item into one of several predefined classes (e.g. "yes", "no"). Several DM algorithms can be used for classifying marketing contacts, each one with its own purposes and capabilities. Examples of popular DM techniques are: Naïve Bayes (NB) (Zhang, 2004), Decision Trees (DT) (Aptéa and Weiss, 1997) and Support Vector Machines (SVM) (Cortes and Vapnik, 1995).

Neural Networks have also been used in response modeling. To overcome the neural networks limitations, Shin and Cho applied Support Vector Machine (SVM) to response modeling. In their study, they introduced practical difficulties such as large training data and class imbalance problem when applying SVM to response modeling. They proposed a neighborhood property based pattern selection algorithm (NPPS) that reduces the training set without accuracy loss. For the other remaining problem they employed different misclassification costs to different class errors in the objective function (Shin 2006).

Although SVM is applied to a wide variety of application domains, there have been only a couple of SVM application reports in response modeling. Cheung, Kwok, Law, and Tsui (2003) used SVM for content-based recommender systems. The system is definitely a form of direct marketing that has emerged by virtue of recent advances in the World Wide Web, e-business, and on-line companies. They compared Naive Bayes, C4.5 and 1-nearest neighbor rule with SVM. The SVM yielded the best results among them. More specific,

SVM application to response modeling was attempted by Viaene et al. (2001b).

2.3 Signature Verification

In the past several decades, a wide variety of approaches have been proposed to attempt to achieve the recognition system of handwritten numerals. These approaches generally fall into two categories: statistical method and syntactic method (C. Y. Suen, et al., 1992). First category includes techniques such as template matching, measurements of density of points, moments, characteristic loci, and mathematical transforms. In the second category, efforts are aimed at capturing the essential shape features of numerals, generally from their skeletons or contours. Such features include loops, endpoints, junctions, arcs, concavities and convexities, and strokes.

Suen et al.,(1992) proposed four experts for the recognition of handwritten digits. In expert one, the skeleton of a character pattern was decomposed into branches. The pattern was then classified according to the features extracted from these branches. In expert two, a fast algorithm based on decision trees was used to process the more easily recognizable samples, and a relaxation process was applied to those samples that could not be uniquely classified in the first phase. In expert three, statistical data on the frequency of occurrence of features during training were stored in a database. This database was used to deduce the identification of an unknown sample. In expert four, structural features were extracted from the contours of the digits. A tree classifier was used for classification. The resulting multiple-expert system proved that the consensus of these methods tended to compensate for individual weakness, while preserving individual strengths. The high recognition rates were reported and compared favorably with the best performance in the field.

The utilization of the Support Vector Machine (SVM) classifier has gained immense popularity in the past years (C. J. C. Burges., et al., 1997 and U. Krebel, 1999). SVM is a discriminative classifier based on Vapnik's structural risk minimization principle. It can be implemented on flexible decision boundaries in high dimensional feature spaces. Generally, SVM solves a binary (two-class) classification problem, and multi-class classification is accomplished by combining multiple binary SVMs. Good results on handwritten numeral recognition by using SVMs can be found in Dong, et al.'s paper.

2.4 Bagging Classifiers

Breiman (1996c) showed that bagging is effective on "unstable" learning algorithms where small changes in the training set result in large changes in predictions. Breiman (1996c) claimed that neural networks and decision trees are example of unstable learning algorithms.

The boosting literature (Schapire, Freund, Bartlett, & Lee, 1997) has recently suggested (based on a few data sets with decision trees) that it is possible to further reduce the test-set error even after ten members have been added to an ensemble (and they note that this result also applies to bagging).

In this work, bagging is evaluated on real and benchmark data sets of intrusion detection, direct marketing, and signature verification in conjunction with support vector machine as the base learner. The performance of the proposed bagged SVM classifier is examined in comparison with standalone SVM.

III. CLASSIFICATION METHODS

3.1 Existing Support Vector Machine

Support vector machines (Cherkassky *et al*., 1998; Burges, 1998) are powerful tools for data classification. Classification is achieved by a linear or nonlinear separating surface in the input space of the dataset. The separating surface depends only on a subset of the original data. This subset of data, which is all that is needed to generate the separating surface, constitutes the set of support vectors. In this study, a method is given for selecting as small a set of support vectors as possible which completely determines a separating plane classifier. In nonlinear classification problems, SVM tries to place a linear boundary between two different classes and adjust it in such a way that the margin is maximized (Vanajakshi and Rilett, 2004). Moreover, in the case of linearly separable data, the method is to find the most suitable one among the hyperplanes that minimize the training error. After that, the boundary is adjusted such that the distance between the boundary and the nearest data points in each class is maximal.

3.2 Proposed Bagged Support Vector Machine

Given a set D, of d tuples, bagging works as follows. For iteration i (i =1, 2,…..k), a training set, D_i, of d tuples is sampled with replacement from the original set of tuples, D. The bootstrap sample D_i, by sampling D with replacement, from the given training data set D repeatedly. Each example in the given training set D may appear repeated times or not at all in any particular replicate training data set D_i. A classifier model, M_i, is learned for each training set, D_i. To classify an unknown tuple, X, each classifier, M_i, returns its class prediction, which counts as one vote. The bagged SVM, M^*, counts the votes and assigns the class with the most votes to X.

Algorithm: SVM ensemble classifier using bagging

Input:

- D, a set of d tuples.
- k = 1, the number of models in the ensemble.
- Base Classifier (Support Vector Machine)

Output: A Bagged SVM, M^*

Method:

1. for i = 1 to k do // create k models
2. Create a bootstrap sample, D_i, by sampling D with replacement, from the given training data set D repeatedly. Each example in the given training set D may

appear repeated times or not at all in any particular replicate training data set D_i

3. Use D_i to derive a model, M_i;
4. Classify each example d in training data D_i and initialized the weight, W_i for the model, M_i, based on the accuracies of percentage of correctly classified example in training data D_i.
5. endfor

To use the bagged SVM model on a tuple, X:

1. if classification then
2. let each of the k models classify X and return the majority vote;
3. if prediction then
4. let each of the k models predict a value for X and return the average predicted value;

IV. PERFORMANCE EVALUATION MEASURES

4.1 Cross Validation Technique

Cross-validation (Jiawei Han and Micheline Kamber, 2003) sometimes called rotation estimation, is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. 10-fold cross validation is commonly used. In stratified K-fold cross-validation, the folds are selected so that the mean response value is approximately equal in all the folds.

4.2 Criteria for Evaluation

The primary metric for evaluating classifier performance is classification Accuracy: the percentage of test samples that are correctly classified. The accuracy of a classifier refers to the ability of a given classifier to correctly predict the label of new or previously unseen data (i.e. tuples without class label information). Similarly, the accuracy of a predictor refers to how well a given predictor can guess the value of the predicted attribute for new or previously unseen data.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Intrusion Detection

5.1.1 Real world Dataset Description

The Acer07 dataset, being released for the first time is a real world data set collected from one of the sensors in Acer e-DC (Acer e-Enabling Data Center). The data used for evaluation is the inside packets from August 31, 2007 to September 7, 2007.

5.1.2 Bench Mark Dataset Description

The data used in classification is NSL-KDD, which is a new dataset for the evaluation of researches in network intrusion detection system. NSL-KDD consists of selected records of the complete KDD'99 dataset (Ira Cohen, et al., 2007). NSL-KDD dataset solve the issues of KDD'99 benchmark [KDD'99 dataset]. Each NSL-KDD connection record contains 41 features (e.g., protocol type, service, and ag) and is labeled as either normal or an attack, with one specific attack type.

5.2 Direct Marketing

5.2.1 Real world Dataset Description

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be (or not) subscribed. The classification goal is to predict if the client will subscribe a term deposit (variable y).

5.2.2 Bench Mark Dataset Description

The data includes all collective agreements reached in the business and personal services sector for locals with at least 500 members (teachers, nurses, university staff, police, etc) in Canada in 87 and first quarter of 88. Data was used to test 2 tier approach with learning from positive and negative examples.

5.3 Signature Verification

5.3.1 Real world Dataset Description

The dataset used to train and test the systems described in this paper was constructed from NIST's Special Database 3 and Special Database 1 which contain binary images of handwritten digits. NIST originally designated SD-3 as their training set and SD-1 as their test set. However, SD-3 is much cleaner and easier to recognize than SD-1. The reason for this can be found on the fact that SD-3 was collected among Census Bureau employees, while SD-1 was collected among high-school students. Drawing sensible conclusions from learning experiments requires that the result be independent of the choice of training set and test among the complete set of samples. Therefore it was necessary to build a new database by mixing NIST's datasets.

5.3.2 Bench Mark Dataset Description

The data used in classification is 10 % U.S. Zip code, which consists of selected records of the complete U.S. Zip code database. The database used to train and test the hybrid system consists of 4253 segmented numerals digitized from handwritten zip codes that appeared on U.S. mail passing through the Buffalo, NY post office. The digits were written by many different people, using a great variety of sizes, writing styles, and instruments, with widely varying amounts of care.

5.3 Experiments and Analysis

5.3.1 Intrusion Detection

5.3.1.1 Real world Dataset

The Acer07dataset is taken to evaluate the proposed bagged SVM for intrusion detection system.

Table 1: The Performance of Existing and Proposed Bagged Classifier for real world dataset

Real Dataset	Classifiers	Classification Accuracy
Acer07 dataset	Existing SVM Classifie	99.80 %
	Proposed Bagged SVM Classifier	99.93 %

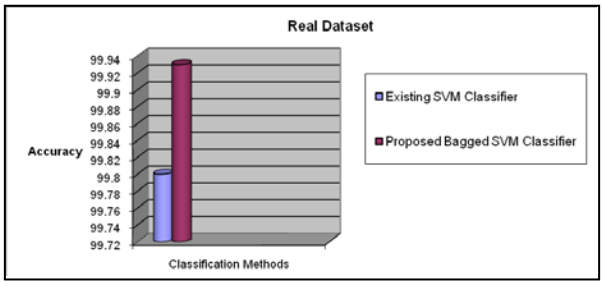

Figure 1: Classification Accuracy of Existing and Proposed Bagged SVM Classifier using Real Dataset

5.3.1.2 Bench Mark Dataset

The NSL- KDD dataset is taken to evaluate the proposed bagged SVM for intrusion detection system.

Table 2: The Performance of Existing and Proposed Bagged Classifier for bench mark dataset

Bench Mark Dataset	Classifiers	Classification Accuracy
NSL- KDD dataset	Existing SVM Classifi	91.81 %
	Proposed Bagged SVM Classifier	92.03 %

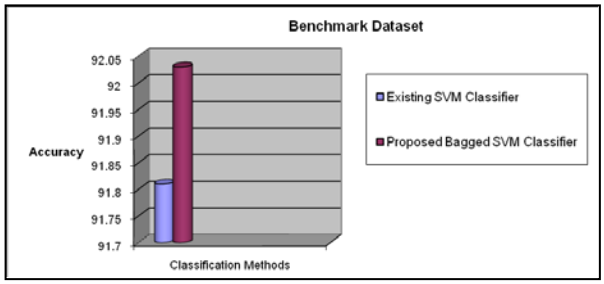

Figure 2: Classification Accuracy of Existing and Proposed Bagged SVM Classifier using Benchmark Dataset

5.3.2 Direct Marketing

In this section, new ensemble classification method is proposed using bagging classifier and its performance is analyzed in terms of accuracy.

5.3.2.1 Real world Dataset

The bank marketing dataset is taken to evaluate the proposed bagged SVM classifier.

Table 3: The Performance of Existing and Proposed Bagged Classifier for real world dataset

Real Dataset	Classifiers	Classification Accuracy
Bank Marketing dataset	Existing SVM Classifier	69.00 %
	Proposed Bagged SVM	73.33 %

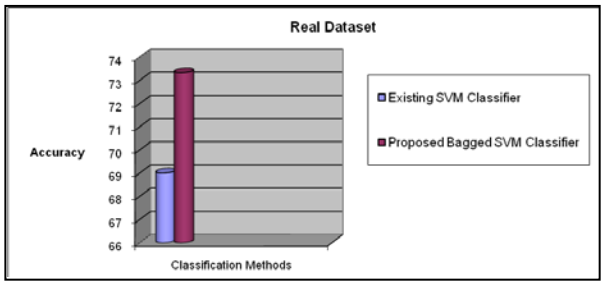

Figure 3: Classification Accuracy of Existing and Proposed Bagged SVM Classifier using Real Dataset

5.3.2.2 Bench Mark Dataset

The labor relations dataset is taken to evaluate the proposed bagged SVM classifier.

Table 4: The Performance of Existing and Proposed Bagged Classifier for benchmark dataset

Benchmark Dataset	Classifiers	Classification Accuracy
Labor Relations Dataset	Existing SVM Classifier	89.47 %
	Proposed Bagged SVM	96.49 %

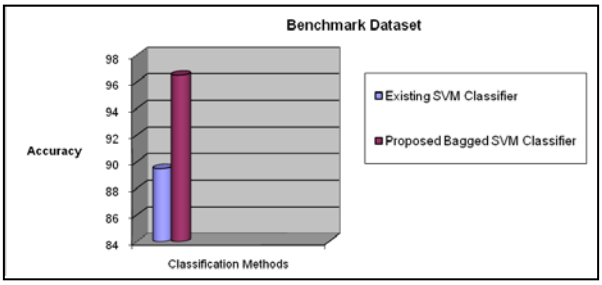

Figure 4: Classification Accuracy of Existing and Proposed Bagged SVM Classifier using benchmark Dataset

5.3.3 Signature Verification

5.3.3.1 Real world Dataset

The NIST dataset are taken to evaluate the proposed bagged SVM for handwriting recognition system.

Table 5: The Performance of Existing and Proposed Bagged Classifier for real world dataset

Real Dataset	Classifiers	Classification Accuracy
NIST dataset	Existing SVM Classifie	89.20 %
	Proposed Bagged SVM Classifier	98.00 %

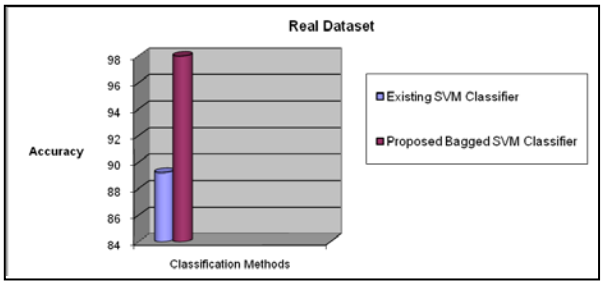

Figure 5: Classification Accuracy of Existing and Proposed Bagged SVM Classifier using Real Dataset

5.3.3.2 Bench Mark Dataset

The U.S. Zip code dataset are taken to evaluate the proposed bagged SVM for handwriting recognition system.

Table 6: Classification Accuracy of Existing and Proposed Bagged Classifier for bench mark dataset

Bench Mark Dataset	Classifiers	Classification Accuracy
U.S. Zip code dataset	Existing SVM Classifier	93.98 %
	Proposed Bagged SVM Classifier	95.45 %

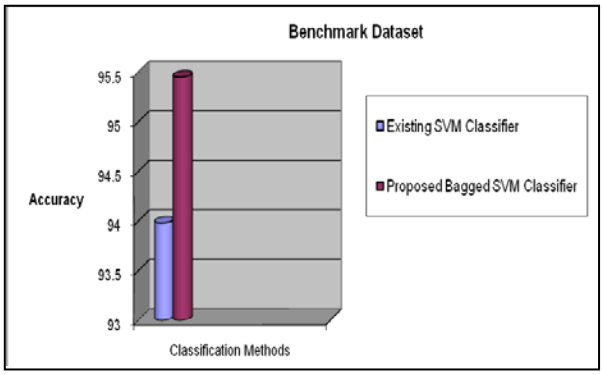

Figure 6: Classification Accuracy of Existing and Proposed Bagged SVM Classifier using benchmark Dataset

In this research work, new ensemble classification method is proposed using bagging classifier in conjunction with support vector machine as the base learner and the performance is analyzed in terms of accuracy. Here, the base classifiers are constructed using support vector machine. 10-fold cross validation (Kohavi, R, 1995) technique is applied to the base classifier and evaluated classification accuracy. Bagging is performed with support vector machine to obtain a very good classification performance. Table 1 to 6 shows classification performance for real and benchmark datasets of intrusion detection, direct marketing, signature verification using existing and proposed bagged support vector machine. The analysis of results shows that the proposed bagged support vector machine are shown to be superior to individual approach for data mining applications in terms of classification accuracy. According to Fig. 1 to 6 proposed combined model show significantly larger improvement of classification accuracy than the base classifier. This means that the combined method is more accurate than the individual method for the data mining applications.

The χ^2 statistic is determined for the above approach and the critical value is found to be less than 0.455. Hence corresponding probability is $p < 0.5$. This is smaller than the conventionally accepted significance level of 0.05 or 5%. Thus examining a χ^2 significance table, it is found that this value is significant with a degree of freedom of 1. In general, the result of χ^2 statistic analysis shows that the proposed classifier is significant at $p < 0.05$ than the existing classifier.

VI. CONCLUSION

In this research work, new combined classification method is proposed using bagging classifier in conjunction with support vector machine as the base learner and the performance comparison has been demonstrated using real and benchmark dataset of intrusion detection, direct marketing, signature verification in terms of accuracy. This research has clearly shown the importance of using ensemble approach for data mining applications like intrusion detection, direct marketing, and signature verification. An ensemble helps to indirectly combine the synergistic and complementary features of the different learning paradigms without any complex hybridization. Since all the considered performance measures could be optimized, such systems could be helpful in several real world data mining applications. The high classification accuracy has been achieved for the ensemble classifier compared to that of single classifier. The proposed bagged support vector machine is shown to be significantly higher improvement of classification accuracy than the base classifier. The real and benchmark dataset of intrusion detection, direct marketing, signature verification could be detected with high accuracy for homogeneous model. The future research will be directed towards developing more accurate base classifier particularly for the data mining applications.

ACKNOWLEDGMENT

Author gratefully acknowledges the authorities of Annamalai University for the facilities offered and encouragement to carry out this work.

REFERENCES

[1] Aptéa, C. and Weiss, S. Data mining with decision trees and decision rules, Future Generation Computer Systems 13, n2-3, 1997, pp.197–210.

[2] J.P. Anderson. Computer security threat monitoring and surveillance, Technical Report, James P. Anderson Co., Fort Washington, PA, 1980.

[3] E. Biermann, E. Cloete and L.M. Venter. A comparison of intrusion detection Systems, Computer and Security, v(20), 2001, pp. 676-683.
[4] Breiman, L. Stacked Regressions, Machine Learning, 24(1), 1996c, pp.49-64.
[5] Burges, C. J. C. A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery, 2(2), 1998, pp.121-167.
[6] C. J. C. Burges and B. Scholkopf. Improving the Accuracy and Speed of Support vector Learning Machine, Advanced in Neural Information Processing Systems 9, MIT Press, Cambridge, MA, 1997, PP. 375-381.
[7] Cherkassky, V. and Mulier, F. Learning from Data - Concepts, Theory and Methods, John Wiley & Sons, New York, 1998.
[8] Cheung, K.-W., Kwok, J. K., Law, M. H., & Tsui, K.-C. Mining customer product rating for personalized marketing, Decision Support Systems, 35, 2003, pp. 231–243.
[9] Cortes, C. and Vapnik, V. Support Vector Networks, Machine Learning, 20, n3, 1995, pp.273–297.
[10] J. X. Dong, A. Krzyzak, and C.Y. Suen. Fast SVM Training Algorithm with Decomposition on Very Large Datasets, IEEE Trans. Pattern Analysis and Machine Intelligence, v(27), n 4, 2005, pp. 603-618.
[11] Heady R, Luger G, Maccabe A, Servilla M. The architecture of a network level intrusion detection system, Technical Report, Department of Computer Science, University of New Mexico, 1990.
[12] T.K.Ho, J.J.Hull, and S.N.Srihari. Combination of Structural Classifiers, Proc. IAPR Workshop Syntatic and Structural Pattern Recog, 1990, pp. 123-137.
[13] Hu, X. A data mining approach for retailing bank customer attrition analysis, Applied Intelligence 22(1), 2005, pp.47-60.
[14] K. Ilgun, R.A. Kemmerer and P.A. Porras. State transition analysis:A rule-based intrusion detection approach, IEEE Trans. Software Eng. V(21),1995, pp. 181-199.
[15] Ira Cohen, Qi Tian, Xiang Sean Zhou and Thoms S.Huang. Feature Selection Using Principal Feature Analysis, Proceedings of the 15th international conference on Multimedia, Augsburg, Germany, September, 2007, pp. 25-29.
[16] Jiawei Han, Micheline Kamber. Data Mining – Concepts and Techniques, Elsevier Publications, 2003.
[17] U. Krebel. Pairwise Classification and Support Vector Machines, Advances in Kernel Methods: Support Vector Learning, MIT Press, Cambridge, MA, 1999, pp. 255-268.
[18] Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection, Proceedings of International Joint Conference on Artificial Intelligence, 1995, pp.1137–1143.
[19] Li, W., Wu, X., Sun, Y. and Zhang, Q. Credit Card Customer Segmentation and Target Marketing Based on Data Mining, Proceedings of International Conference on Computational Intelligence and Security, 2010, pp.73-76.
[20] Ling, X. and Li, C. Data Mining for Direct Marketing: Problems and Solutions. Proceedings of the 4th KDD conference, AAAI Press, 1998, pp.73–79.
[21] E. Lundin and E. Jonsson. Anomaly-based intrusion detection: privacy concerns and other problems, Computer Networks, v(34), 2002, pp. 623-640.
[22] D. Marchette. A statistical method for profiling network traffic, proceedings of the First USENIX Workshop on Intrusion Detection and Network Monitoring (Santa Clara), CA, 1999, pp. 119-128.
[23] Mukkamala S, Sung AH, Abraham A. Intrusion detection using ensemble of soft computing paradigms, third international conference on intelligent systems design and applications, intelligent systems design and applications, advances in soft computing. Germany: Springer, 2003, pp. 239–48.
[24] Mukkamala S, Sung AH, Abraham A. Modeling intrusion detection systems using linear genetic programming approach, The 17th international conference on industrial & engineering applications of artificial intelligence and expert systems, innovations in applied artificial intelligence. In: Robert O., Chunsheng Y., Moonis A., editors. Lecture Notes in Computer Science, vol. 3029. Germany: Springer, 2004a, pp. 633–42.
[25] Mukkamala S, Sung AH, Abraham A, Ramos V. Intrusion detection systems using adaptive regression splines. In: Seruca I, Filipe J, Hammoudi S, Cordeiro J, editors. Proceedings of the 6th international conference on enterprise information systems, ICEIS'04, vol. 3, Portugal, 2004b, pp. 26–33.
[26] S. Mukkamala, G. Janoski and A.Sung. Intrusion detection: support vector machines and neural networks" In proceedings of the IEEE International Joint Conference on Neural Networks (ANNIE), St. Louis, MO, 2002, pp. 1702-1707.
[27] Schapire, R., Freund, Y., Bartlett, P., and Lee, W. (1997). Boosting the margin: A new explanation for the effectives of voting methods. In proceedings of the fourteenth International Conference on Machine Learning, Nashville, TN, 1997, pp. 322-330.
[28] Shah K, Dave N, Chavan S, Mukherjee S, Abraham A, Sanyal S. (2004), Adaptive neuro-fuzzy intrusion detection system. IEEE International Conference on Information Technology: Coding and Computing (ITCC'04), v(1), USA: IEEE Computer Society, 2004, pp. 70–74.
[29] Shin, H., Cho, S. "Response Modeling with Support vector Machines", Expert Systems with Applications, 30, 2006, pp. 746-760.
[30] T. Shon and J. Moon. A hybrid machine learning approach to network anomaly detection", Information Sciences, v(177), 2007, pp. 3799-3821.
[31] C.Y.Suen, C.Nadal, T.A.Mai, R.Legault, and L.Lam. Recognition of totally unconstrained handwritten

numerals based on the concept of multiple experts, Frontiers in Handwriting Recognition , C.Y.Suen, Ed., IN Proc.Int.Workshop on Frontiers in Handwriting Recognition, Montreal, Canada, Apr. 2-3, 1990, pp. 131-143.
[32] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai, and L. Lam. Computer recognition of unconstrained handwritten numerals, Proc. IEEE, v(80), 1992, pp. 1162–1180.
[33] Summers RC. Secure computing: threats and safeguards. New York, McGraw-Hill, 1997.
[34] Sundaram A. An introduction to intrusion detection. ACM Cross Roads, 2(4), 1996.
[35] W. Stallings. Cryptography and network security principles and practices, USA: Prentice Hall, 2006
[36] C. Tsai , Y. Hsu, C. Lin and W. Lin. Intrusion detection by machine learning: A review, Expert Systems with Applications, v(36), 2009, pp.11994-12000.
[37] T. Verwoerd and R. Hunt. Intrusion detection techniques and approaches, Computer Communications, v(25), 2002, pp.1356-1365.
[38] Vanajakshi, L. and Rilett, L.R. A Comparison of the Performance of Artificial Neural Network and Support Vector Machines for the Prediction of Traffic Speed, IEEE Intelligent Vehicles Symposium, University of Parma, Parma, Italy, IEEE, 2004, pp.194-199.
[39] Viaene, S., Baesens, B., Van Gestel, T., Suykens, J. A. K., Van den Poel, D., Vanthienen, J., et al. Knowledge discovery in a direct marketing case using least squares support vector machines, International Journal of Intelligent Systems, 16, 2001b, pp.1023–1036.
[40] Vapnik, V. (1998). Statistical learning theory, New York, John Wiley & Sons, 1998.
[41] S. Wu and W. Banzhaf. The use of computational intelligence in intrusion detection systems: A review, Applied Soft Computing, v(10), 2010, pp. 1-35.
[42] Zhang, H. The Optimality of Naïve Bayes, Proceedings of the 17th FLAIRS conference, AAAI Press, 2004.

M.Govindarajan received the B.E and M.E and Ph.D Degree in Computer Science and Engineering from Annamalai University, Tamil Nadu, India in 2001 and 2005 and 2010 respectively. He did his post-doctoral research in the Department of Computing, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey, United Kingdom in 2011 and pursuing Doctor of Science at Utkal University, orissa, India. He is currently an Assistant Professor at the Department of Computer Science and Engineering, Annamalai University, Tamil Nadu, India. He has presented and published more than 75 papers at Conferences and Journals and also received best paper awards. He has delivered invited talks at various national and international conferences. His current Research Interests include Data Mining and its applications, Web Mining, Text Mining, and Sentiment Mining. He was the recipient of the Achievement Award for the field and to the Conference Bio-Engineering, Computer science, Knowledge Mining (2006), Prague, Czech Republic, Career Award for Young Teachers (2006), All India Council for Technical Education, New Delhi, India and Young Scientist International Travel Award (2012), Department of Science and Technology, Government of India New Delhi. He is Young Scientists awardee under Fast Track Scheme (2013), Department of Science and Technology, Government of India, New Delhi and also granted Young Scientist Fellowship (2013), Tamil Nadu State Council for Science and Technology, Government of Tamil Nadu, Chennai. He has visited countries like Czech Republic, Austria, Thailand, United Kingdom, Malaysia, U.S.A, and Singapore. He is an active Member of various professional bodies and Editorial Board Member of various conferences and journals.

