International Journal of Engineering Research and Development
e-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com
Volume 6, Issue 6 (March 2013), PP. 16-20

Meta Learning Algorithms for Credit Card Fraud Detection

[1] Sanjay Kumar Sen, [2] Prof. Dr Sujata Dash,
[1]Asst. Professor (Comp Sc & Engg.), OEC, Bhubaneswar
[2] DEAN(R & D) & PROFESSOR(CSE), GIFT,Bhubaneswar

Abstract:- Due to the rapid advancement of electronic commerce technology, there is a great and dramatic increase in credit card transactions. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising; to detect credit card frauds in electronic transactions becomes the focus of risk of control of banks. The proposed work in this paper is the combination of five supervised machine learning algorithms, Classification and Regression Tree (CART), Adaboost and Logitboost,Bagging and Dagging are proposed for classification of credit card data. These resulted forms help researchers to detect fraud in credit card. The experimental result shows the performance analysis of different meta-learning algorithms and also compared on the basis of misclassification and correct classification rate. Smaller misclassification reveals that bagging algorithm performs better classification of Credit card fraud detection technique.

Keywords:- Classification, Fraud detection, meta-learning algorithm, Machine Learning Algorithm, misclassification, classification.

I. INTRODUCTION

Data mining helps the use of complicated data analysis to discover valid patterns which are previously unknown and relationships among large data sets. These tools have mathematical algorithms, statistical models and machine learning methods. Data mining comprises of more than collection and management of data, representation in textual, qualitative or multimedia forms. Data mining applications uses a range of parameters to observe and analyse the data which includes association, classification, sequence or path analysis, clustering and forecasting.

The popularity of online shopping is growing day by day. According to an ACNielsen study conducted in 2005, one-tenth of the world's population is shopping online [1].Germany and Great Britain have the largest number of online shoppers, and credit card is the most popular mode of payment (59 percent). About 350 million transactions per year were reportedly carried out by Barclaycard, the largest credit card company in the United Kingdom, toward the end of the last century [2]. Retailers like Wal-Mart typically handle much larger number of credit card transactions including online and regular purchases. As the number of credit card users rises world-wide, the opportunities for attackers to steal credit card details and, subsequently, commit fraud are also increasing. The total credit card fraud in the United States itself is reported to be $2.7 billion in 2005 and estimated to be $3.0 billion in 2006, out of which $1.6 billion and $1.7 billion, respectively, are the estimates of online fraud [3].

The use of credit cards is common in modern day society. Fraud is a millions dough business and it is rising every year. Fraud presents significant cost to our financial prudence measure world wide .Modern techniques based on Data mining, Machine learning, Sequence Alignment technique, Fuzzy Logic, Genetic Programming, Artificial Intelligence (AI) etc. ,has been introduced for detecting & preventing credit/ATM card, CHEQUE book type of fraudulent transactions.

With great increase in credit cards, fraud has increasing excessively now a days. Since credit card becomes the most general mode of payment or both online as line as regular not only capturing the deceptive events but capturing such activities as rapidly as possible.

II. RELATED WORK ON CREDIT CARD FRAUD DETECTION

Credit card fraud detection has drawn a lot of research interest and a number of techniques, with special emphasis on data mining and neural networks, have been suggested. Ghosh and Reilly [4] have proposed credit card fraud detection with a neural network. They have built a detection system, which is trained on a large sample of labelled credit card account transactions. These transactions contain example fraud cases due to lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud, and non received issue (NRI) fraud. Recently, Syeda et al. [5] have used parallel granular neural networks (PGNNs) for improving the speed of data mining and knowledge discovery process in credit card fraud detection. A complete system has been implemented for this purpose. Stolfo et al. [6] suggest a credit card fraud detection system (FDS) using meta learning techniques to learn models of fraudulent credit card transactions. Meta learning is a general strategy that provides a means for combining and integrating a number of separately built classifiers or models. Aleskerov et al. [7] present CARDWATCH, a database mining system used for credit card fraud detection. The system, based on a neural learning module, provides an interface to a variety of commercial databases. Kim and Kim have identified skewed distribution of data and mix of legitimate and fraudulent transactions as the two main reasons for the complexity of credit card fraud detection [8]. Fan et al. [9] suggest the application of distributed data mining in credit card fraud detection. Brause et al. [10] have developed an approach that involves advanced data mining techniques and neural network algorithms to obtain high fraud coverage. Chiu and Tsai [11] have proposed Web services and data mining techniques to establish a collaborative scheme for fraud detection in the banking industry. [12] have done an extensive survey of existing data-mining-based FDSs and published a comprehensive report. Prodromidis and Stolfo [13] use an agent-based approach with distributed learning for detecting frauds in credit card transactions. It is based on artificial intelligence and combines inductive learning algorithms and meta learning methods for achieving higher accuracy. Phua et al. [14] suggest the use of meta classifier similar to [6] in fraud detection problems. They consider naive Bayesian, C4.5, and Back Propagation neural networks as the base classifiers. Vatsa et al. [15] have recently proposed a game-theoretic approach to credit card fraud detection. They model the interaction between an attack erandan FDS as a multistage game between two players, each trying to maximize his payoff.

III. MACHINE LEARNING ALGORITHMS

i. Classification and Regression Tree(CART)

It was introduce by Breimann 1984.It builds both classification and regression tree (Gini index measure is used for selecting splitting attribute. Pruning is done on training data set. It can deal with both numeric and categorical attributes and can also handle missing attributes. [16].

The CART monograph focus on the Gini rule which is similar to the better know entropy or information gain criterion [17]. For binary (0/1) target the 'Gini measure of impurity" of a node t is:

Classification and regression free provide automatic construction of new features within each node and for the binary target.

ii. Adaboost

AdaBoost, short for Adaptive boosting is a machine algorithm, formulated by Yeave Freud and Robert Scapire. It is a meta-algorithm and can be used in conjunction with many other learning algorithms to improve their performance. AdaBoost is an algorithm for constructing a"strong" classifier as linear combination. AdaBoost is adaptive in the sense that subsequent classifiers built are tweaked in favour of those instances misclassified by previous classifiers. AdaBoost is sensitive to noisy data and outliers. In some problems, however, it can be less susceptible to the over fitting problem than most learning algorithms. The classifiers it uses can be weak (i.e., display a substantial error rate), but as long as their performance is not random (resulting in an error rate of 0.5 for binary classification), they will improve the final model. Even classifiers with an error rate higher than would be expected from a random classifier will be useful, since they will have negative coefficients in the final linear combination of classifiers and hence behave like their inverses.

iii. Bagging

The way of combining the decisions of different models means amalgamating the various outputs into a single prediction. The way of doing to do this is to calculate the average. In bagging the models receives equal weights. In case of bagging suppose that several training datasets of the same size are chosen at random from the problem domain. Suppose using a particular machine learning technique to build a decision dtree for each dataset, we might expect these trees to be practically identically and to make the same prediction for each new test instance. This is a disturbing fact and seems to cast a shadow over the whole enterprise.

In bagging the models receive equal weights, whereas in boosting weighting is used to give more influence to the more successful one just as an executive might place different values on the advice of different experts depending on how experienced they are. To introduce bagging, several training datasets of the same size are chosen at random from the problem domain. Suppose using a particular machine learning technique to build a decision tree for each dataset, we might expect these trees to be practically identically and to make the same prediction for each new test instance.

iv. Logitboost

LogitBoost is a boosting algorithm formulated by Jerome Friedmome, Trevor Hastie, and Robert Tibshirani. The original paper casts the Adaboost algorithm into a statistical framework. Specifically, if one considers AdaBoost as a generalized additive model and then applies the cost functional of logistic regression one can derive the LogitBoost algorithm.

v. Grading

We investigate another technique, which we call *grading*. The basic idea is to learn to predict for each of the original learning algorithms whether its prediction for a particular example is correct or not. We therefore train one classifier for each of the original learning algorithms on a training set that consists of the original examples with class labels that encode whether the prediction of this learner was correct on this particular example. The algorithm may also be viewed as an attempt to extend the work of Bay and Pazzani (2000)[18] who propose to use a meta-classification scheme for characterizing model errors. Hence in contrast to stacking—we leave the original examples unchanged, but instead modify the class labels. The algorithm may also be viewed as an attempt to extend the work of Bay and Pazzani (2000)[19] who propose to use a meta-classification scheme for characterizing model errors. Their suggestion is to learn a comprehensible theory that describes the regions of errors of a given classifier. While the step of constructing the training set for the meta classifier is basically the same as in our approach,1 their approach is restricted to learning descriptive characterizations but cannot be directly used for improving classifier performance. The reason is that negative feedback when the meta classifier predicts that the base classifier is wrong only rules out the class predicted by the base classifier, but does not help to choose among the remaining classes (except, of course, for two-class problems).

IV. PERFORMANCE COMPARISONS

In literature classification algorithm, classifier performance can be measured on the same data. On the basis of results obtained Bagging algorithm is found better than other four algorithms. Comparisons of machine learning algorithms have been done on the basis of the misclassification and correct classification rate. It is observed that BAGGING machine learning classifier performance is better than Classification and Regression Technique, Adaboost, Logitboost, Bagging and Grading linear and quadratic dis-criminant analysis classifier in context of misclassification rate and correct classification rate.

Table 1

Algorithm misclassification	Correct Classification Rate	Mis- Classification Rate
CART	0.834	0.166
ADABOOST	0.847	0.153
LOGITBOOST	0.855	0.145
BAGGING	0.877	0.123
GRADING	0.536	0.464

According to obtained results of classification in table 1 following graph can be drawn.

Graph for Mis-classification rate

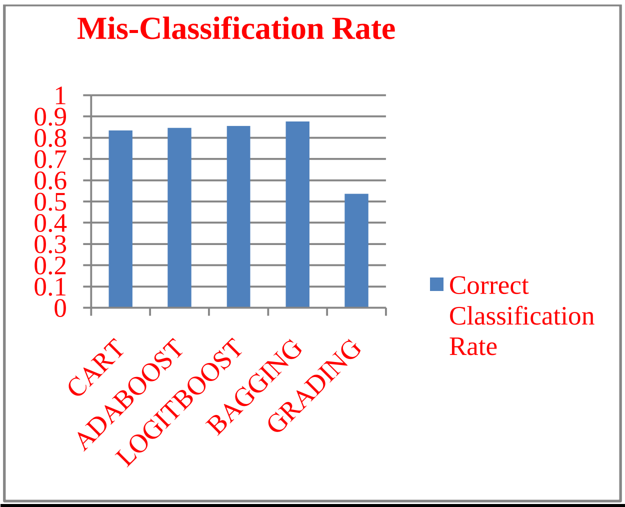

Fig.1. Performace comparision graph between between CART,ADABOOST,LOGITBOOST,BAGGING, GRADING with respect to Mis-classification error rate

Graph for correct classification rate

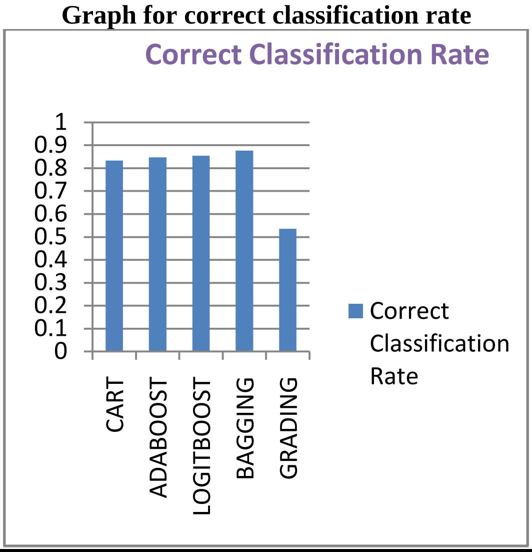

Fig.2. Performace comparision graph between between CART,ADABOOST,LOGIBOOST,BAGGING, GRADING with respect to correct classification error rate

Graph for Correct Classification and Misclassification rate

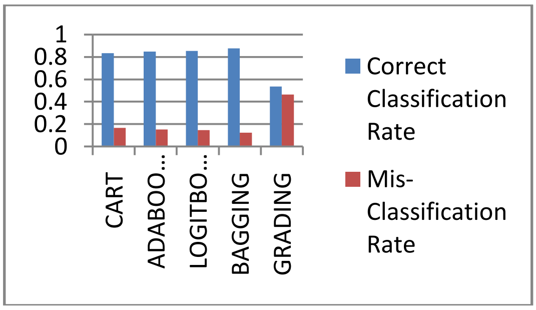

Figure 3. Performance comparison graph between CART, ADABOOST, LOGIBOOST, BAGGING AND GRADING algorithm with respect to misclassification error rate and correct classification error rate.

V. CONCLUSIONS

This paper represents computational issues of five supervised machine learning algorithms i.e. Adaboost algorithm, Logitboost algorithm, Classification and Regression technique algorithm, Bagging algorithm and Dagging algorithm with dedicating role of detection of credit fraud rating on the basis of classification rule. Among five algorithms, Bagging algorithm is the best because the Bagging algorithm is easier to interpret and understand as compared to Adaboost,Llogitboost,Classification and Regression Tree algorithm and Dagging algorithm. In order to compare the classification performance of three machine learning algorithm, classifiers are applied on same data and results are compared on the basis of misclassification and correct classification rate and according to experimental results in table 1, it can be concluded that bagging algorithm is the best as compared to classification and regression tree, is best as compared to adaboost, logitboost, classification and regression algorithm and Grading algorithm.

REFERENCES

[1]. "Global Consumer Attitude Towards On-Line Shopping, "http://www2.acnielsen.com/reports/documents/2005_cc_online shopping.pdf, Mar. 2007.

[2]. D.J. Hand, G. Blunt, M.G. Kelly, and N.M. Adams, "Data Mining for Fun and Profit," Statistical Science, vol. 15, no. 2, pp. 111-131,2000.

[3]. "Statistics for General and On-Line Card Fraud,"http://www.epaynews.com/statistics/fraud.html, Mar. 2007.

[4]. S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int'l Conf. System Sciences:Information Systems: Decision Support and Knowledge-Based Systems, vol. 3, pp. 621-630, 1994.

[5]. M. Syeda, Y.Q. Zhang, and Y. Pan, "Parallel Granular Networks for Fast Credit Card Fraud Detection," Proc. IEEE Int'l Conf. Fuzzy Systems, pp. 572-577, 2002.

[6]. S.J. Stolfo, D.W. Fan, W. Lee, A.L. Prodromidis, and P.K. Chan, "Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results," Proc. AAAI Workshop AI Methods in Fraud and Risk Management, pp. 83-90, 1997.

[7]. E. Aleskerov, B. Freisleben, and B. Rao, "CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection," Proc. IEEE/IAFE: Computational Intelligence for Financial Eng., pp. 220-226, 1997.

[8]. M.J. Kim and T.S. Kim, "A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection," Proc. Int'l Conf. Intelligent Data Eng. and Automated Learning, pp. 378-383, 2002.

[9]. W. Fan, A.L. Prodromidis, and S.J. Stolfo, "Distributed Data Mining in Credit Card Fraud Detection," IEEE Intelligent Systems, vol. 14, no. 6, pp. 67-74, 1999.

[10]. R. Brause, T. Langsdorf, and M. Hepp, "Neural Data Mining for Credit Card Fraud Detection," Proc. IEEE Int'l Conf. Tools with Artificial Intelligence, pp. 103-106, 1999.

[11]. C. Chiu and C. Tsai, "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection," Proc. IEEE Int'l Conf.e-Technology, e-Commerce and e-Service, pp. 177-181, 2004.

[12]. C. Phua, V. Lee, K. Smith, and R. Gayler, "A Comprehensive Survey of Data Mining-Based Fraud Detection Research," http://www.bsys.monash.edu.au/people/cphua/, Mar. 2007.

[13]. S. Stolfo and A.L. Prodromidis, "Agent-Based Distributed Learning Applied to Fraud Detection," Technical Report CUCS-014-99, Columbia Univ., 1999.

[14]. C. Phua, D. Alahakoon, and V. Lee, "Minority Report in Fraud Detection: Classification of Skewed Data," ACM SIGKDD Explorations Newsletter, vol. 6, no. 1, pp. 50-59, 2004.

[15]. V. Vatsa, S. Sural, and A.K. Majumdar, "A Game-theoretic Approach to Credit Card Fraud Detection," Proc. First Int'l Conf. Information Systems Security, pp. 263-276, 2005.

[16]. Matthew N. Anyanwu and Sajjan G. Shiva," Comparative Analysis of Serial Decision Tree Classification Algorithms," International Journal of Computer Science and Security, (IJCSS) Volume (3) : Issue (3) ,.pp:230-240.

[17]. XindongWu ,Vipin Kumar , J. Ross Quinlan ,Joydeep Ghosh , Qiang Yang ,Hiroshi Motoda , et al "Top 10 algorithm in data mining according to the survey paper of Xindong wu et al know (inf syst (2008) @springer verlog London limited 2007,pp:1-37, Dec. 2007.

[18]. Bay, S. D., & Pazzani, M. J. (2000). Characterizing model errors and differences. In *Proceedings of the 17th International Conference on Machine Learning (ICML- 2000)*. Morgan Kaufmann.

[19]. Bay, S. D., & Pazzani, M. J. (2000). Characterizing model errors and differences. In *Proceedings of the 17th International Conference on Machine Learning (ICML- 2000)*. Morgan Kaufmann.