Detection of Credit Card Fraud Transactions using Machine Learning based Algorithm

Swetha T[1], Bellam Narendra Nath[2], Manjunath[3], Govinda N[4], H V Naveen Kumar[5]
Project Guide, Department of Computer Science and Engineering[1]
Projecties, Department of Computer Science and Engineering[2,3,4,5]
Sri Jagadguru Chandrashekaranatha Swamiji Institute of Technology, Chikkaballapura, Karnataka, India

Abstract: *The rapid growth in E-Commerce industry has lead to anexponential increase in the use of credit cards for online purchases and consequently they has been surge in the fraud related to it .In recent years, For banks has become verydifficult for detecting the fraud in credit card system. Machinelearning plays a vital role for detecting the credit card fraud in the transactions. For predicting these transactions banks make use of various machine learning methodologies, past data has been collected and new features are been used for enhancing the predictive power. The performance of fraud detecting in credit card transactions is greatly affected by the sampling approach on data-set, selection of variables and detection techniques used. This paper investigates the performance of SVM, decision tree and random forest for credit card fraud detection. Dataset of credit card transactions iscollected from kaggle and it contains a total of 2,84,808 credit card transactions of a European bank data set. It considers fraud transactions as the "positive class" and genuine ones as the "negative class" .The data set is highly imbalanced, it has about 0.172% of fraud transactions and the rest are genuine transactions. The author has been done oversampling to balance the data set, which resulted in 60% offraud transactions and 40% genuine ones. The three techniques are applied for the dataset and work is implemented in R language. The performance of the techniques is evaluated for different variables based on sensitivity, specificity, accuracy and error rate. The result shows of accuracy for SVM, Decision tree and random forest classifier are 90.0, 94.3, 95.5 respectively. The comparative results show that the Random forest performs better than the SVM and decision tree techniques.*

Keywords: Fraud detection, Credit card, SVM, Decision tree, Random forest.

REFERENCES

[1]. Andrew. Y. Ng, Michael. I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes", Advances in neural information processing systems, vol. 2, pp. 841-848, 2002.

[2]. A. Shen, R. Tong, Y. Deng, "Application of classification models on credit card fraud detection", Service Systems and Service Management 2007 International Conference, pp. 1-4, 2007.

[3]. A. C. Bahnsen, A. Stojanovic, D. Aouada, B. Ottersten, "Cost sensitive credit card fraud detection using Bayes minimum risk", Machine Learning and Applications (ICMLA). 2013 12th International Conference, vol. 1, pp. 333-338, 2013.

[4]. B.Meena, I.S.L.Sarwani, S.V.S.S.Lakshmi," Web Service mining and its techniques in Web Mining" IJAEGT,Volume 2,Issue 1 , Page No.385-389.

[5]. F. N. Ogwueleka, "Data Mining Application in Credit Card Fraud Detection System", Journal of Engineering Science and Technology, vol. 6, no. 3, pp. 311-322, 2011.

[6]. G. Singh, R. Gupta, A. Rastogi, M. D. S. Chandel, A. Riyaz, "A Machine Learning Approach for Detection of Fraud based on SVM", International Journal of Scientific Engineering and Technology, vol. 1, no. 3, pp. 194-198, 2012, ISSN ISSN: 2277-1581.

[7]. K. Chaudhary, B. Mallick, "Credit Card Fraud: The study of its impact and detection techniques", International Journal of Computer Science and Network (IJCSN), vol. 1, no. 4, pp. 31-35, 2012, ISSN ISSN: 2277-5420.

[8]. M. J. Islam, Q. M. J. Wu, M. Ahmadi, M. A. Sid- Ahmed, "Investigating the Performance of Naive-Bayes Classifiers and KNearestNeighbor Classifiers", IEEE International Conference on Convergence Information

Technology, pp. 1541-1546, 2007.

[9]. R. Wheeler, S. Aitken, "Multiple algorithms for fraud detection" in Knowledge-Based Systems, Elsevier, vol. 13, no. 2, pp. 93-99, 2000.

[10]. S. Patil, H. Somavanshi, J. Gaikwad, A. Deshmane, R. Badgujar, "Credit Card Fraud Detection Using Decision Tree Induction Algorithm", International Journal of Computer Science and Mobile Computing (IJCSMC), vol. 4, no. 4, pp. 92-95, 2015, ISSN ISSN: 2320-088X.

[11]. S. Maes, K. Tuyls, B. Vanschoenwinkel, B. Manderick, "Credit card fraud detection using Bayesian and neural networks", Proceedings of the 1st international naiso congress on neuro fuzzy technologies, pp. 261-270, 2002.

[12]. S. Bhattacharyya, S. Jha, K. Tharakunnel, J. C. Westland, "Data mining for credit card fraud: A comparative study", Decision Support Systems, vol. 50, no. 3, pp. 602-613, 2011.

[13]. Y. Sahin, E. Duman, "Detecting credit card fraud by ANN and logistic regression", Innovations in Intelligent Systems and Applications (INISTA) 2011 International Symposium, pp. 315-319, 2011.

[14]. Selvani Deepthi Kavila,LAKSHMI S.V.S.S.,RAJESH B " Automated Essay Scoring using Feature Extraction Method " IJCER ,volume 7,issue 4(L), Page No. 12161-12165.

[15]. S.V.S.S.Lakshmi,K.S.Deepthi,Ch.Suresh   "Text Summarization basing on Font and Cue-phrase