Credit Card Fraud Detection Using AdaBoost and Majority Voting

ABSTRACT

Credit card fraud is a serious problem in financial services. Billions of dollars are lost due to credit card fraud every year. There is a lack of research studies on analyzing real-world credit card data owing to confidentiality issues. In this paper, machine learning algorithms are used to detect credit card fraud. Standard models are firstly used. Then, hybrid methods which use AdaBoost and majority voting methods are applied. To evaluate the model efficacy, a publicly available credit card data set is used. Then, a real-world credit card data set from a financial institution is analyzed. In addition, noise is added to the data samples to further assess the robustness of the algorithms. The experimental results positively indicate that the majority voting method achieves good accuracy rates in detecting fraud cases in credit cards.

**EXISTING SYSTEM**

* A credit card fraud detection system was proposed in [8], which consisted of a rule-based filter, Dumpster–Shafer adder, transaction history database, and Bayesian learner. The Dempster–Shafer theory combined various evidential information and created an initial belief, which was used to classify a transaction as normal, suspicious, or abnormal. If a transaction was suspicious, the belief was further evaluated using transaction history from Bayesian learning [8]. Simulation results indicated a 98% true positive rate [8]. A modified Fisher Discriminant function was used for credit card fraud detection in [9]. The modification made the traditional functions to become more sensitive to important instances. A weighted average was utilized to calculate variances, which allowed learning of profitable transactions. The results from the modified function confirm it can eventuate more profit [9].
* Association rules are utilized for extracting behavior patterns for credit card fraud cases in [10]. The data set focused on retail companies in Chile. Data samples were defuzzified and processed using the Fuzzy Query 2+ data mining tool [10]. The resulting output reduced excessive number of rules, which simplified the task of fraud analysts [10]. To improve the detection of credit card fraud cases, a solution was proposed in [11]. A data set from a Turkish bank was used. Each transaction was rated as fraudulent or otherwise. The misclassification rates were reduced by using the Genetic Algorithm (GA) and scatter search. The proposed method doubled the performance, as compared with previous results [11].

**Disadvantages**

* + There is no Majority Voting technique for credit card fraud detection.
  + There is no Machine Learning Techniques in the existing system.

**PROPOSED SYSTEM**

* In the proposed system, a total of twelve machine learning algorithms are used for detecting credit card fraud. The algorithms range from standard neural networks to deep learning models. They are evaluated using both benchmark and real world credit card data sets. In addition, the AdaBoost and majority voting methods are applied for forming hybrid models. To further evaluate the robustness and reliability of the models, noise is added to the real-world data set.
* The key contribution of this paper is the evaluation of a variety of machine learning models with a real-world credit card data set for fraud detection. While other researchers have used various methods on publicly available data sets, the data set used in this paper is extracted from actual credit card transaction information over three months.

**Advantages**

* The system is very fast due to AdaBoost Technique.
* Effective Majority Voting techniques.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**Software Requirements:**

* Operating System - Windows XP
* Coding Language - Java/J2EE(JSP,Servlet)
* Front End - J2EE
* Back End - MySQL