A Novel Approach for Credit Card Fraud Detection using Decision Tree and Random Forest Algorithms

ABSTRACT

In the world of finance, as the technology grown, new systems of business making came into picture. Credit card system is one among them. But because of lot of loop holes in this system, lot of problems are aroused in this system in the method of credit card scams. Due to this the industry and customers who are using credit cards are facing a huge loss. There is a deficiency of investigation lessons on examining practical credit card figures in arrears to privacy issues. In the manuscript an attempt has been made for finding the frauds in the credit card business by using the algorithms which adopted machine learning techniques. In this regard, two algorithms are used viz Fraud Detection in credit card using Decision Tree and Fraud Detection using Random Forest. The efficiency of the model can be decided by using some public data as sample. Then, an actual world credit card facts group from a financial institution is examined. Along with this, some clatter is supplemented to the data samples to auxiliary check the sturdiness of the systems. The significance of the methods used in the paper is the first method constructs a tree against the activities performed by the user and using this tree scams will be suspected. In the second method a user activity based forest will have constructed and using this forest an attempt will be made in identifying the suspect. The investigational outcomes absolutely show that the mainstream elective technique attains decent precision degrees in sensing scam circumstances in credit cards.

**EXISTING SYSTEM**

A. A. Akinyelu and O. Adewumi [1] has made a detail study on fraud detection using the method of natural observation of the events happened from the customer side. A. Kundu, S. A. Srivastava Sural and A. Majumdar [2], has made a detailed study on scam detection in credit card business by adopting hidden Markov model. Singh, P. K. Saraswat et al [3], worked on the machine learning oriented techniques for swarm intelligence, since the algorithms discussed in the paper is concerned about the machine learning methods, the scope of the referenced paper may be taken into considerations. Jung, J. J et al [4], has worked on the methods of collecting the data from the social media and framing them in terms of big data models and working on the challenges existed the field.

Bharill N et al [5], has made detailed study on Apache Spark which used fuzzy based clustering logic for big data analysis. Y. Sahin et al [6], demonstrated a comprehensive nominal cost model for scam detection in busines s field. The Nilson Report [7], states a detailed report on the various methods of possibility of occurrence of the fraud or scam in the field of credit card business and various methods of identifying them and the adverse effects of the scams on the business environments. J. T. Quah et al [8], described narrated towards the development of an automated model for the detection of the frauds in the business. S. Jha et al [9], implemented a system which supports in the detection of the scams or frauds in the

field of the business by recording the transactions and there by building a model using data mining models. S. Panigrahi et al [10], demonstrated Use of Dempster-Shafer theory and Bayesian inferencing for fraud detection in communication networks. T.

Fawcett et al [11], has demonstrated the use of Adaptive fraud detection, Data Mining and Knowledge Discovery. Y. Wang et al [12], has made a detailed study on Distributed Intrusion Detections Based on data fusion method. Yakub K et al [13], has demonstrated Application of GA Feature Selection on Naive Bayes Random Forest and SVM for Credit Card Fraud Detection. Zhang, R et al [14], has gave a wide explanation about Sequential Behavioral Data Processing Using Deep Learning and the Markov Transition Field in Online Fraud Detection. Zhong fang Zhuang et al [15], has demonstrated a system called Attributed Sequence Embedding where the different data sets are constructed using system.

Disadvantages

1) .The system doesn’t have technique to analyze large number of datasets.

2). There is no technique Random decision forests and Random forests which are the group learning techniques for categorization, prediction and additional jobs that function by building a gigantic volume of decision trees at exercise time and outputting the class.

**PROPOSED SYSTEM**

The proposed system defines the procedure used to hostage the credit card scam. The numerous competent approaches like arrangement orientation, device learning, neural networks, artificial intelligence, fuzzy logic are employed to detect and encounter scams in credit card businesses. Credit card fraud has become progressively widespread in modern years. In Current day, the fraud is one of the key causes of excessive business losses, not only for merchants, distinct clients are also affected. So there are some methods to detect such kind of frauds. Initially, clustering model was adopted to categorize the authorized and deceitful operation by means of data clusterization of areas of factor value. Furthermore, Gaussian mixture model is used to model the possibility thickness of credit card operator's past performance such that the chance of present actions can be intended to perceive any irregularities from the historical behavior. Finally, Bayesian networks are used to define the measurements of a specific user and the pointers of different scam circumstances.

**Advantages**

1. The proposed system offered several innovative approaches that have vastly increased the efficiency of cyber threat identification.
2. The system is more effective due to presence of Random Forest based Credit Card Fraud Detection Algorithm using Machine Learning.

**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 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).