

Credit Card Fraud Prediction

INTRODUCTION

OVERVIEW:

The credit card fraud Prediction features uses user behavior and location scanning to check for unusual patterns. These patterns include user characteristics such as user spending patterns as well as usual user geographic locations to verify his identity.

If any unusual pattern is detected, the system requires revivification. The system analyses user credit card data for various characteristics. These characteristics include user country, usual spending procedures. Based upon previous data of that user the system recognizes unusual patterns in the payment procedure. So now the system may require the user to login again or even block the user for more than 3 invalid attempts.

Core Features:

- The system stores previous transaction patterns for each user.
- Based upon the user spending ability and even country, it calculates user's characteristics.
- More than 20 -30 %deviation of users transaction(spending history and operating country) is considered as an invalid attempt and system takes action.

Credit Card Fraud Prediction

ABSTRACT

Along with increasing credit cards and growing trade volume in China, credit card fraud rises sharply. How to enhance the detection and prevention of credit card fraud becomes the focus of risk control of banks. This paper proposes a credit card fraud detection model using outlier detection based on distance sum according to the infrequency and unconventionality of fraud in credit card transaction data, applying outlier mining into credit card fraud detection. Experiments show that this model is feasible and accurate in detecting credit card fraud.

purpose:

Credit card fraud costs consumers and the financial company billions of dollars annually, and fraudsters continuously try to find new rules and tactics to commit illegal actions. Thus, **fraud detection, prediction** systems have become essential for banks and financial institution, to minimize their losses.

However, there is a lack of published literature on credit card fraud detection techniques, due to the unavailable credit card transactions dataset for researchers. The most commonly techniques used fraud detection methods are Naïve Bayes (NB), Support Vector Machines (SVM), K-Nearest Neighbor algorithms (KNN). These techniques can be used alone or in collaboration using ensemble or meta-learning techniques to build classifiers. But amongst all existing method, ensemble learning methods are identified as popular and common method, not because of its quite straightforward implementation, but also due to its exceptional predictive performance on practical problems. In this paper we trained various data mining techniques used in credit card fraud detection and evaluate each methodology based on certain design criteria. After several trial and comparisons; we introduced the bagging classifier based on decision tree, as the best classifier to construct the fraud detection model. The performance evaluation is performed on real life credit card transactions dataset to demonstrate the benefit of the bagging ensemble algorithm.

Credit Card Fraud Prediction

2.Literature survey

Existing system:

existing method to detect the credit card was on the mode like the detection occurs only after the complaint of the card holder about fraud done. It is not a convenient way to avoid the loss happens to the card holder. After getting the complaint they detected the fraud on the basis of the IP address. For this they need the help of the Cyber crime to detect the fraud and make action on it. It takes so much man power.

Disadvantages

1) The main disadvantage of the existing system is the detection occurs only after gets a written complaint.

2) In the existing system there is physical inconvenience exists.

3) The period occurs to detect the fraud will cause so many losses to the cardholder.

4) There is no particular security system in the existing so a hacker can easily access others card.

PROPOSED SYSTEM

Here we are introducing a project for the credit cards fraud detection using Hidden Markov Model (HMM). It is done on the basis of the spending profile of the card holder. The usual spending of the cardholder is being checked by the FDS (Fraud Detection system) in the bank. The system checks all the spending of the user. When it turns unusual the method blocks the transaction on the card. And it alerts the bank. It occurs automatically. It doesn't need any man power.

Advantages

1) The main advantage is that the detection occurs much faster than any other method.

2) In all the existing systems the real card holder should be checked for the fraud detection. But in our method there is no need of the physical inconveniences of the card holder. All the checking and the detection occur automatically.

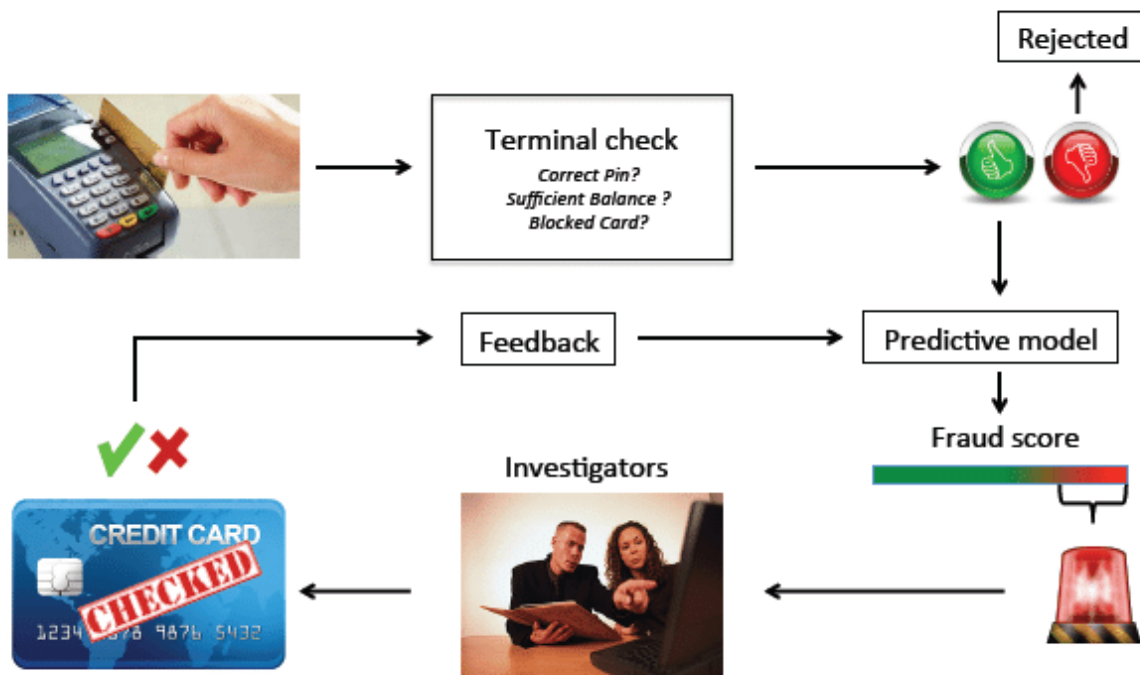
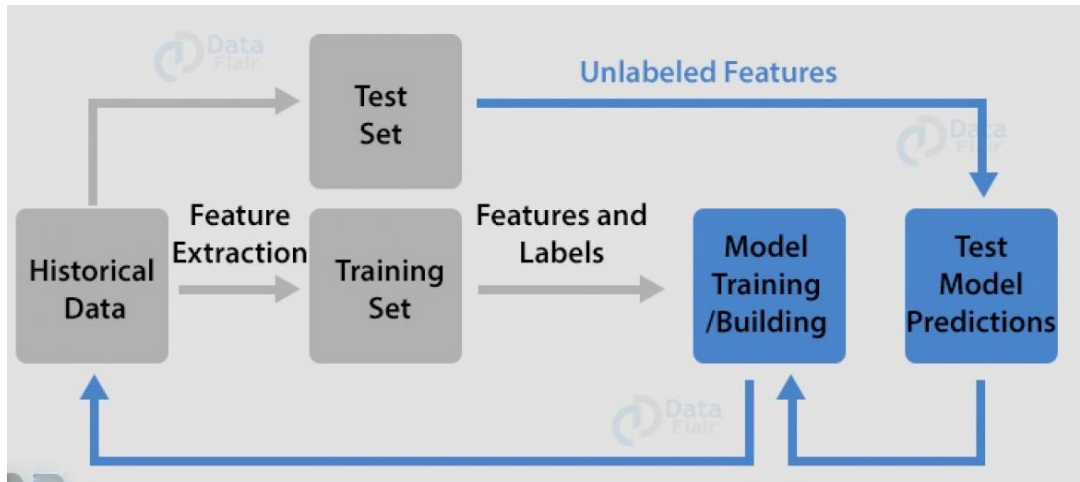
3) This project needs no man power for the detection.

4) This project provides most accurate method in credit card fraud detection.

Credit Card Fraud Prediction

3.Theoretical Analysis

Block Diagram:



Credit Card Fraud Prediction

The fraud detection module will work in the following steps:

- 1)The Incoming set of transactions and amount are treated as credit card transactions.
- 2)The credit card transactions are given to machine learning algorithms as an input.
- 3)The output will result in either fraud or valid transaction by analyzing the data and observing a pattern and using macforest to do anomaly detection.
- 4)The fraud transactions are given to alarm which alerts the user that fraud transaction has occurred and the user can block the card to prevent further financial loss to him as well as the credit card company.
- 5)The valid transactions are treated as genuine transactions.hine learning algorithms such as local outlier factor and isolation.

SYSTEM REQUIREMENTS

Hardware Requirements:

SYSTEM : Pentium IV 2.4 GHz

HARD DISK : 40 GB

FLOPPY DRIVE : 1.44 MB

MONITOR : 15 VGA colour

MOUSE : Logitech.

RAM : 256 MB

KEYBOARD : 110 keys enhanced.

Software Requirements:

OPERATING SYSTEM : Windows 10/7

PLATFORM:Eclipse IDE for java developers,WEKA

CODING LANGUAGE : java

Credit Card Fraud Prediction

Experimental Investigations

During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity.

The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

Design is concerned with identifying software components specifying relationships among components. Specifying software structure and providing blue print for the document phase. Modularity is one of the desirable properties of large systems. It implies that the system is divided into several

Credit Card Fraud Prediction

parts. In such a manner, the interaction between parts is minimal clearly specified. Design will explain software components in detail. This will help the implementation of the system. Moreover, this will guide the further changes in the system to satisfy the future requirements

INPUT DESIGN:

Input design is the process of converting user-originated inputs to a computer-based format. Input design is one of the most expensive phases of the operation of computerized system and is often the major problem of a system.

In the project input forms are login, registration, fraud user login, security details, giving product.

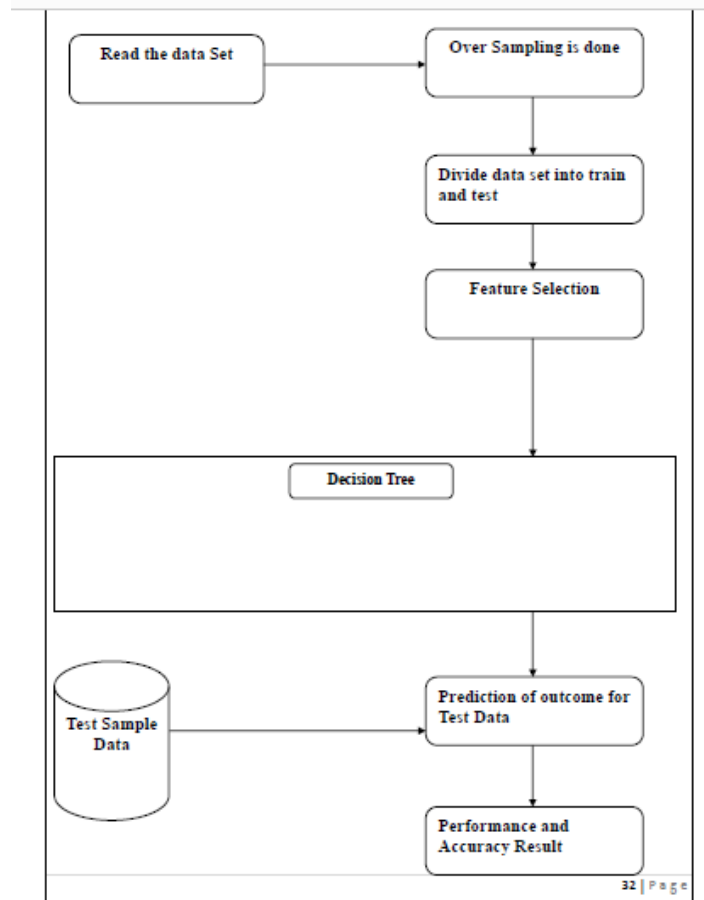
OUTPUT DESIGN:

Output design generally refers to the results and information that are generated by the system for many end-users; output is the main reason for developing the system and the basis on which they evaluate the usefulness of the application.

In the project output forms are view product, viewing purchased product, connection, node information, path node.

Credit Card Fraud Prediction

Flowchart



Credit Card Fraud Prediction

Advantages and Disadvantages:

Advantages:

1. Artificial Immune System (AIS) :Self-Organization/easy in integration with other systems/fault tolerance
2. Hidden Markov Model (HMM): Fast in detection
3. Neural Network: High accuracy/ Portability/ high speed in detection
4. Genetic Algorithm is Inexpensive/fast in detection
5. Decision Tree :High flexibility/easy to Implement
6. Support Vector Machines (SVM) SVMs can be robust, even when the training sample has some bias.

Disadvantages:

1. Need high training time in NSA
2. Low accuracy/not scalable to large size data sets
3. High expense/ Sensitivity to data format.
4. Requires extensive tool knowledge to set up and operate and difficult to understand.
5. Requirements to check each condition one by one. In fraud detection condition is transaction
6. Expensive/Poor in process large dataset

Credit Card Fraud Prediction

Conclusion

Credit card fraud is without a doubt an act of criminal dishonesty. This article has listed out the most common methods of fraud along with their detection methods and reviewed recent findings in this field.

This has also explained in detail, how machine learning can be applied to get better results in fraud detection along with the algorithm, pseudocode, explanation its implementation and experimentation results. While the algorithm does reach over 99.6% accuracy, its precision remains only at 28% when a tenth of the data set is taken into consideration. However, when the entire dataset is fed into the algorithm, the precision rises to 33%. This high percentage of accuracy is to be expected due to the huge imbalance between the number of valid and number of genuine transactions.

Since the entire dataset consists of only two days' transaction records, its only a fraction of data that can be made available if this project were to be used on a commercial scale. Being based on machine learning algorithms, the program will only increase its efficiency over time as more data is put into it.

Credit Card Fraud Prediction

Future Enhancements

While we couldn't reach our goal of 100% accuracy in fraud detection, we did end up creating a system that can, with enough time and data, get very close to that goal. As with any such project, there is some room for improvement here. The very nature of this project allows for multiple algorithms to be integrated together as modules and their results can be combined to increase the accuracy of the final result. This model can further be improved with the addition of more algorithms into it. However, the output of these algorithms needs to be in the same format as the others. Once that condition is satisfied, the modules are easy to add as done in the code. This provides a great degree of modularity and versatility to the project. More room for improvement can be found in the dataset. As demonstrated before, the precision of the algorithms increases when the size of dataset is increased. Hence, more data will surely make the model more accurate in detecting frauds and reduce the number of false positives. However, this requires official support from the banks themselves.

References

1. Datasets. (n.d.). Retrieved from <https://www.kaggle.com/datasets>
2. A. Srivastava, M. Yadav, S. Basu, S. Salunkhe and M. Shabad, "Credit card fraud detection at merchant side using neural networks," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 667-670.
3. W. Yu and N. Wang, "Research on Credit Card Fraud Detection Model Based on Distance Sum," 2009 International Joint Conference on Artificial Intelligence, Hainan Island, 2009, pp. 353-356. doi: 10.1109/JCAI.2009.146
4. Eduonix. (2018, July 26). Eduonix/creditcardML. Retrieved from <https://github.com/eduonix/creditcardML>