

Fraud detection in credit card transaction using hybrid model

Cite as: AIP Conference Proceedings **2277**, 130010 (2020); <https://doi.org/10.1063/5.0025561>
Published Online: 06 November 2020

A. Maria Nancy, G. Senthil Kumar, S. Veena, N. A. S Vinoth, and Moinak Bandyopadhyay



View Online



Export Citation

ARTICLES YOU MAY BE INTERESTED IN

QoS based validation for web services

AIP Conference Proceedings **2277**, 130011 (2020); <https://doi.org/10.1063/5.0025449>

Recommendation of web services using implicit feedback and collaborative filtering technique

AIP Conference Proceedings **2277**, 130008 (2020); <https://doi.org/10.1063/5.0025452>

Effective system for software requirement management

AIP Conference Proceedings **2277**, 240010 (2020); <https://doi.org/10.1063/5.0025463>



Your Qubits. Measured.

Meet the next generation of quantum analyzers

- Readout for up to 64 qubits
- Operation at up to 8.5 GHz, mixer-calibration-free
- Signal optimization with minimal latency

Find out more

 Zurich Instruments

Fraud Detection in Credit Card Transaction using Hybrid Model

A.Maria Nancy ^{a)}, G.senthil Kumar, S.Veena, N.A.S Vinoth, Moinak
Bandyopadhyay

Department of Software Engineering, SRM Institute of Science and Technology, Chennai, India

a)corresponding Author : marianancyrajg@gmail.com,
chenthi2004@gmail.com,

ABSTRACT: Credit card payment is a popular mode of online transaction. It is one of the simplest and easiest mode of payment across the internet. However, with growing popularity of credit card transactions, there is an exponential growth in fraudulent payments. Every year we lose billions of dollars due to fraudulent acts. These activities look like a genuine transaction; hence, simple pattern techniques and less complex methods don't notgo to work to minimize the fraudulent act and minimize the chaos we are proposing a Hybrid model consisting of Deep Learning Algorithm of Convolutional Neural Network followed by K-Nearest Neighbors Classification. These two approaches are proved to decrease the false alarm rates and increase the fraud detection rate and expected to be more efficient than other relevant algorithms.

INTRODUCTION

The popularity of Internet-based transactions is growing day by day, and one of the most sought-after method of payment is by Credit Card. The ease and convenience of use has made the market shift towards credit card as the mode of payment. However, with a growing trend in the development of credit card technology, fraudulent activities have also soared. In worldwide transaction improvement, economic fraud is found to be drastically increasing. Every year huge amount of dollars are lost due to these fraudulent acts. The elegance by which these activities are carried out make them seem like a genuine transaction. An effective and systematized method for the detection of illegal transactions has become a necessity for all banking systems to minimize disorder and bring order in place.

The proposed system works on using Machine Learning techniques like KNN- Classification and Convolutional Neural Network on data-features such as Customer ID (to understand the number of consecutive transactions), Merchant ID (to understand the beneficiary account of the transaction), Merchant type, age and gender of customer, etc.

RELATED WORK

Credit card fraud detection is an application that analyze and detect whether transaction is done genuine or fraudulent act. When an unknown or fraudulent act occurs, it reports directly to the bank and rejects its transaction request.

There has been a lot of research on transaction Fraud Detection on various e-commerce websites. [1] In this paper, they studied the behavior pattern based on their previous transaction records. They classify all the attributes of transaction and construct the logical graph of behavior profile (LGBP). They define the state transition probability matrix to attain the features of transaction and construct behavior pattern for each user and propose a BP based fraud detection method to determine whether the transaction is done by genuine user or not.

[2]Credit card fraud identification using KNN and Outlier detection. They don't require predictive model before classification and their outlier detection mechanism helps to detect the credit card fraud using less memory and computation requirements. Their accuracy depends on the measure of distance (k value). They cannot detect the fraud at the time of transaction.

PROPOSED SYSTEM

Our CCFDS uses a serialized approach for the fraud detection. This means that the model was trained in such a way that, the predicted outcome of the CNN model was feeded into the training set for the KNN. This therefore, increases the accuracy rate as the probability of decreases by cumulative error rate of both primitive models. For this model, the dataset used was congregated by the research of Edgar Alonso Lopez-Rojas and Stefan Axelsson on “BankSim : A Bank Payment Simulation for Fraud Detection Research”. The main factor with real-time fraud detection is, it has to be a continuous learning process with unlabelled data. To solve this constraint, CNN was used as the first layer of the model. It was paired with Long Short TermMemory (LSTM). The combination of CNN along with LSTM is more inclined towards detecting Fraudulent transaction attempts. The LSTM layer allows the model to analyze a sequence of data with the added benefit of memory checking. This allows the model to realise the illegal transactions based on recurring logs of similar transactions. The CNN model thus comprises of a structure as follows:

- Two convolution layers
- One max-pooling layer
- Two convolution layers
- One max-pooling layer
- One LSTM layer
- One dropout layer
- One dense layer

The output of this layer is also stored as the classification label for the training set, to feed into the KNN model that follows. The KNN layer is used to quickly classify through the resultant set, making the model faster and more accurate. The model was trained with a parameter value of $k=5$. This was found to be the ideal value in terms of accuracy. The hybrid model is then ready to be used for predicting fraudulent transaction data. The following diagram (Fig 1) shows the implementation of the hybrid model.

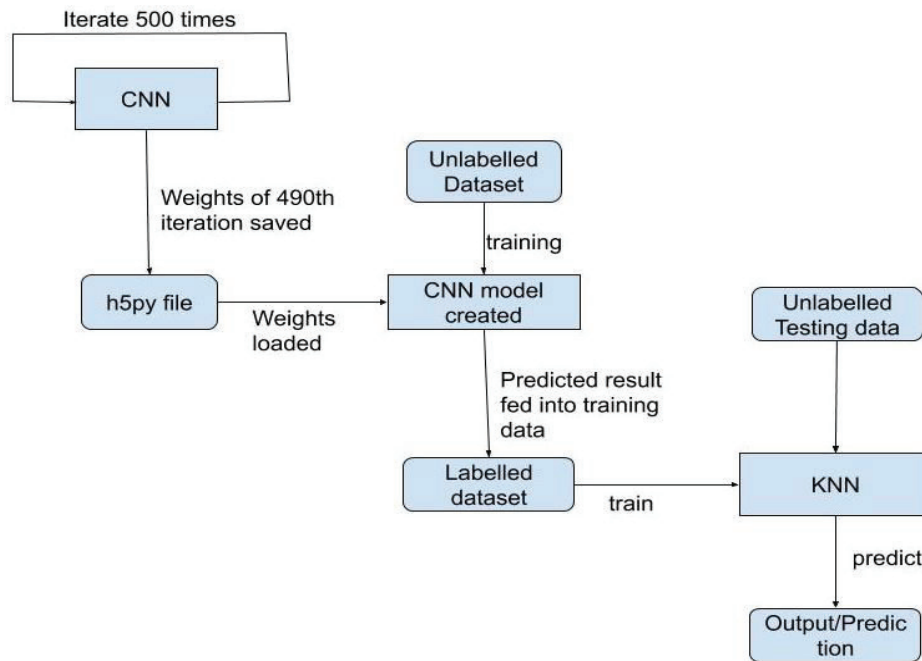


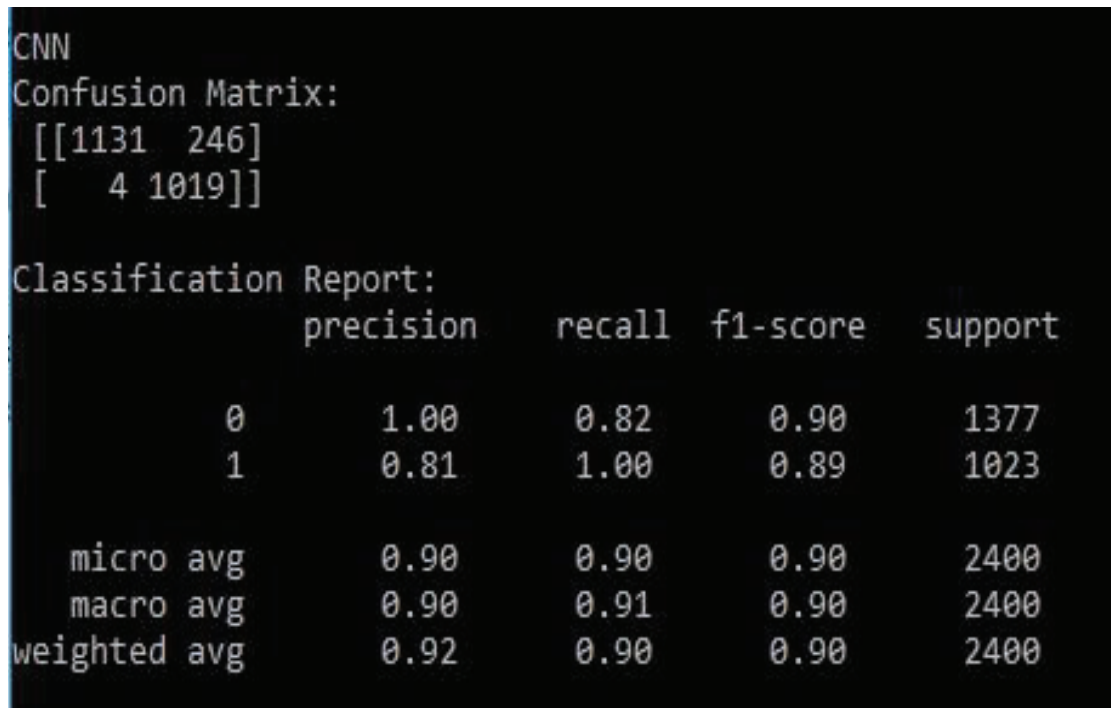
FIGURE-1 Implementation Diagram

Algorithm Hybrid Model

Step1: Import the input data
Step2: Preprocessing the data, removing false data
Step3: Using CNN feature selected
Step4: Using KNN clustering data
Step5: Prediction model

RESULTS

We have used CNN as the first stage of the hybrid model. The individual accuracy of CNN after training the data for 490 iterations has 87.79% and a logarithmic loss of 3.90. It is followed by K-Nearest Neighbor classification, which has a 90.5%. Upon successful hybridization, the resultant CCFDS model has an accuracy of 98% with a logarithmic loss of 0.647.



```
CNN
Confusion Matrix:
[[1131 246]
 [   4 1019]]

Classification Report:
              precision    recall  f1-score   support

      0           1.00       0.82       0.90       1377
      1           0.81       1.00       0.89       1023

   micro avg       0.90       0.90       0.90       2400
   macro avg       0.90       0.91       0.90       2400
weighted avg       0.92       0.90       0.90       2400
```

	precision	recall	f1-score	support
0	1.00	0.82	0.90	1377
1	0.81	1.00	0.89	1023
micro avg	0.90	0.90	0.90	2400
macro avg	0.90	0.91	0.90	2400
weighted avg	0.92	0.90	0.90	2400

FIGURE-2Confusion matrix and classification report for CNN

```

KNN
Confusion Matrix:
[[1204  173]
 [   19 1004]]

Classification Report:
              precision    recall  f1-score   support

     0       0.98       0.87       0.93       1377
     1       0.85       0.98       0.91       1023

 micro avg       0.92       0.92       0.92       2400
 macro avg       0.92       0.93       0.92       2400
weighted avg       0.93       0.92       0.92       2400

```

FIGURE-3Confusion Matrix and classification for CNN

```

Combination
Confusion Matrix:
[[1354   23]
 [   27  996]]

Classification Report:
              precision    recall  f1-score   support

     0       0.98       0.98       0.98       1377
     1       0.98       0.97       0.98       1023

 micro avg       0.98       0.98       0.98       2400
 macro avg       0.98       0.98       0.98       2400
weighted avg       0.98       0.98       0.98       2400

```

FIGURE-4 Confusion Matrix and Classification for Hybrid model (CNN and KNN)

CONCLUSION

We have done a comparative study of the results of K-Nearest Neighbors and Convolutional Neural Network and the hybrid model of both. Among the individual models, The KNN had the highest accuracy rate of 90.66% , followed by CNN with 88.12% accuracy. Upon hybridization, the resultant model had accuracy of 98%.The accuracy of the Convolutional Neural Network increased by 10% when made into a hybrid model with K-Nearest Neighbors, and would only improve if trained over larger balanced dataset.

REFERENCES

1. Luta Zheng, Guan Jun Liu, Chang Jun Jiang, Chungang, "Transaction Fraud Detection Based on Total Order Relation and Behavior Diversity" in IEEE Transactions on Computational Social Systems, 2018.
2. N. Malini, M. Pushpa, "Analysis on credit card fraud identification techniques based on KNN and outlier detection" in Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-informatics (AEEICB), 2017
3. N. Balasupramanian, Ben George Ephrem, Imad Salim Al-Barwani, "User pattern based online fraud detection and prevention using big data analytics and self-organizing maps" in International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), 2017.
4. Krishna Modi, Reshma Dayma, "Review on fraud detection methods in credit card transactions" in International Conference on Intelligent Computing and Control (I2C2), 2017.
5. Jisha Shaji, Dakshata Panchal, "Improved fraud detection in e-commerce transactions" in 2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA), 2017
6. B. B. Sagar, Pratibha Singh, S. Mallika, "Online transaction fraud detection techniques: A review of data mining approaches" in 3rd International Conference on Computing for Sustainable Global Development (INDIA Com), 2016.
7. Dataset from Kaggle link, <https://www.kaggle.com/ntnu-testimon/banksim1>.
8. Edgar Alonso Lopez-Rojas, Stefan Axelsson "BankSim: A Bank Payment Simulation for Fraud Detection Research" in 26th European Modeling and Simulation Symposium, 2016.
9. B. B. Sagar, Pratibha Singh "BLAST-SSAHA Hybridization for Credit Card Fraud Detection" in IEEE Transactions on Dependable and Secure Computing, 2016.
10. S. Benson Edwin Raj ; A. Annie Portia, "Analysis on credit card fraud detection methods" in International Conference on Computer, Communication and Electrical Technology (ICCCET), 2011.
11. Krishna Modi ; Reshma Dayma, "Review on fraud detection methods in credit card transactions" in International Conference on Intelligent Computing and Control (I2C2), 2017.