Paper Presentation

Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework

Authors

Petr Hajek
Mohammad Zoynul Abedin
Uthayasankar Sivarajah

Published on

"Information Systems Frontiers"
A Journal of Research and Innovation
Accepted: 23 September 2022

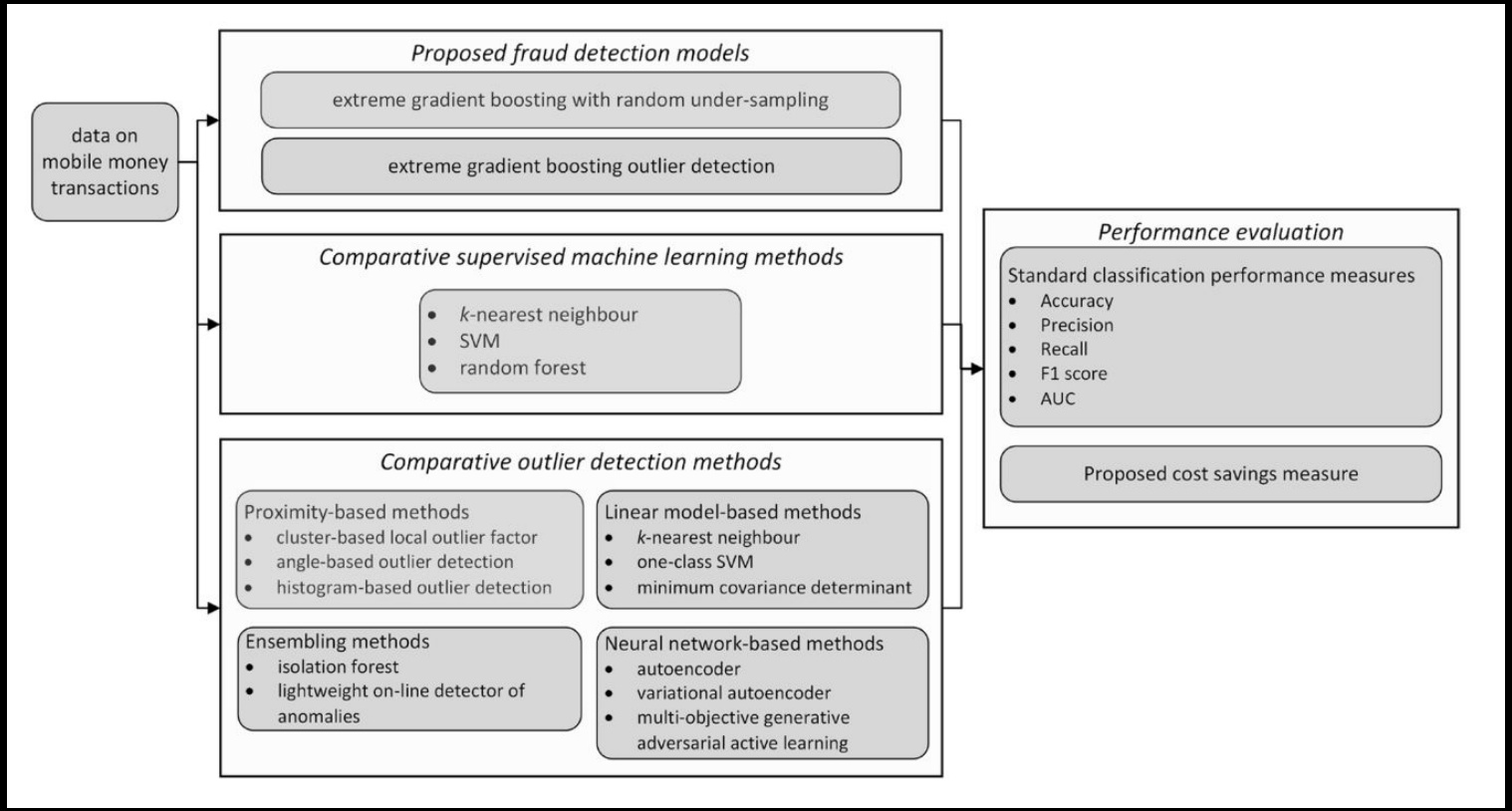
Introduction

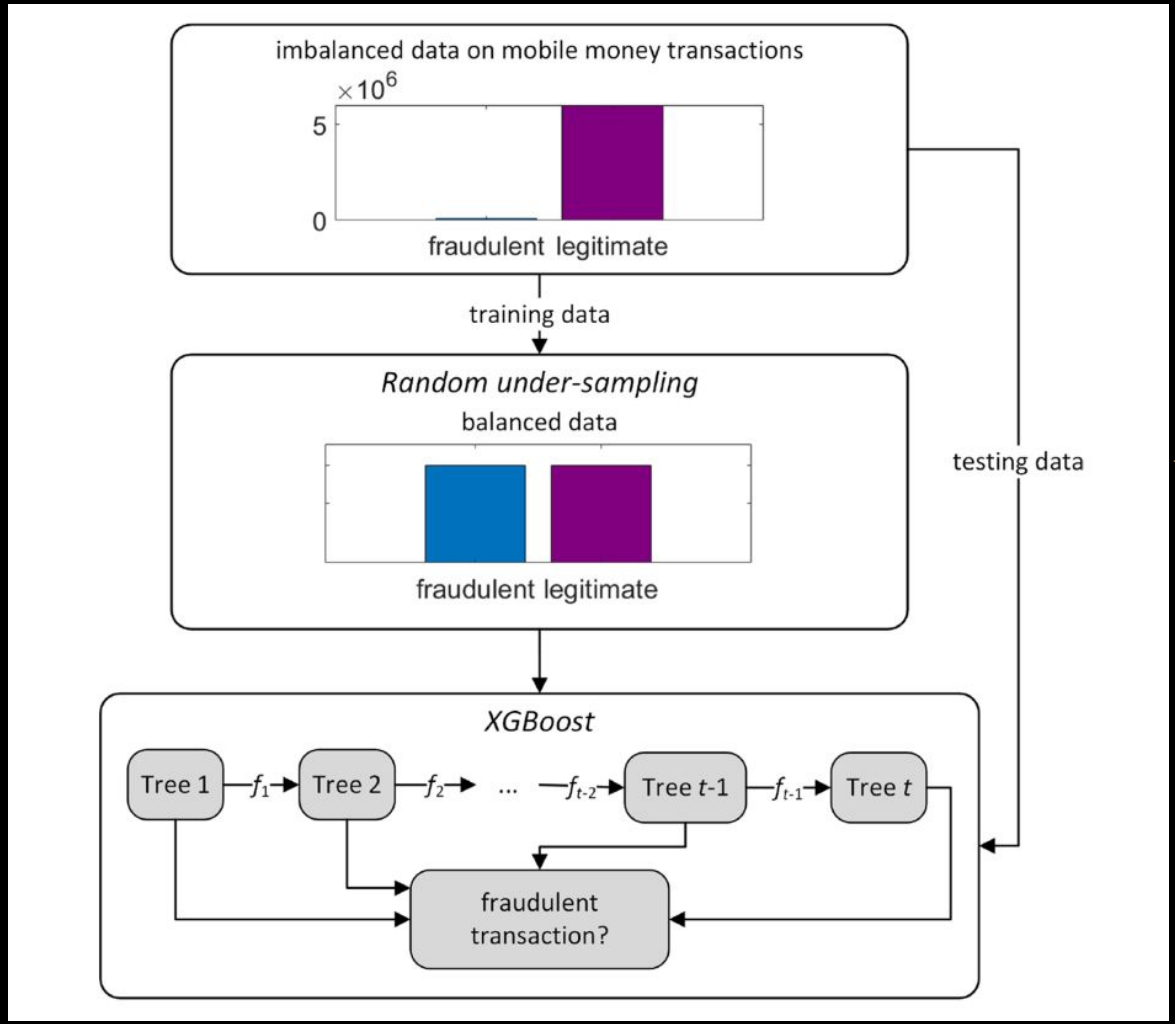
- ➔ **Problem :** Fraud Transaction Detection
- ➔ **Proposed Solution :** XGBoost-based Framework
- ➔ **Dataset :** Simulated Spanish Bank Dataset

Challenges

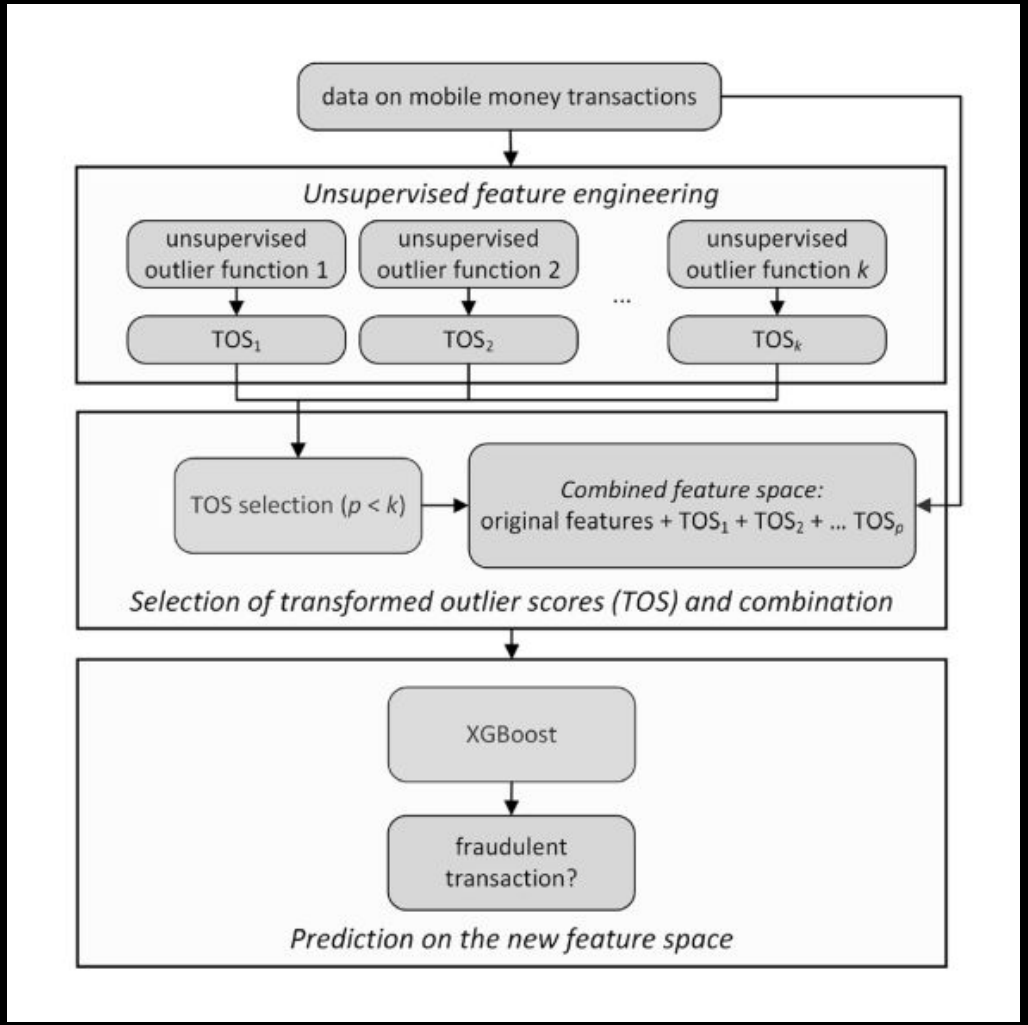
- extreme class imbalance
- changing patterns of fraud over time - *leads to decrease the performance and efficiency of the detection model. Therefore machine learning model must be constantly updated.*
- Inadequate selection of performance metrics - *ratio between two errors(high false positive and low false negative)*

Proposed Framework




imbalanced data on mobile money transactions
×10^6
5
0
fraudulent legitimate
training data
Random under-sampling
balanced data
fraudulent legitimate
testing data
XGBoost
Tree 1
f1
Tree 2
f2
...
ft-2
Tree t-1
ft-1
Tree t
fraudulent transaction?

- **RUS + XGBoost**


data on mobile money transactions
Unsupervised feature engineering
unsupervised outlier function 1
unsupervised outlier function 2
...
unsupervised outlier function k
TOS1
TOS2
TOSk
TOS selection (p < k)
Combined feature space:
original features + TOS1 + TOS2 + ... TOSp
Selection of transformed outlier scores (TOS) and combination
XGBoost
fraudulent transaction?
Prediction on the new feature space

XGBOD

Experimental Setup

- 75% training data, 25% testing data
- 5-fold cross-validation on the training data
- Implementation tool :
  - supervised learning methods in the Python library Scikit-Learn 0.23.0.
  - the RUS algorithm available in the library Imbalanced-Learn 0.6.2.
  - the outlier detection methods available in the library PyOD (Zhao et al., 2009).

Empirical Results

- Performance of Supervised learning
- Random Under sampling effectiveness
- Outlier detection methods performance evaluation
- Financial consequences
- Robustness test using credit card fraud dataset

Conclusion

- Proposed an XGBoost-based fraud detection framework while considering the financial impact of fraud detection.
  - First, the XGBoost model was combined with under-sampling to effectively address the problem of extreme class imbalance and avoid overfitting.
  - Second, to fully exploit the large amount of underlying data, unsupervised outlier detection methods were integrated into the XGBoost-based model

Future Research

- Ensemble methods combined with alternative under-sampling
- And alternative unsupervised outlier detection methods should be further investigate.
- XGBoost method enhanced with weighted and focal losses

Paper Presentation

Credit Card Fraud Detection using Machine Learning and Data Science

Authors

Aditya Saini
Swarna Deep Sarkar
Shadab Ahmed
S P Maniraj

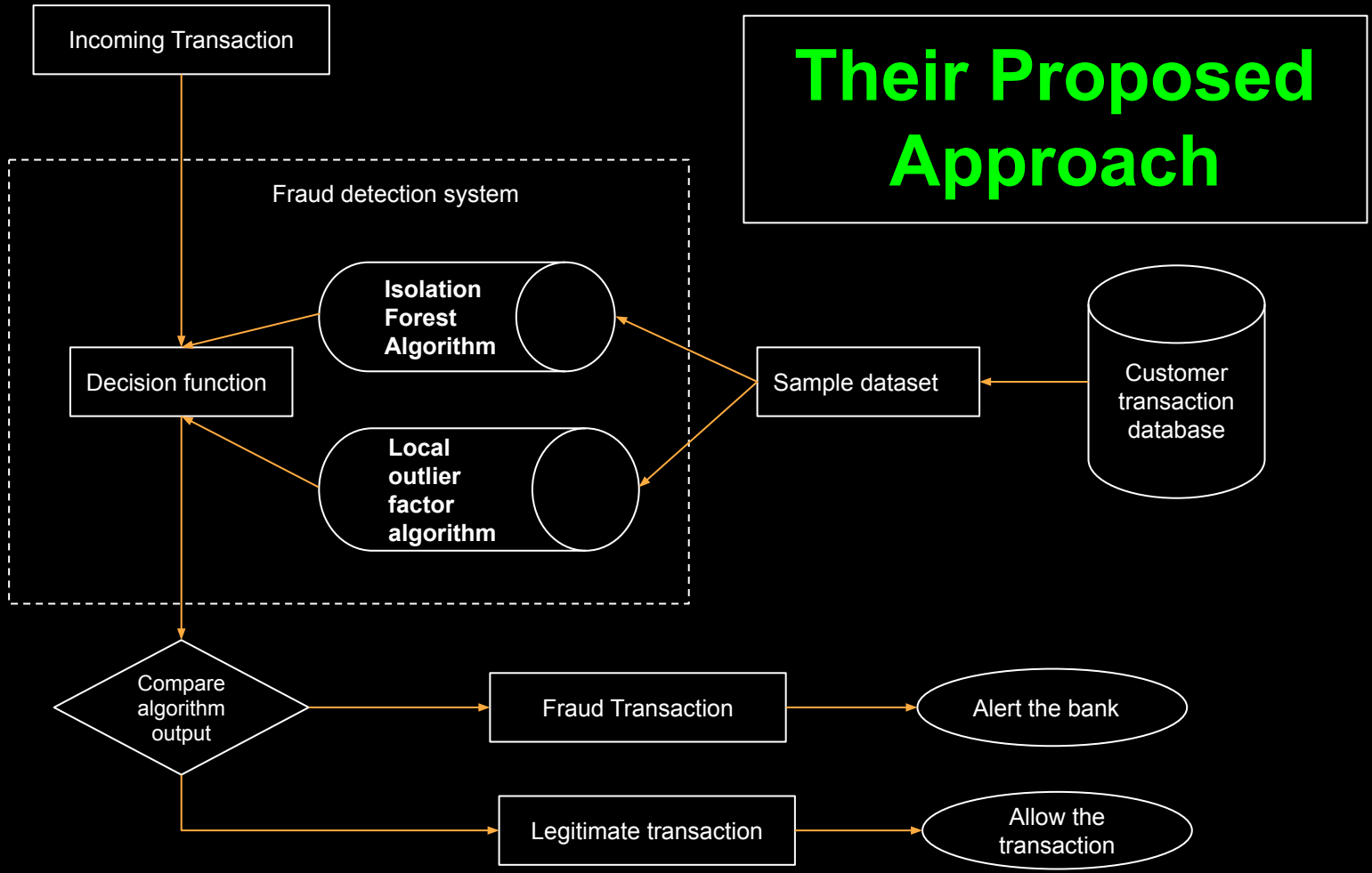
Published on

International Journal of Engineering Research & Technology (IJERT)
Accepted: 09, September-2019

Introduction

- ➔ **Problem :** Credit Card Fraud Detection.
- ➔ **Proposed Solution :** Local Outlier Factor and Isolation Forest Algorithm on the PCA transformed Credit Card Transaction data.
- ➔ **Dataset :** Credit card dataset(Kaggle).

Their Proposed Approach



Problem in my perspective

- They mentioned here that focused on data analysis and preprocessing but “Data Imbalance are not handled.”
- Their technique was applied to a full application data set supplied by a German bank in 2006.Their approach has no proved.
- Evaluation has no proved.
- Not compare with other approach.
- Only experiment 2 methods. Not to try others way or methods and not doing any comparison.
- I think, ultimately I could not find any useful methods or way or insightful information what can i use in my problem.