Geospatial Analysis of Credit Card Fraudulent Transactions

Team 7 · Chanyoung Park, Nithiya Varsha Markandan Rajedren, Samhita Sarikonda, Vishal Orsu, Ravi Datta Rachuri

ABSTRACT

The geospatial analysis of credit card fraudulent transactions identifies how geospatial data can be harnessed through the traditional fraud-detection techniques to point out trends in locations of fraudulent activities of credit cards so that the systems may be working more effectively in detecting such frauds. It leads to great potentiality for fraud, which has resulted in vast consumers and financial institutions facing great challenges due to the pertinent transactions of credit cards that relate to instant financial exchanges in today's financial landscape. Our project proposes integrating geospatial analysis with existing fraud detection methodologies in order to make predictions more precise and generate valuable actionable insights for financial institutions. The initiative draws from earlier researches that have emphasized the necessity for big data technology in mitigating credit card fraud, especially with regard to developing machine learning techniques, which can find and mitigate instances of fraudulent transactions on the move. We used a rich data set with details of individual transactions made using credit cards. This dataset includes anonymized data that has its date, time, merchant information, cardholder demographics, and transaction latitude and longitude. It covers geospatial data, such as the latitude and longitude of transactions, which is key in our analysis. This information will give an in-depth study into the spending behavior of consumers on a geographical scale and be able to identify high-risk areas. The geospatial analysis included in the analytical framework of this project will be used to identify spatial distributions of fraud, including the use of statistical modeling for the identification of critical predictors related to fraudulent transactions in the form of transaction size, timing, and geographical location. To be processed and analyzed effectively through the vast amount, the project applies PySpark technology in distributed data processing, Spark MLlib in the development of a machine learning model, and with all the project management and scalability under the domain of Databricks. It should be mentioned that advanced machine learning algorithms, such as Random Forest and Gradient Boosting Machines, should be used. We choose these algorithms as they are the most powerful in working with big data and solving complex classification problems, corresponding to this project. We will employ the sophisticated analytics techniques described in the preceding approaches in trying to develop a robust model that can better predict cases of potential fraud and adapt to new patterns as they emerge. In sum, "Geospatial Analysis of Credit Card Fraudulent Transactions" is a mammoth step towards the integration of modern analytical techniques with conventional ones, and this should be practiced more to have an efficient fraud detection and prevention regime in today's digitally changing financial environment. Finally, the arms of the financial institutions will be made stronger through this project, giving them robust frameworks of technological fraud detection software to make their operations and customers secure from fraudulent activities with more pinpoint accuracy in tools.

Keywords Geospatial Analysis · Fraudulent Transactions · Machine Learning · Predictive Analytics · Big Data Technologies · Consumer Behavior · Statistical Modeling

References

1. **A review of credit card fraud detection using Machine Learning techniques | IEEE** conference publication | IEEE Xplore. (n.d.). https://ieeexplore.ieee.org/document/9365916
2. Credit Card Transactions Fraud Detection Dataset. (2020, August 5). https://www.kaggle.com/datasets/kartik2112/fraud-detection/code