

Comparing Graph Theoretic Features for Automated Credit Card Fraud Detection

Michael Holtz

October 15, 2023

1 Abstract

In 2022, the Federal Trade Commission received over 440,000 reports of credit card fraud, an increase of over 50,000 over the previous year [1, 2]. The annual cost of credit card fraud in the United States has been estimated at XXXXXXXX. These factors have lead researchers and institutions to develop novel techniques to detect credit card fraud. Using machine learning techniques to identify fraudulent transactions has been an area of research since at least 1994 and has been a hot topic as of late as the number of credit card transactions continues to rise year over year [3, 4]. Specifically, Prusti, Das, and Rath found that the inclusion of three graph features increased the effectiveness of five supervised and two unsupervised machine learning algorithms [5]. In this paper we seek to evaluate many different graph features and feature selection algorithms to judge the best subset of features for each of the seven machine learning algorithms.

2 Introduction

Credit card data is some of the most regulated data in the world. Finding quality datasets, even for academic use, is nearly impossible. In an attempt to solve this, some have turned to simulations such as BankSim[6]. These simulations start with real anonymized transaction data and simulate a market of buyers, sellers, and fraudsters such that the resulting transactions contain the same fraud indicators as the real world data. The resulting dataset can be freely used and shared as it does not contain any private information.

Prusti, Das, and Rath used the Banksim simulation to train several machine learning models, first on features derived directly from the transactions, and secondly incorporating features derived from a graph model of the transactions, namely degree centrality, PageRank, and label propagation algorithm (LPA) community

References

- [1] C. S. Network, *Data Book 2021*. Federal Trade Commission, 2021.
- [2] —, *Data Book 2022*. Federal Trade Commission, 2022.
- [3] Ghosh and Reilly, “Credit card fraud detection with a neural-network,” in *1994 Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, vol. 3, 1994, pp. 621–630.
- [4] F. Reserve, “Federal reserve payments study (frps),” Jul 2023. [Online]. Available: <https://www.federalreserve.gov/paymentsystems/fr-payments-study.htm>
- [5] D. Prusti, D. Das, and S. K. Rath, “Credit card fraud detection technique by applying graph database model.” *Arabian Journal for Science and Engineering*, vol. 46, no. 1-20, 2021.
- [6] E. A. Lopez-Rojas and S. Axelsson, “Banksim: A bank payment simulation for fraud detection research,” in *26th European Modeling and Simulation Symposium, EMSS 2014*, 09 2014.