Sample abstract…

Ciaran Finnegan

D21124026 – TU060 – First Year MSc Computer Science (Data Science)

***Title***

An analysis to determine if in the past decade (2012-2022) that the evolution of Machine Learning techniques for Credit Card Fraud detection have improved in accuracy and performance.

Maybe change ‘evolution’ to ‘application’…

Credit Card Fraud – what specific context – geographical, etc.

What is credit card fraud? Define it…

Machine Learning techniques – include the datasets in the abstract…and how these datasets will have changed in terms of range of features…

Add sample accuracy levels…

Include reference to unbalancing and resampling – this is in there but re-phrase.

Useful at start of report - Add key words based on [Computing Classification System (acm.org)](https://dl.acm.org/ccs)

***Abstract***

Credit Card fraud remains a multi-billion euro challenge each year for Financial Institutions and their customers. Loses grow annually, and the patterns of fraud execution continue to adapt to new payment channels. This review describes Machine Learning techniques that were frequently employed from 2012 – 2017 and compares them against emerging approaches documented circa 2020. The objective is to determine if the more modern ML strategies are delivering significantly better performance, despite possible limitations because of their inherent complexities.

As of 2021, there are still relatively few historical credit card fraud datasets upon which to conduct Machine Learning experiments for fraud detection. A high proportion of these datasets are also highly imbalanced with less that 1% of records reflecting incidents of fraud. Therefore, the drive in this area of research is to produce Machine Learning models that will work with these challenges and offer the best detection performance.

This article looks at five major studies, conducted from 2012-2020, that applied a range of feature engineering and algorithm selections techniques. The first two papers employed what the authors themselves described as ‘traditional’ ML Classification approaches. The third paper looks at an ensemble approach to resampling and anomaly detection. The fourth paper moves into the more contemporary approach of using Neural Networks. The last paper looks at recent algorithm optimisations to avoid resampling of imbalanced data and circumvent possible data corruption.

All the research referenced in this article produced metrics to advance the benefits of their techniques.

It is not possible to compare the findings directly as the authors use a range of criteria from AUC to F1 Scores. However, the later papers make some compelling arguments that Neural Networks, more often used for image classification, can deliver solutions that are both more accurate and, crucially, much faster at detection credit card fraud. This has implications in terms of real-time detection applications.

Neural Network solutions can be criticised for lack of auditability in terms of classifications, and that remains a problem in this domain, but the potential is obvious.