Fraud detection Project

Introduction

ATMs, introduced in the late 1960s, have revolutionized banking by providing convenience for transactions. However, this has also led to a rise in fraudulent activities during ATM transactions. These include skimming and card trapping. To combat these threats, banks have implemented secure authentication methods, chip, and PIN cards. Criminals adapt with new tactics, such as malware attacks and hidden cameras.

Project Objectives

In this project, our primary objectives can be summarized as follows:

1. **Studying the State of the Art**: We will begin by conducting an extensive review of existing literature and methodologies related to ATM (Automated Teller Machine) management using machine learning. This step aims to provide us with valuable insights into the various models and approaches employed in similar contexts.
2. **Developing a Machine Learning Solution**: Building upon the knowledge gathered from the literature review, our goal is to design and create a machine learning solution tailored specifically for ATM management. This solution will harness advanced algorithms to optimize cash management processes and improve overall operational efficiency.
3. **Constructing a Universal ATM Model**: We aspire to develop a versatile machine learning model that can be applied to a wide range of ATM setups, although there may be certain limitations based on practical considerations. This model will be designed to handle diverse transaction patterns and ATM usage scenarios.
4. **Hyperparameter Optimization and Data Preprocessing**: We will fine-tune our machine learning model through rigorous hyperparameter optimization and advanced data preprocessing techniques. This step is aimed at enhancing the model's accuracy, resilience, and predictive capabilities.
5. **API Development with ASP.NET in C#**: To seamlessly integrate our machine learning solution into real-world banking operations, we will create an application programming interface (API) using ASP.NET in the C# programming language. This API will act as a communication bridge between the machine learning model and the ATM systems.
6. **Simulation and Performance Comparison**: We will simulate the operation of our machine learning solution in various scenarios and compare its performance with traditional, non-automated methods. This comparative analysis will provide valuable insights into the tangible benefits of our approach.
7. **Financial Impact Reporting**: A crucial aspect of this project involves assessing the financial impact of implementing our machine learning solution within the banking environment. We will conduct a thorough evaluation of the gains achieved in terms of cost savings, increased operational efficiency, and enhanced customer satisfaction.

Project Guide

As previously mentioned, our project has two primary objectives. First, we aim to implement two models: one for classifying incoming transactions as fraudulent or legitimate, and the other for

identifying the specific reason behind fraud using multi-class classification. The project is divided into two distinct phases: the training phase and the testing phase.

Training Phase:

The training phase involves the development and training of the machine learning models. This phase consists of three code files. The "fraude cases.ipynb" file calculates additional features using data provided by the company and saves this new data in a file called "fraudecases.csv." This dataset will serve as the basis for training our models. Additionally, there are two more code files in this phase. The first, "unsupervised learning.ipynb," applies unsupervised learning techniques to detect anomalies or fraud cases within our dataset. The second file, "supervised learning.ipynb," is dedicated to developing the two models mentioned earlier. It first utilizes "scaled_df.csv," which is the previously calculated data, but scaled, and applies various models for binary and multi-output classification.

Testing Phase:

The testing phase focuses on the deployment of our solution. It involves the creation of a web application that directly imports data from a SQL server, calculates additional features, imports pre-trained models for prediction, and presents the results as a RESTful API. The results obtained from this phase classify incoming transactions as either fraudulent or not and provide insights into the reasons for fraud.

Conclusion

In summary, our project encompasses the development and training of machine learning models, followed by the deployment of these models through a web application to classify transactions and identify fraud reasons efficiently.