Fraudulent Transaction Detection

Presented by: Tamanna Grover
Taarini Dhulipala
Suryansh Priyadarshi

Overview

Problem Statement

Develop a machine learning model to detect fraudulent financial transactions by analyzing historical transaction data. The model should accurately classify transactions in real-time, minimizing false positives and maximizing detection accuracy. It should adapt to evolving fraud patterns and handle imbalanced datasets effectively.

Abstract

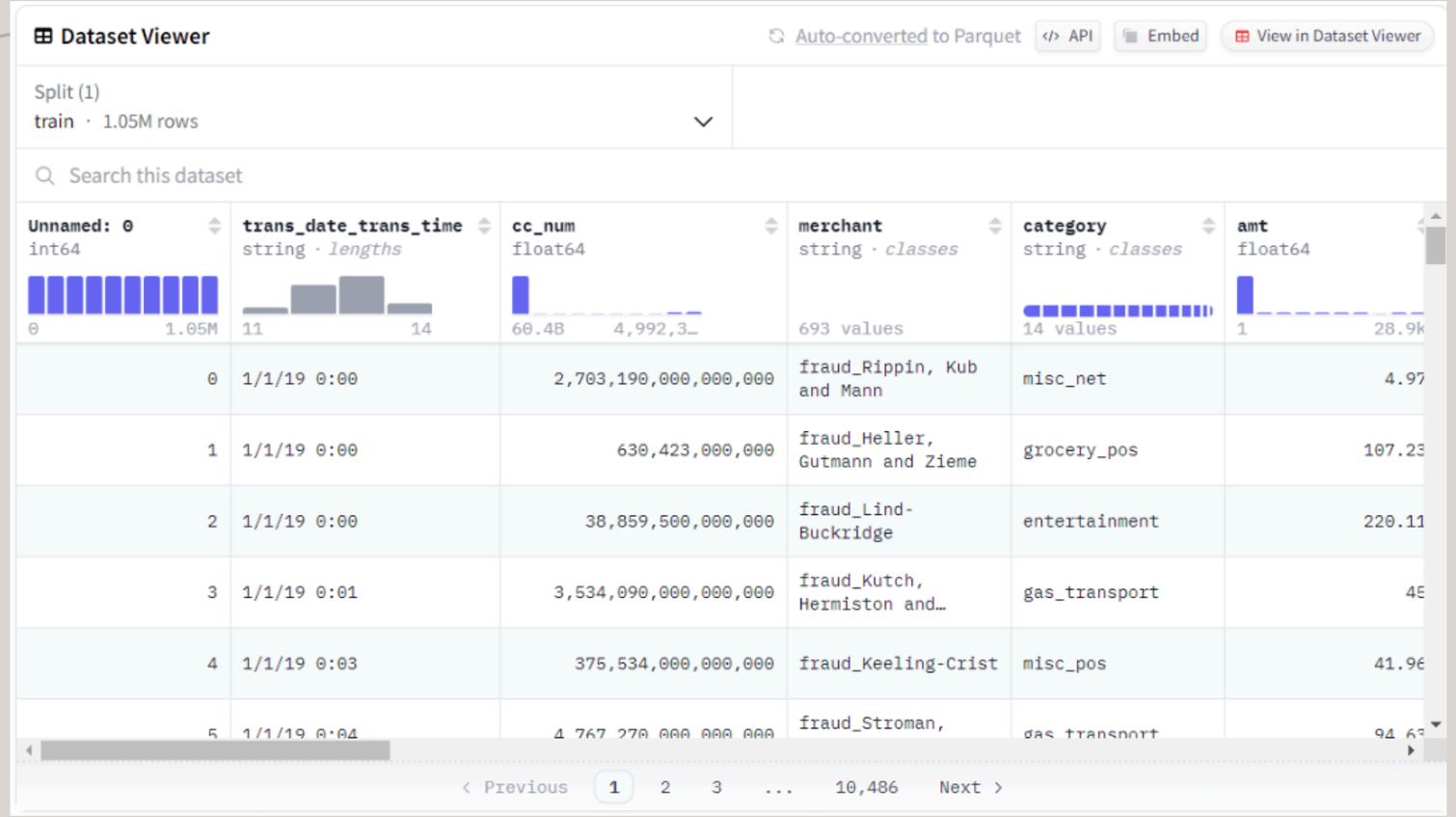
This project aims to detect fraudulent financial transactions using quantum machine learning. We will use a dataset from Hugging Face, applying preprocessing and splitting it into training and test sets. Quantum Approximate Optimization Algorithm (QAOA) will be used for feature selection to reduce computational complexity. Instead of a neural network, we will implement a Quantum Support Vector Machine (QSVM), which utilizes quantum kernels for more accurate classification. The model will be trained on selected features, and its performance will be evaluated and fine-tuned to minimize false positives and maximize accuracy.

Literature Survey

1.Title: A Survey of Machine Learning Techniques for Fraud Detection(2009)
Authors: M. Chandola, A. Banerjee, V. Kumar
Summary: This paper reviews various machine learning techniques used for fraud detection, including supervised, unsupervised, and semi-supervised methods. It addresses challenges like imbalanced datasets and discusses feature engineering strategies and practical applications.

2.Title: Quantum Machine Learning: A Survey and Current Directions(20140
Authors: S. Lloyd, P. Mohseni, C. Preskill
Summary: This survey explores quantum machine learning, covering quantum algorithms and their potential benefits over classical methods. It highlights quantum models like Quantum Support Vector Machines and discusses future directions in the field.

Dataset Used



Dataset Viewer

Auto-converted to Parquet | API | Embed | View in Dataset Viewer

Split (1)
train · 1.05M rows

Search this dataset

Unnamed: 0 int64 0 – 1.05M	trans_date_trans_time string · lengths 11 – 14	cc_num float64 60.4B – 4,992,3…	merchant string · classes 693 values	category string · classes 14 values	amt float64 1 – 28.9k
0	1/1/19 0:00	2,703,190,000,000,000	fraud_Rippin, Kub and Mann	misc_net	4.97
1	1/1/19 0:00	630,423,000,000	fraud_Heller, Gutmann and Zieme	grocery_pos	107.23
2	1/1/19 0:00	38,859,500,000,000	fraud_Lind-Buckridge	entertainment	220.11
3	1/1/19 0:01	3,534,090,000,000,000	fraud_Kutch, Hermiston and…	gas_transport	45
4	1/1/19 0:03	375,534,000,000,000	fraud_Keeling-Crist	misc_pos	41.96
5	1/1/19 0:04	4,767,270,000,000,000	fraud_Stroman,	gas_transport	94.63

‹ Previous 1 2 3 ... 10,486 Next ›

- dazzle-nu/CIS435-CreditCardFraudDetection
- This dataset will be preprocessed to split the dataset into train and test data used for training and testing the data respectively.

Feature Selection

QAOA for feature selection in fraud detection models optimizes the selection of important features by framing it as a combinatorial optimization problem. It leverages quantum computing to efficiently explore feature subsets, improving model accuracy while handling large datasets faster than traditional methods. QAOA helps reduce the computational complexity by focusing on the most critical features for fraud detection.

Quantum Support Vector Machine

The quantum kernel function is essential in QSVM. It computes the inner product of data points in the high-dimensional quantum space, providing a powerful way to evaluate the similarities between data points. This capability enhances the model's ability to discriminate between legitimate and fraudulent transactions by making the separation boundaries more distinct.

Training with Data

During training, QSVM uses the features selected by QAOA. The quantum processor is leveraged to perform computations that are infeasible for classical systems, particularly for large or complex datasets. The training process involves optimizing the model parameters to find the best separating hyperplane between the two classes: legitimate and fraudulent transactions.

Evaluation and Opitmisation

After training, the QSVM model is evaluated using metrics such as accuracy, precision, recall, and F1 score. The quantum kernel parameters and other hyperparameters are tuned to enhance the model's performance. The goal is to maximize the detection of fraudulent transactions while minimizing false positives, ensuring the model is both accurate and practical for real-world applications.

Thank You