Financial Fraud Detection for Ibotta

Identify fraudulent activity in bank account transactions to exclude fraudsters from its platform to make new programs more effective

Objectives

Ibotta's mission is to make all purchases rewarding via digital coupons. Unfortunately, fraudsters like to take advantage of our platform. We would like a model that:

(1) Takes as input one more example accounts
and
(2) Outputs one or more accounts that are very similar to the supplied accounts

This would allow our fraud investigators to quickly identify the same fraudulent activity in real-time and allow them to take action.

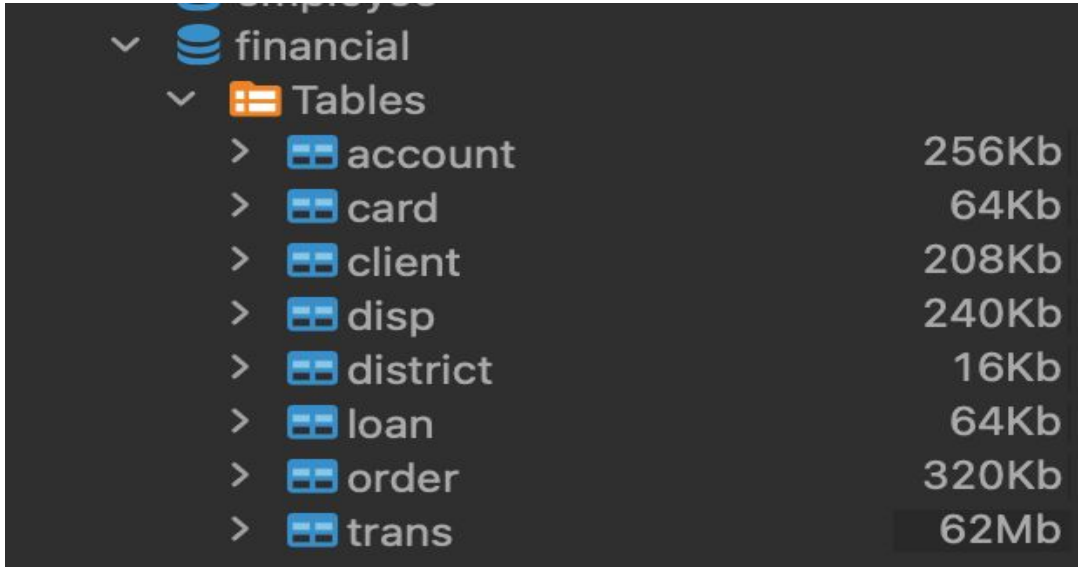
In this project the goal is: 1) To process customer account data and detect similarity to other accounts/transaction combinations in a scalable manner, 2) To process large volumes of data with minimal latency and high accuracy for fraud detection/outlier detection tasks.

Questions to answer

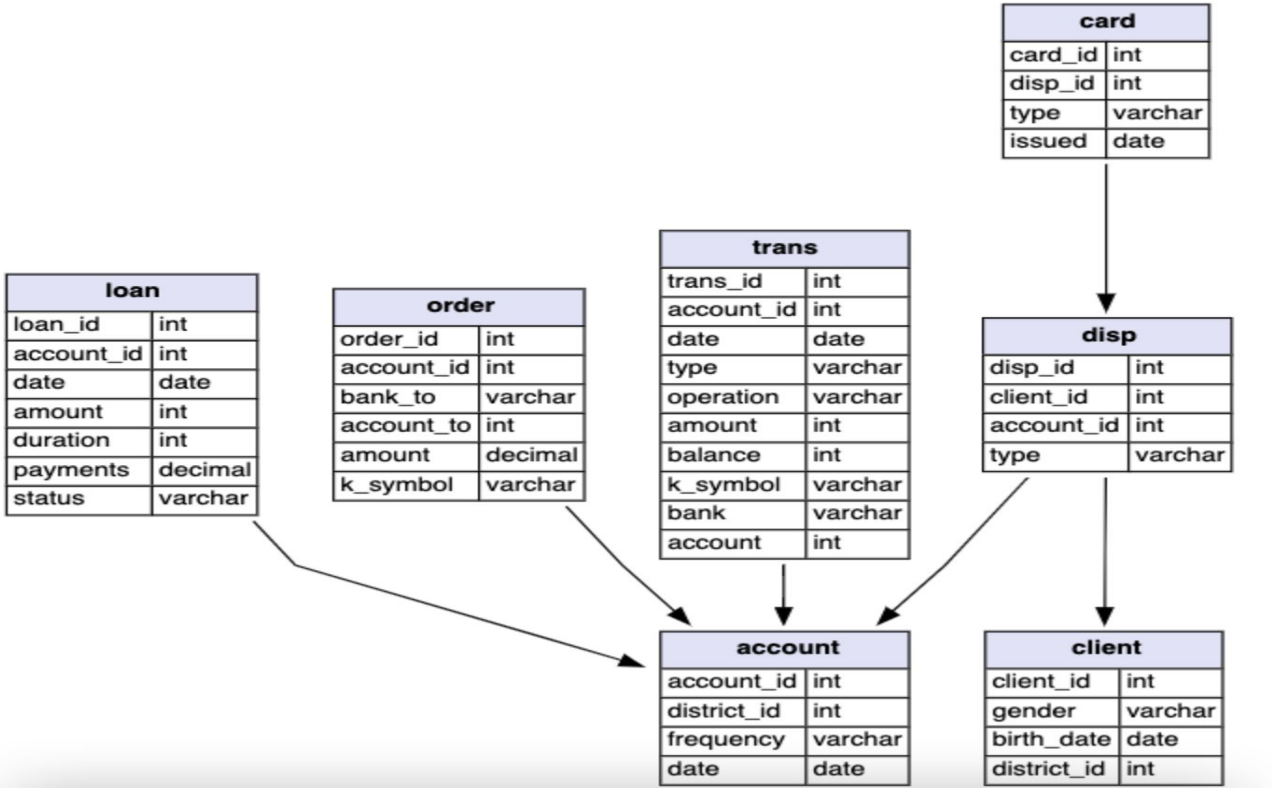
1. End-to-end pipeline that identifies fraudulent accounts.
2. Quantitative and qualitative ways to identify fraudulent accounts and similar accounts.
3. Deployed ML model
4. Technical report on System setup and performances.

Dataset

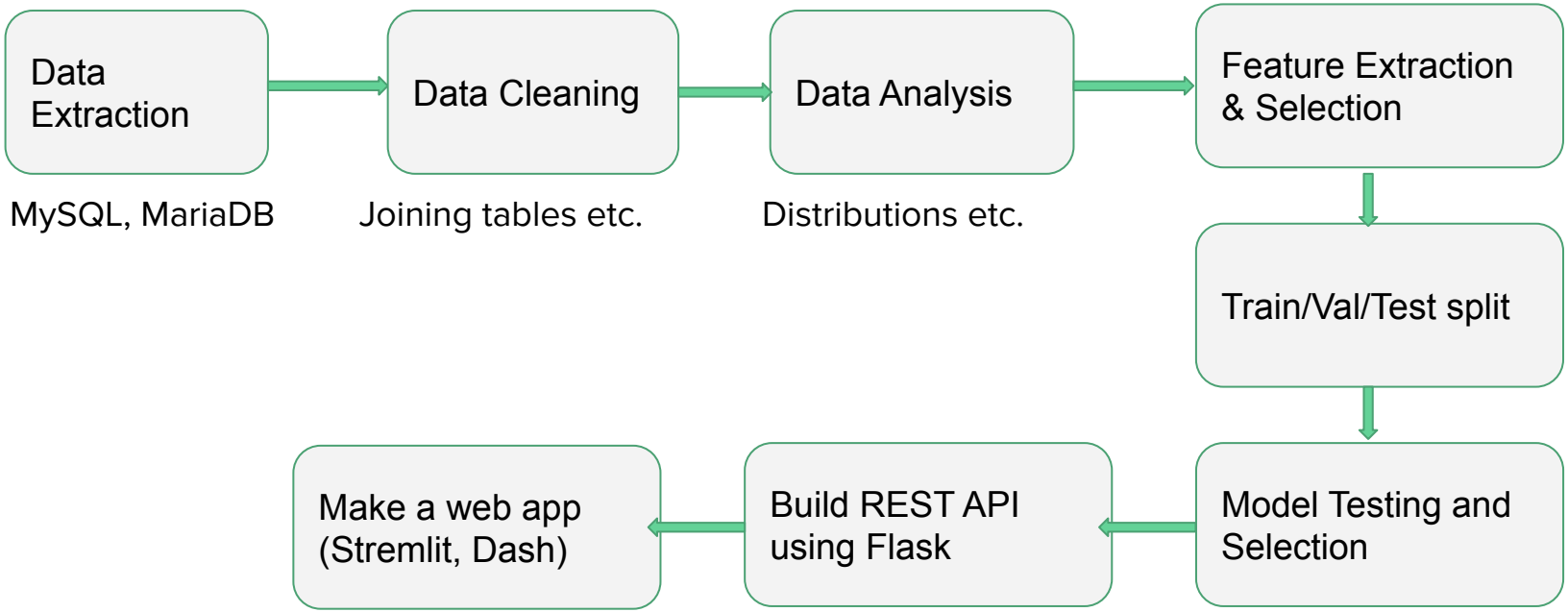
https://relational.fit.cvut.cz/dataset/Financial



Dataset (Tables and Relations)



Model Pipeline



Development Timeline

Task	Start Date	End Date	Duration
Data Extraction			0.5 weeks
Data Cleaning			0.5 week
Data Analysis			0.5 week
Feature Extraction and Selection			1 weeks
Train/Test/Val Split			0.5 week
Model Testing and Selection			0.5 week
RestAPI			1 weeks
Web App			1 weeks