

Ibotta - Financial Fraud Detection

Tiziano Franza
Unmesh Mali

What are we building? What is the product?

- ML model to detect anomalies within the transactions database.
- Single transactions + sequences of transactions.
- A web app with easy-to-use interface that summarizes patterns in the transactions database on a bank.

Who/Why is this product for?

- For **Fraud Investigations** team at Ibotta.
- Investigate each account individually & find similar accounts with the bank
- E.g. Fraud investigators know an account to have fraudulent activity. Our product return s accounts that have higher likelihood of containing fraudulent activities as well.

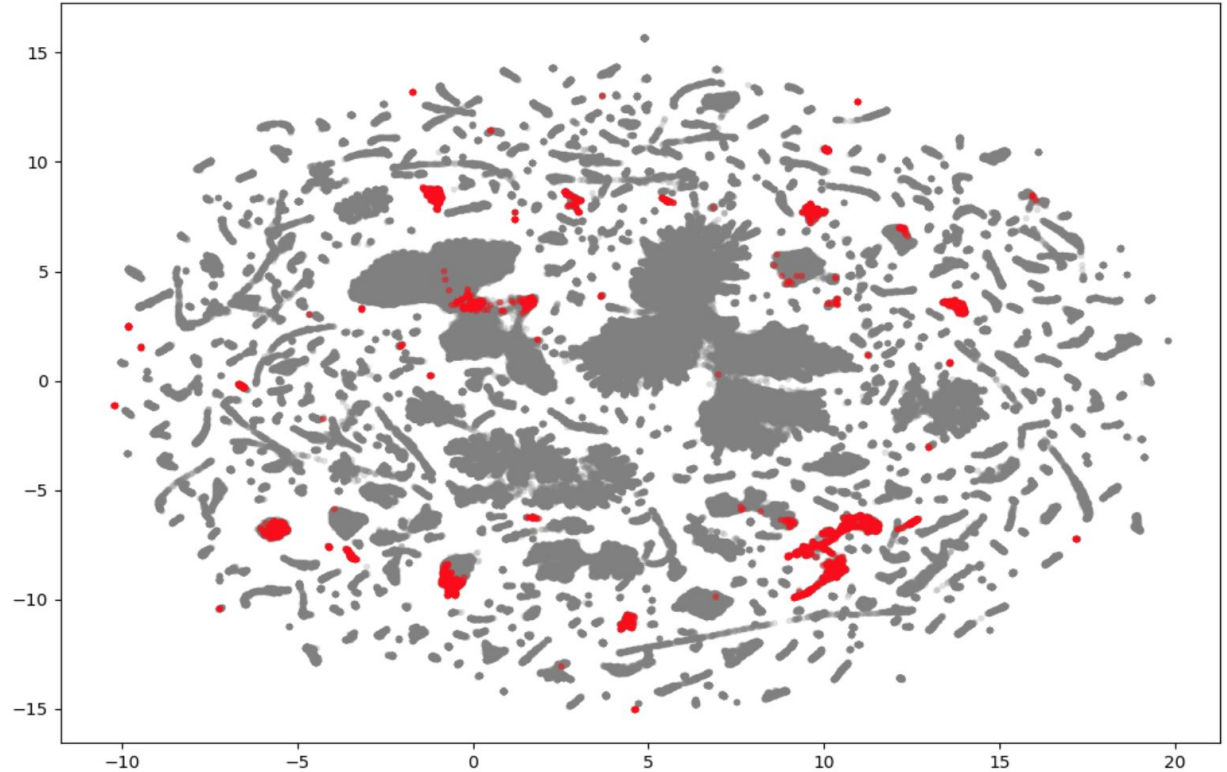
Feature Engineering

Raw Dataset	(1005200 x 7)	Date, categorical columns, etc
Feature Extraction	(1005200 x 25)	Splitting date column, one-hot encoding
Feature Selection	(1005200 x 21(k))	Removing highly correlated features. (v2 - sequence of length k)
Feature Scaling	(1005200 x 21)	Standard Scaler (z-score normalization)

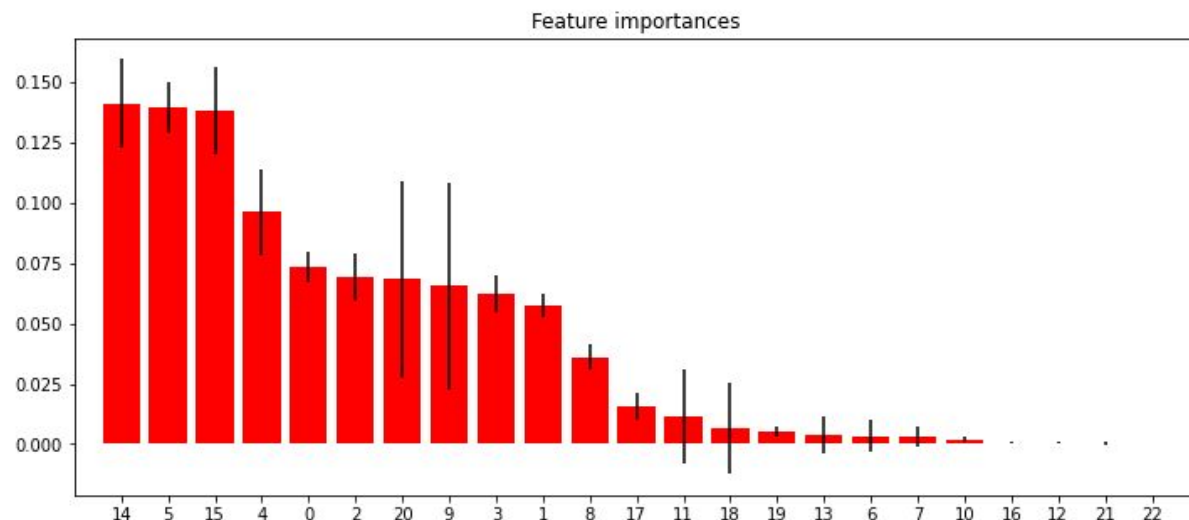
Model Selection

Considerd models:

- KNN
- Local Outlier Factor
- PCA
- HDBSCAN
- Isolation Forests
- Extended Isolation Forests
- Autoencoders



Outlier Analysis



Feature ranking:

1. amount (0.141260)
2. date_days_from_last_trans (0.139548)
3. balance (0.138207)
4. date_year (0.096114)
5. date_dayofweek (0.073401)
6. date_week (0.069108)

Development Timeline

Task	Start Date	End Date	Duration
Data Extraction	Jul, 26	Jul, 28	0.5 weeks
Data Cleaning	Jul, 29	Aug, 1	0.5 week
Data Analysis & Feature Extraction	Aug, 2	Aug, 4	0.5 week
Feature Selection	Aug, 5	Aug, 12	1 weeks
Model Testing and Selection	Aug, 13	Aug, 18	0.5 week
RestAPI	Aug, 19	Aug, 26	1 weeks
Web App	Aug, 26	Sep, 2	1 weeks

 - Completed  - Subject to change with model tuning