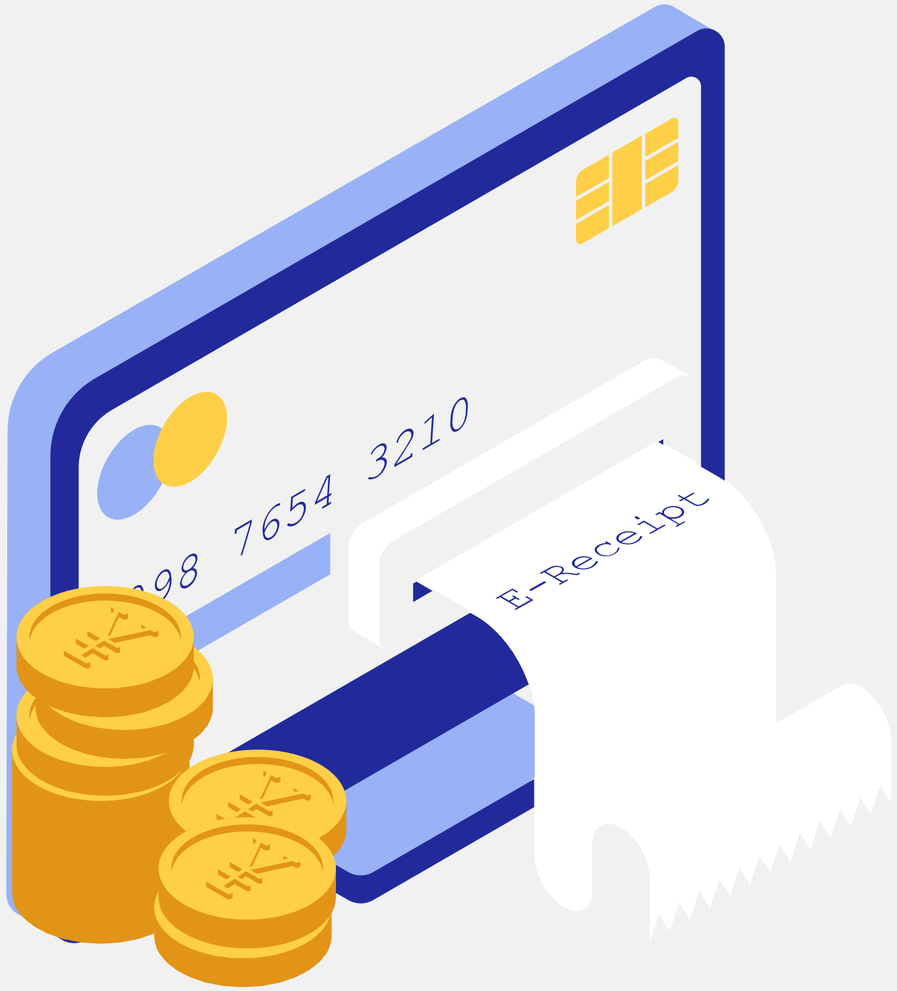
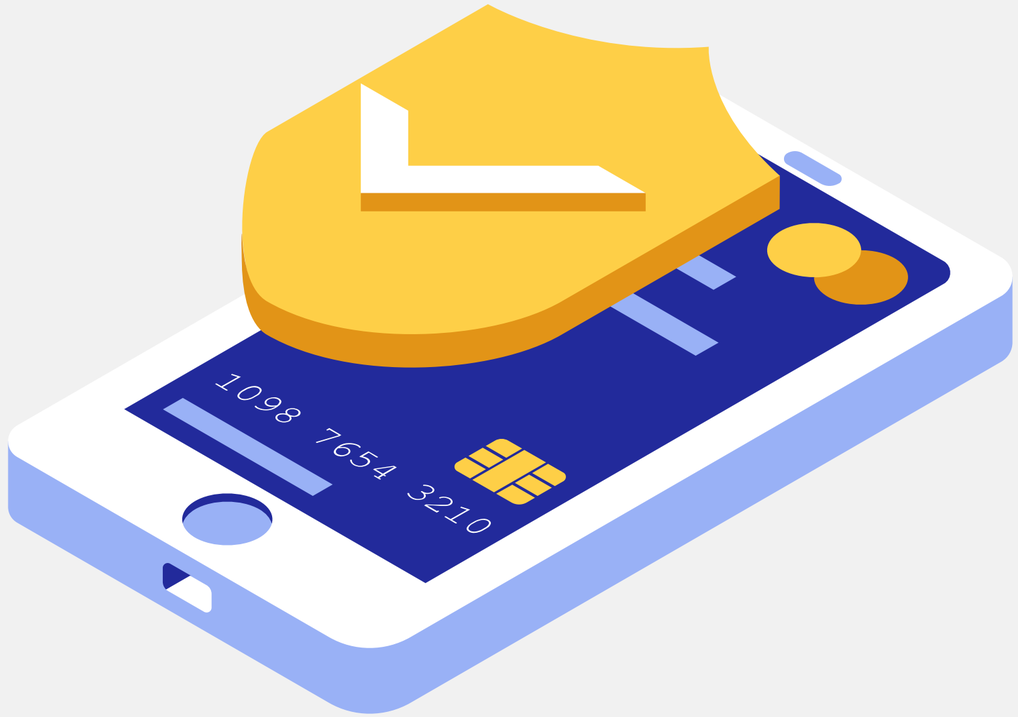


Skyline Financial Services (SFS)

Fraud Investigation

Objective

Analyse **user data**, including **demographics, spending habits**, and **transaction history**, to help SFS **better detect and prevent** future fraud, **reducing revenue loss**



Agenda

01

Data Analysis Findings

Who are more likely to commit fraud?

02

Fraud Detection Model

ML model to detect and prevent fraud

03

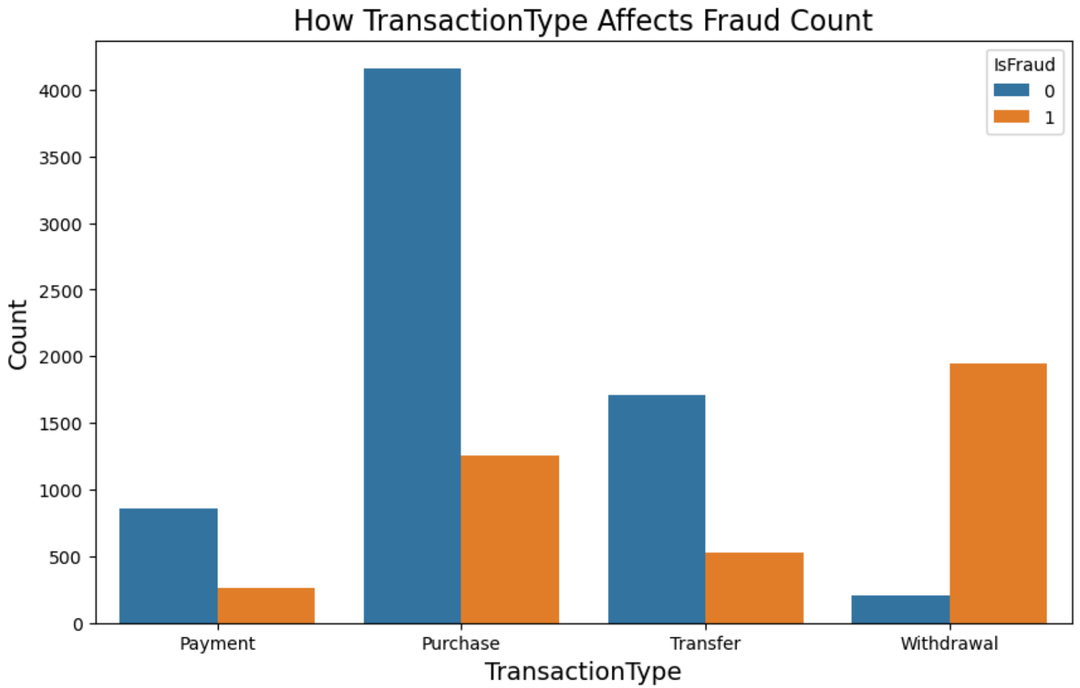
Conclusions & Recommendations

01
Data Analysis Findings

What type of transaction?

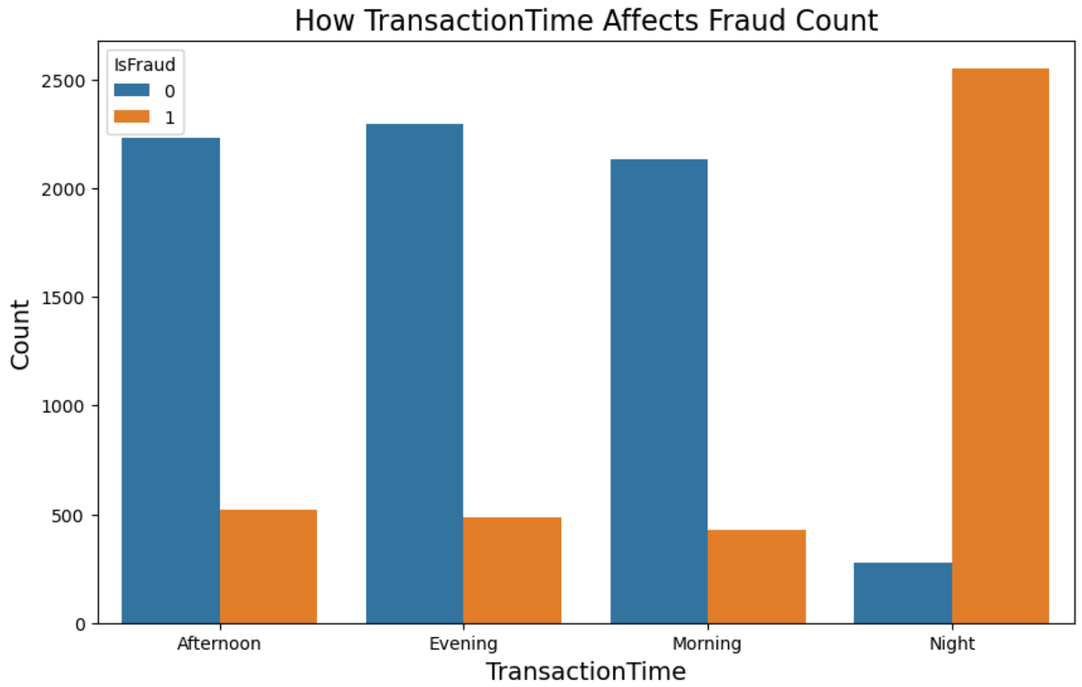
Transaction Type

Most of fraudulent activities are through **withdrawal transactions**



Transaction Time

Most of fraudulent activities happens at **night time (12 am to 6 am)**

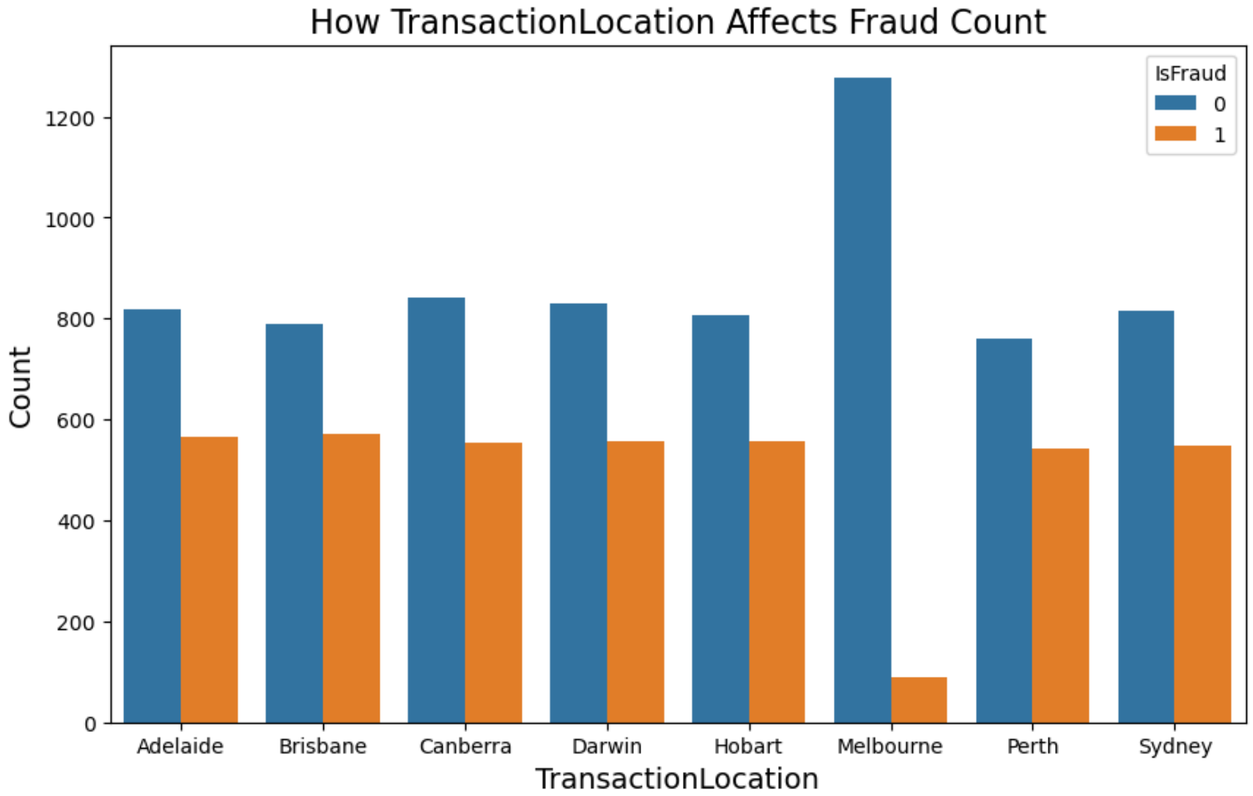


What type of transaction?

Transaction Location

Melbourne has exceptionally low fraudulent transactions

Further investigations needed to understand why



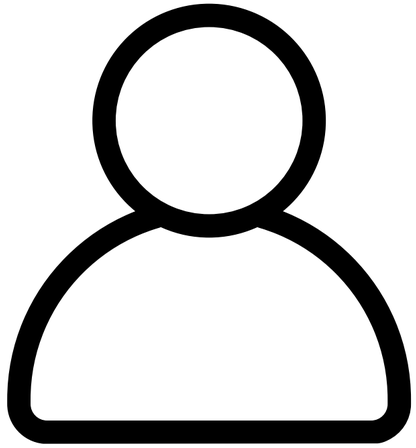
Who is more likely to commit fraud?

Use Yahoo

Divorced or Widowed

Elderly people

Hold a PhD



Student or Unemployed

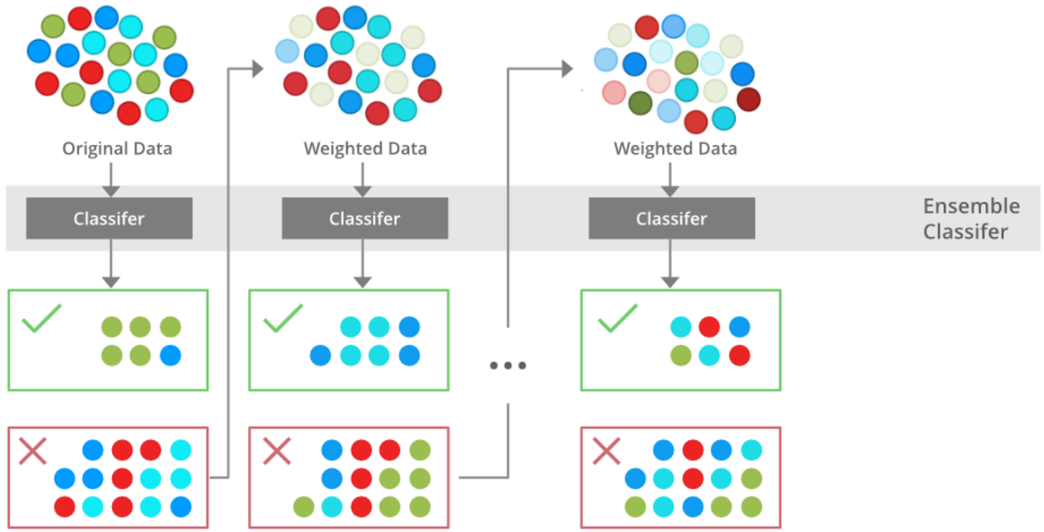
High-Risk Users

UserID	Count
2479	8
2616	7
4052	7
1077	7
1951	7
495	7
491	7

02
Fraud Detection Model

XGBoost Algorithm

An extremely effective model for tabular data



Why we use it?

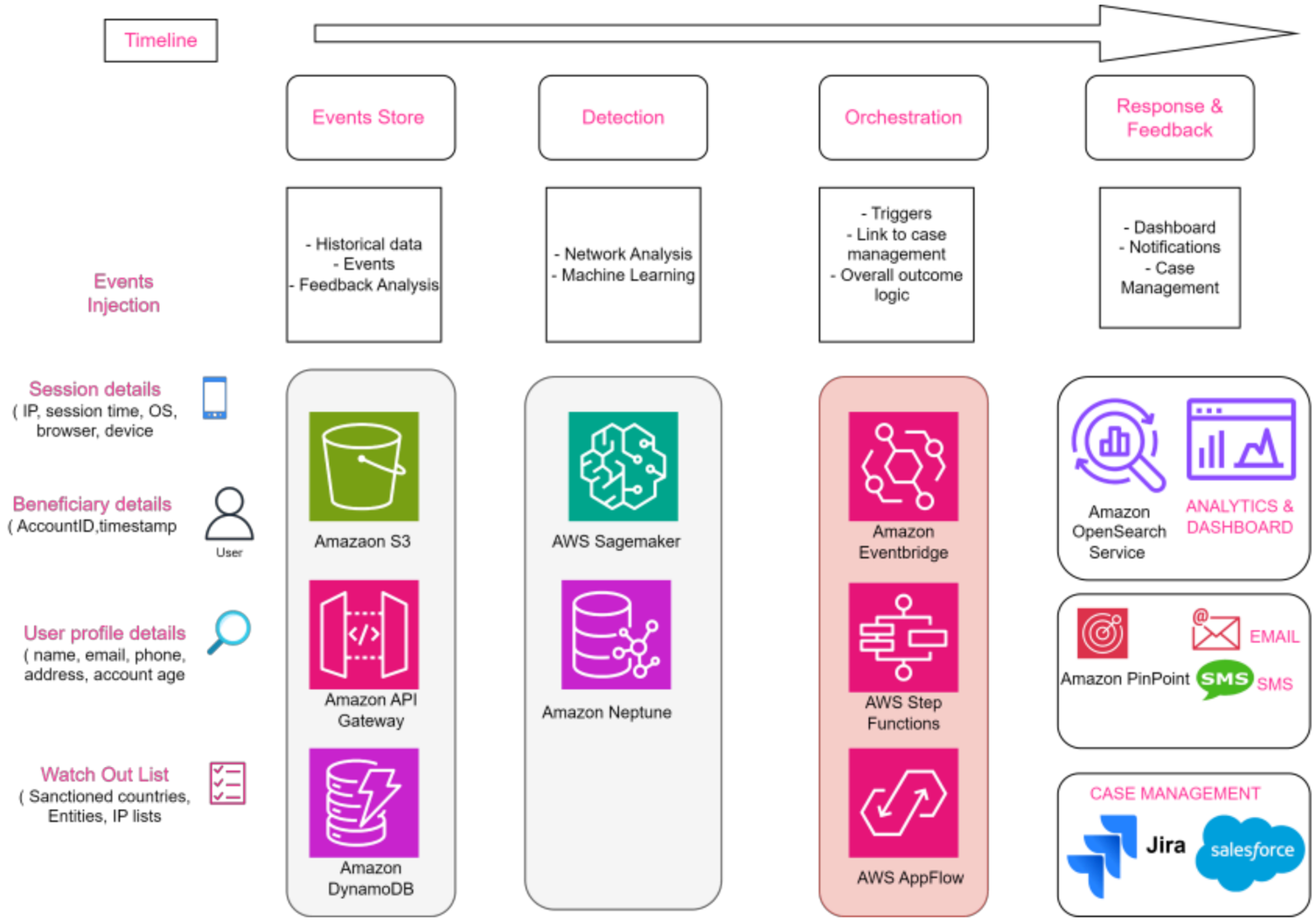
- As a boosting model, XGBoost creates very accurate predictions with low bias and a concentration on detecting edge cases.
- XGBoost is light weight, easy to train, and is more suitable for small dataset (~10k) than more complex models.
- Each step of the algorithm is interpretable, making it easier to be audited by human experts.

Model Performance

How our XGBoost Model performs on the public testing dataset

F1 Score	Training Speed	Inference Speed
0.998 / 1	**10.5 seconds**	**0.03 seconds**
Our F1 score indicates that the model for fraud detection achieves near-perfect precision and recall, effectively balancing false positives and false negatives.	The training speed, which includes time for cross-validation and randomized grid search for hyperparameter tuning, demonstrates high efficiency.	The includes both preprocessing and prediction time. It indicates an extremely fast model that is highly suitable for real-time fraud detection applications.

A Real Time Fraud Detection System with AWS



03
Conclusions & Recommendations

Conclusions

- Nighttime withdrawals under $50 are common fraud indicators.
- Melbourne data may be unreliable.
- Some users have 5+ fraudulent transactions with no valid ones.
- The elderly group is likely to commit fraud, but it might be by mistake

Recommendations

- Be more vigilant and track transactions that happen at night.
- Track and monitor transactions that deals with withdrawals.
- Continue monitoring high-risk users.
- Further investigate Melbourne situation to understand why there are not that many fraudulent transactions.

Thank you!

Skyline Financial Services (SFS)
Data Analysis Team 5/12