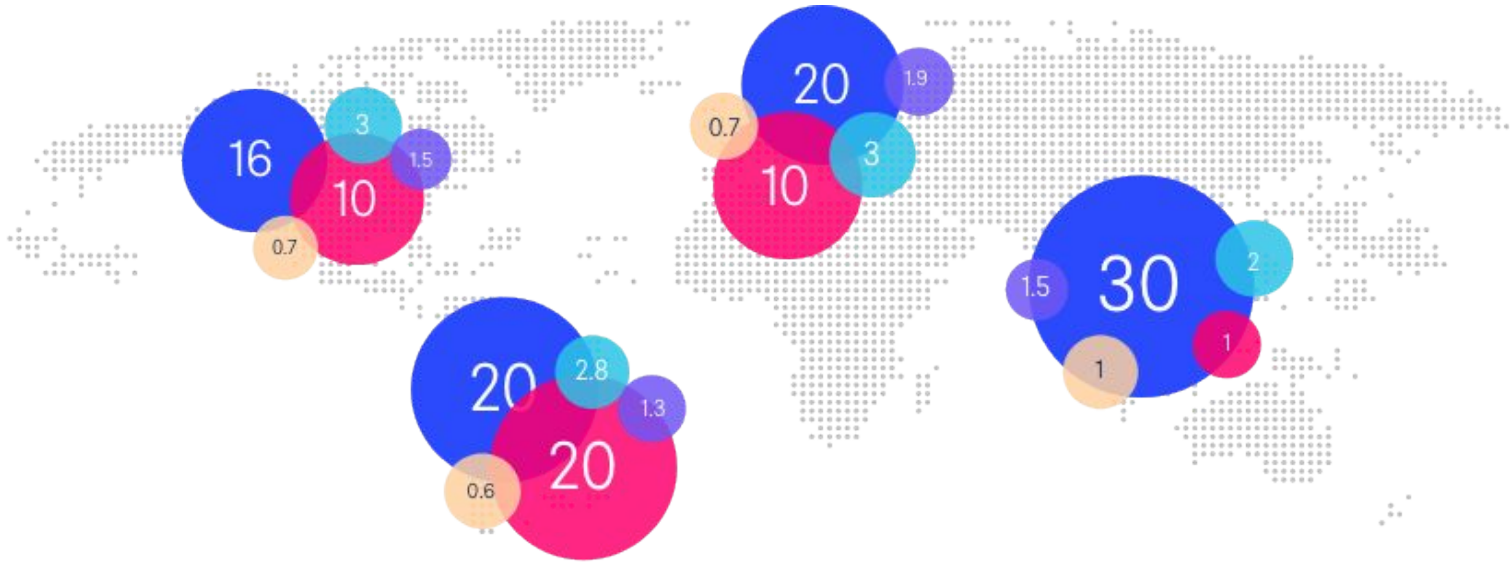
CREDIT CARD FRAUD DETECTION

QI RUI, JAYA PILLAI



METHODOLOGY

01

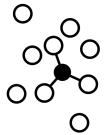
EDA

Exploring Data

02

SMOTE

Sampling & Borderline



03

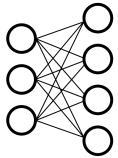
CLASSIFIER

Classifying models

04

STACKING

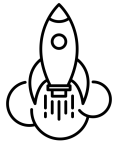
Multiple models



05

NEURAL NETWORK

Keras/ ANN

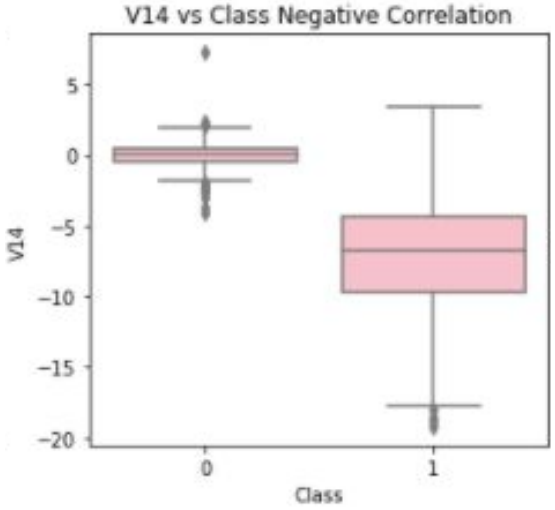


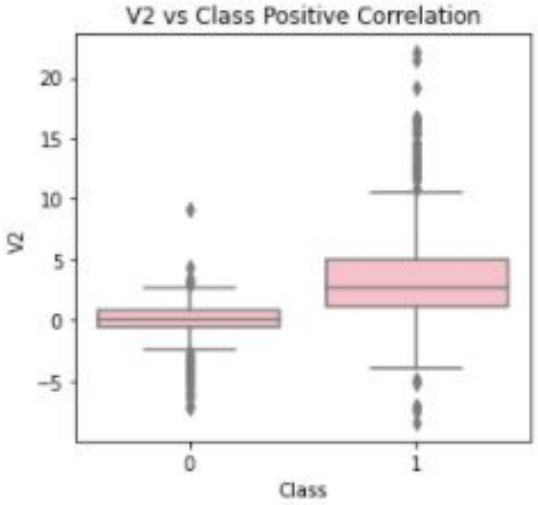
06

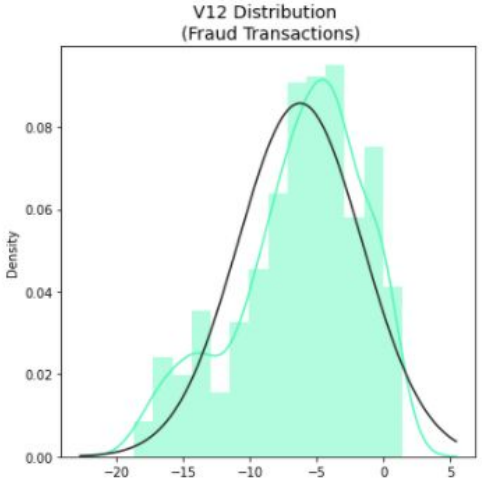
BOOSTING

Multiple models

THE DATASET


V14 vs Class Negative Correlation
V14
Class
5
0
−5
−10
−15
−20
0
1


V2 vs Class Positive Correlation
V2
Class
20
15
10
5
0
−5
0
1


V12 Distribution
(Fraud Transactions)
Density
0.08
0.06
0.04
0.02
0.00
−20
−15
−10
−5
0
5

FINDINGS AND RESULTS

OVERFITTING & CROSS VALIDATION

```
Logistic Regression Cross Validation Score:  93.54%
Knears Neighbors Cross Validation Score 93.01%
Support Vector Classifier Cross Validation Score 93.67%
DecisionTree Classifier Cross Validation Score 91.43%
```

SUCCESS CRITERIA

TRUE POSITIVE

Percent Correct Classification

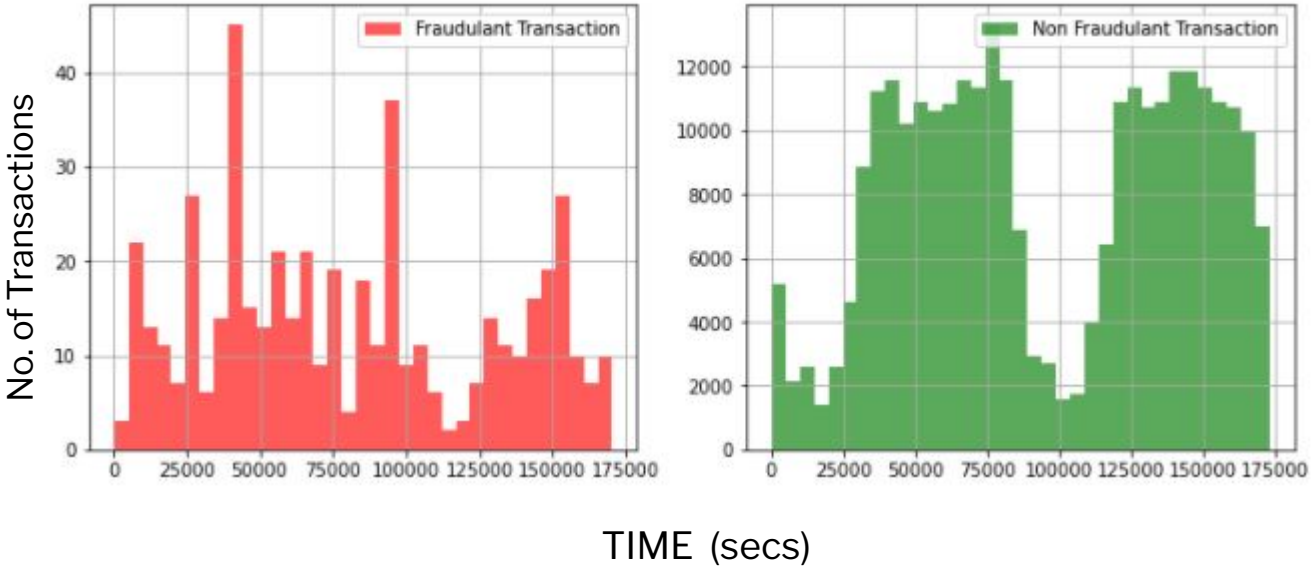
ROC

False Positive vs True Positive

CLASSIFICATION

Lesser Imbalance

FINDINGS AND RESULTS



RECOMMENDATIONS

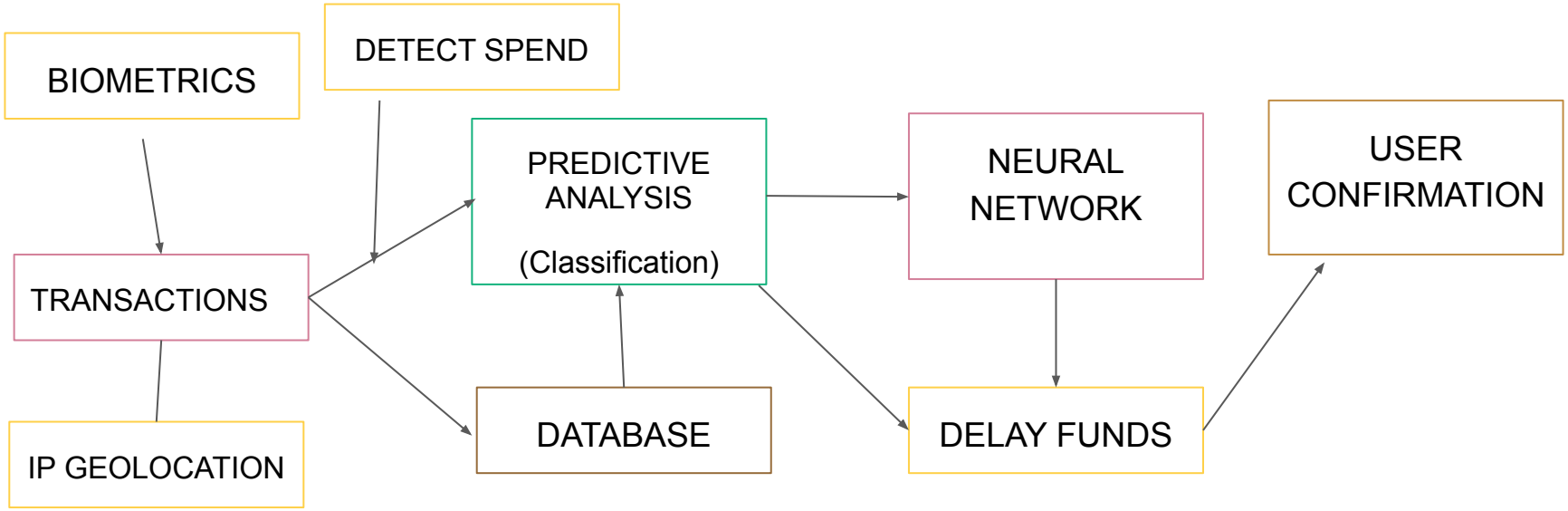
IP GEOLOCATION

BIOMETRICS IDENTIFICATION

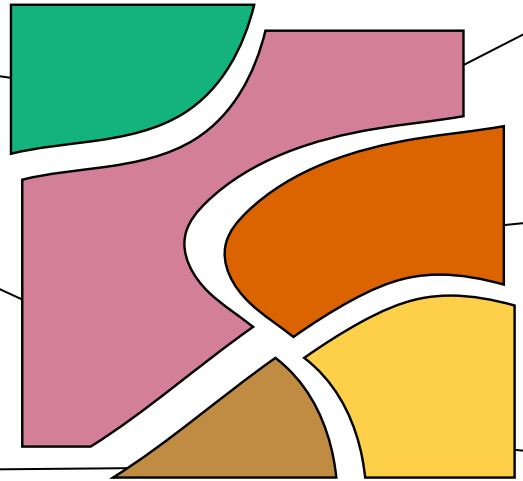
DETECT SPEND AMOUNT

DELAY FUNDS TRANSFER

BUSINESS FRAMEWORK



CHALLENGES



DATASET

PERFORMANCE MEASURE

HIGH COST OF MISCLASSIFICATION

ALGORITHM USED

ALGORITHM MAINTENANCE

DYNAMIC BEHAVIOUR OF FRAUDSTERS

FUTURE STEPS

01 SUPERVISED & UNSUPERVISED MODELS

02 CLUSTERING METRIC

03 REAL TIME ANALYTICS

04 QUALITY DATA

05 USER BEHAVIOUR

06 PINPOINTING FRAUD TIMINGS

THANKS!

Questions?

Appendix

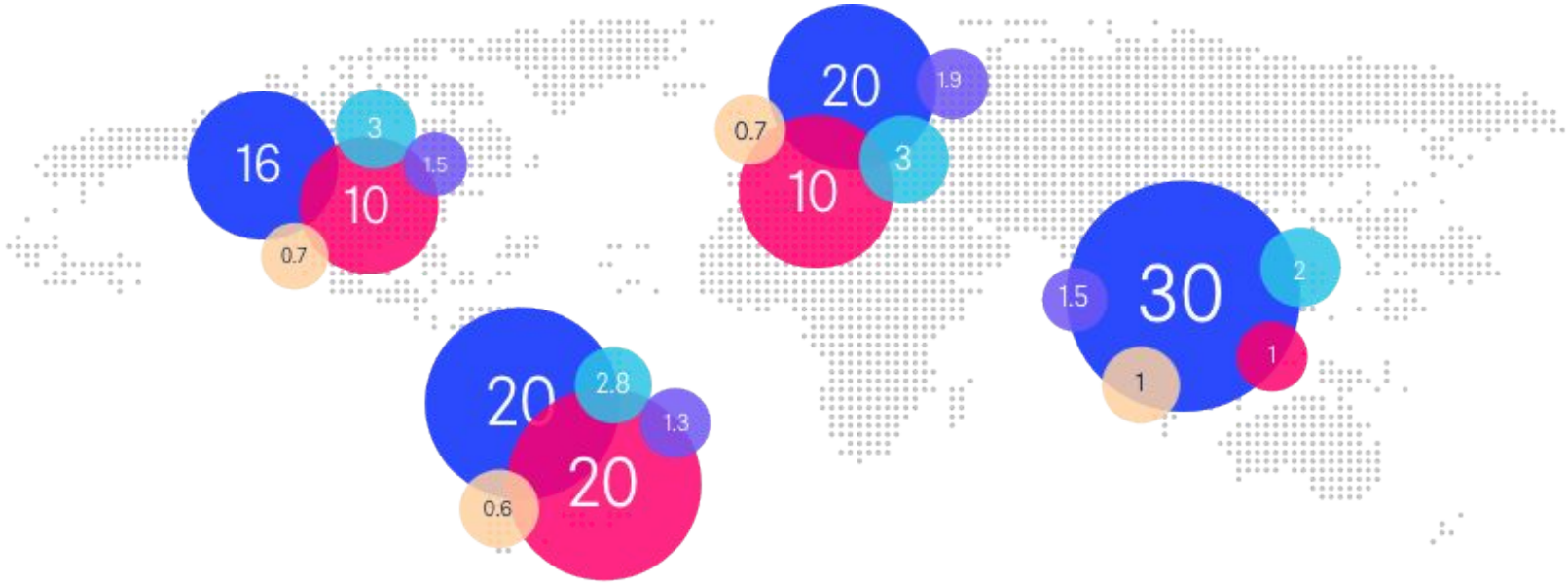
https://thenounproject.com/search/?q=exploring+&i=3382694

1. Exposure by Christopher T. Howlett from the Noun Project
2. Explore by Monkik from the Noun Project
3. k nearest Neighbours by sachin modgekar from the Noun Project
4. Stacking by Sachin Modgekar from the Noun Project
5. RBM by sachin modgekar from the Noun Project
6. Thumb scanning by Vectors Point from the Noun Project
7. Warning by myladkings from the Noun Project

Business Architecture
https://wso2.com/whitepapers/fraud-detection-and-prevention-a-data-analytics-approach/

https://geotargetly.com/blog/can-credit-card-fraud-be-minimised-with-ip-geolocation

https://www.cardrates.com/advice/credit-card-fraud-statistics/



Total % of revenue lost to fraud

Prevented % due to suspected fraud

Chargebacks as % of revenue

% of orders manually reviewed

% of reviewed orders declined