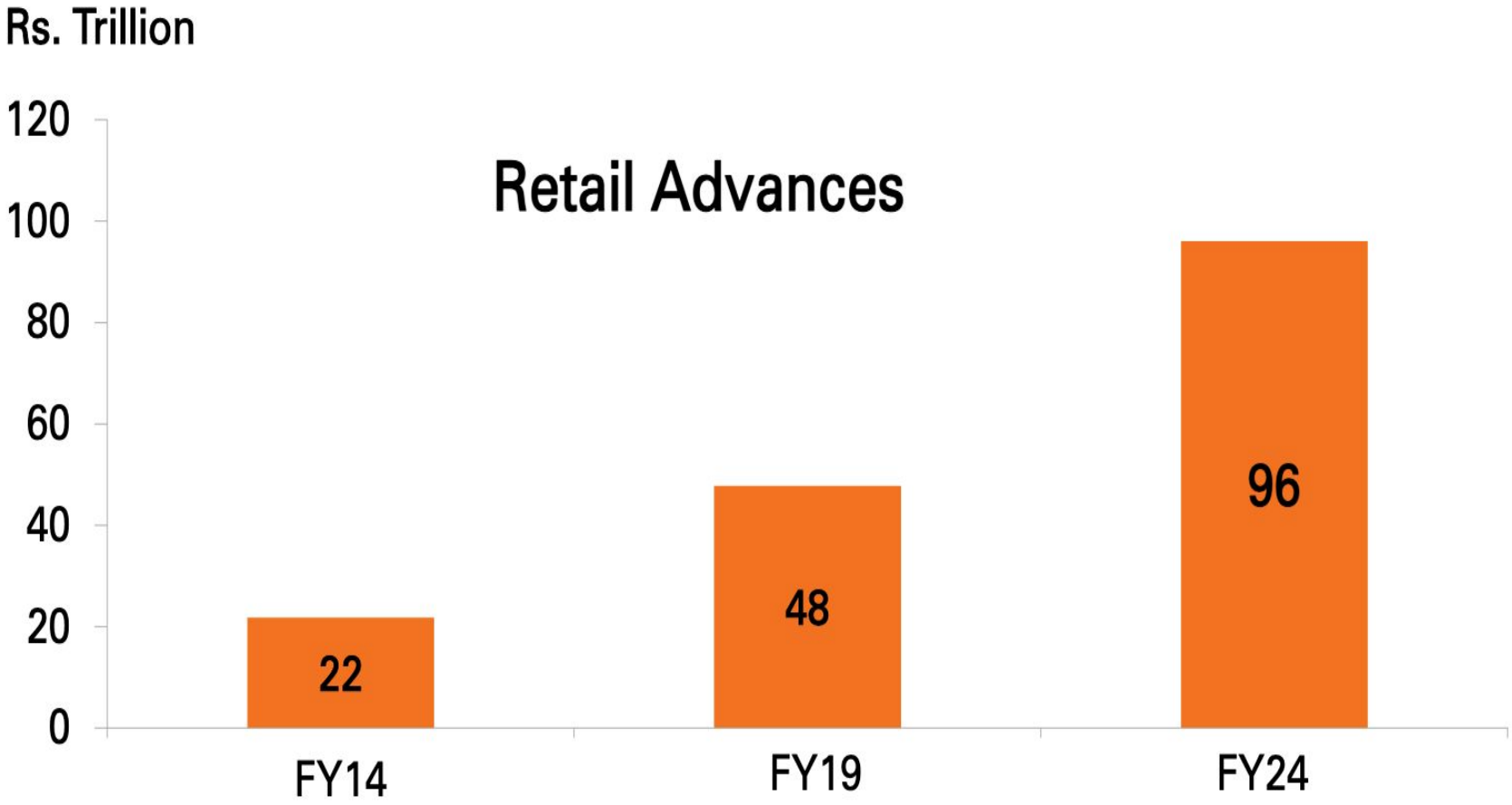
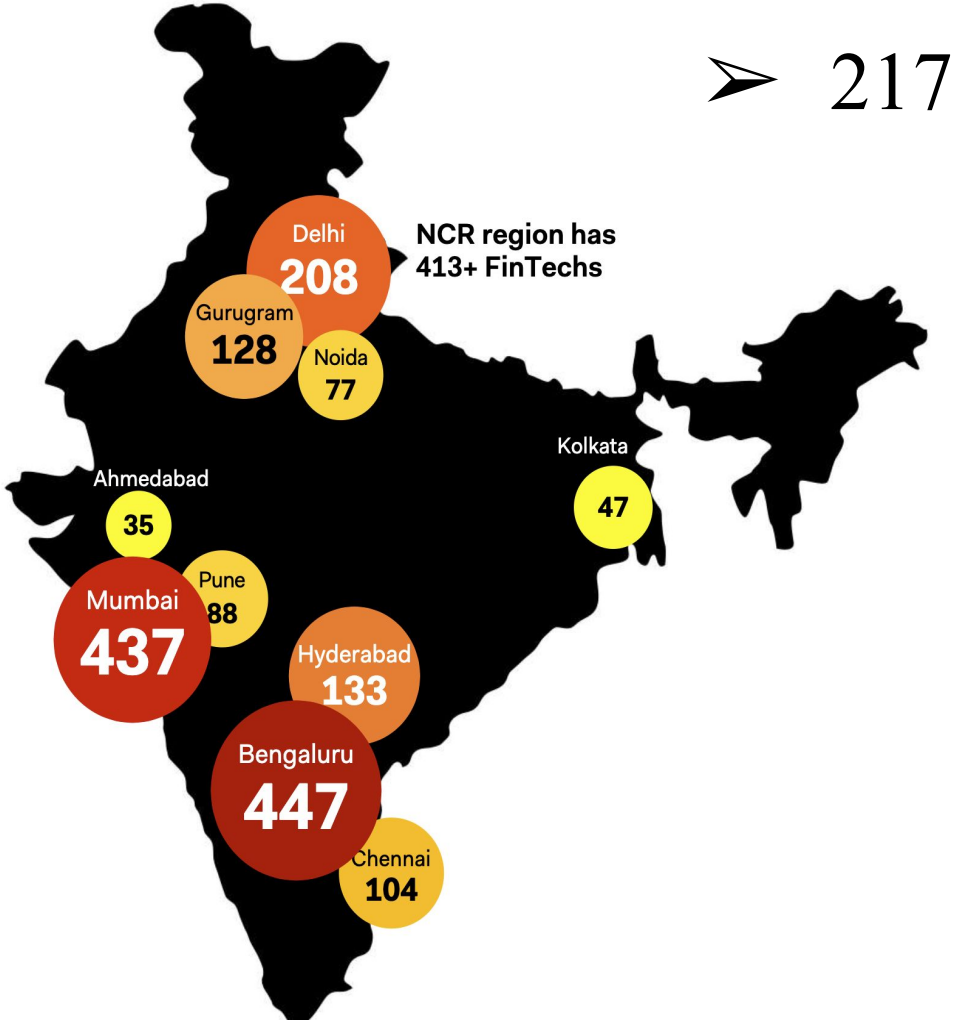
PREDICTING CONSUMER LOAN DEFAULTS IN INDIA


भारतीय रिज़र्व बैंक
RESERVE BANK OF INDIA


Rs. Trillion
120
100
80
60
40
20
0
Retail Advances
22
48
96
FY14
FY19
FY24



➢ 2174 Fin-tech Start-ups

TYPES

- Online Marketplaces
- Peer-to-Peer Lending
- Instant Loans
- Non-Bank Financing Corps
- Micro-Loan Crowd Funding

DATA

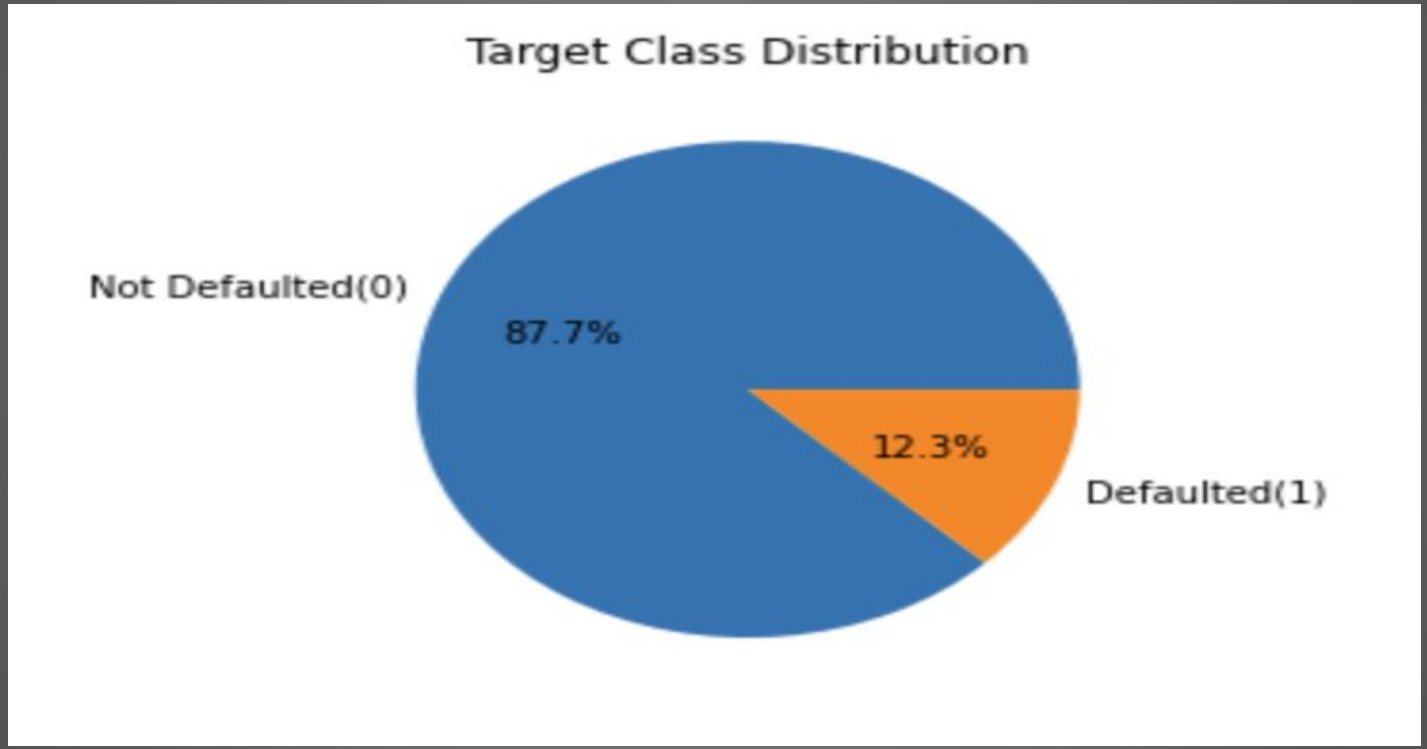
- University of Alabama
- Kaggle

- Observations 252,000

- Features: 10
  - Quantitative: 5
  - Categorical: 5

- Data Cleaning/EDA

CLASS IMBALANCE

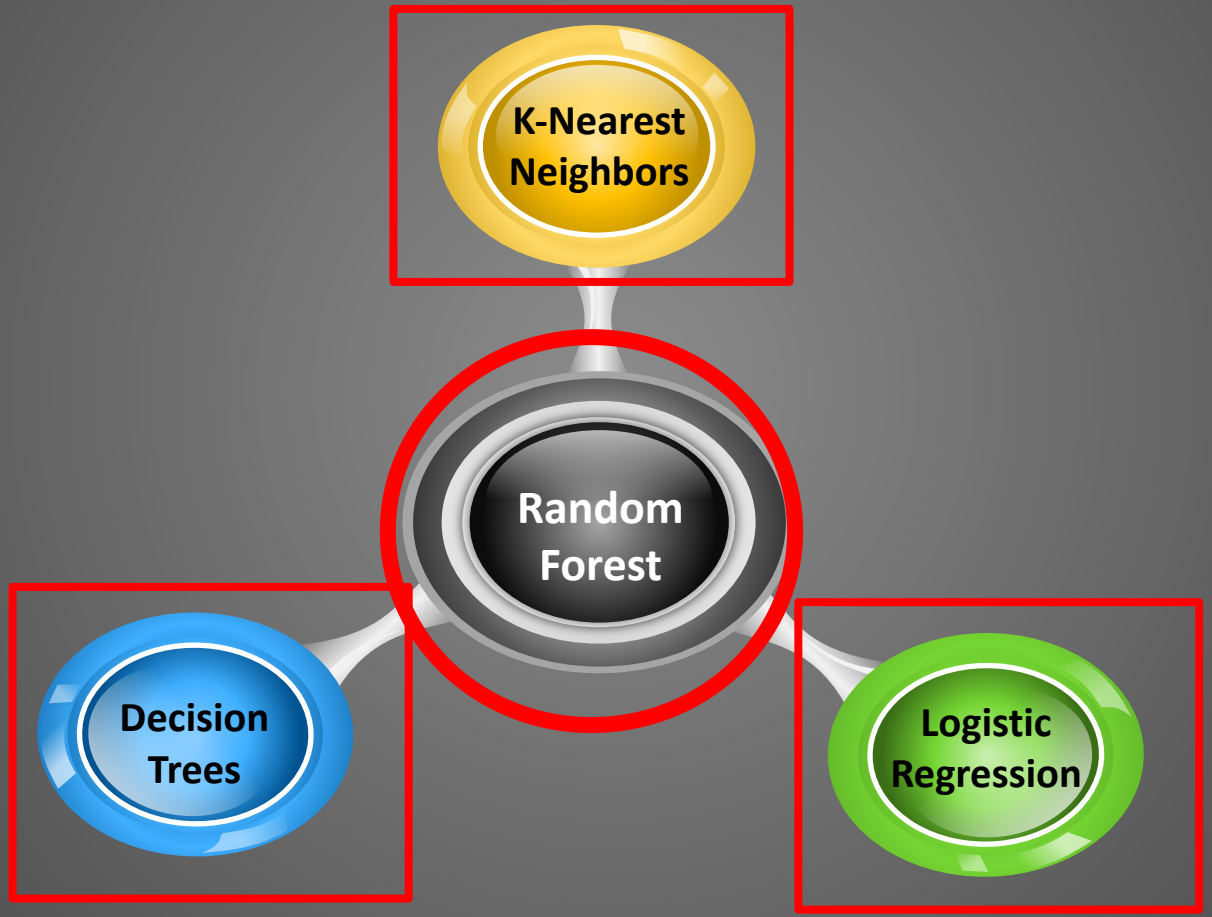


EVALUATION METRICS

- False Negatives(Recall)
  - Don’t approve those highly likely to default
    - Principle is a priority
    - Low opportunity cost of False Positive(Precision)

- F-Beta Score
  - Scale: 0 to 1
  - Beta Hyperparameter: Change relative importance of precision v. recall

Modeling Choices



scikit learn

matplotlib

seaborn

RANDOM FOREST

Bootstrapping

Decreases Variance/Limited Bias Increase

Corrects Overfitting

Decorrelate Trees/Feature Subsampling

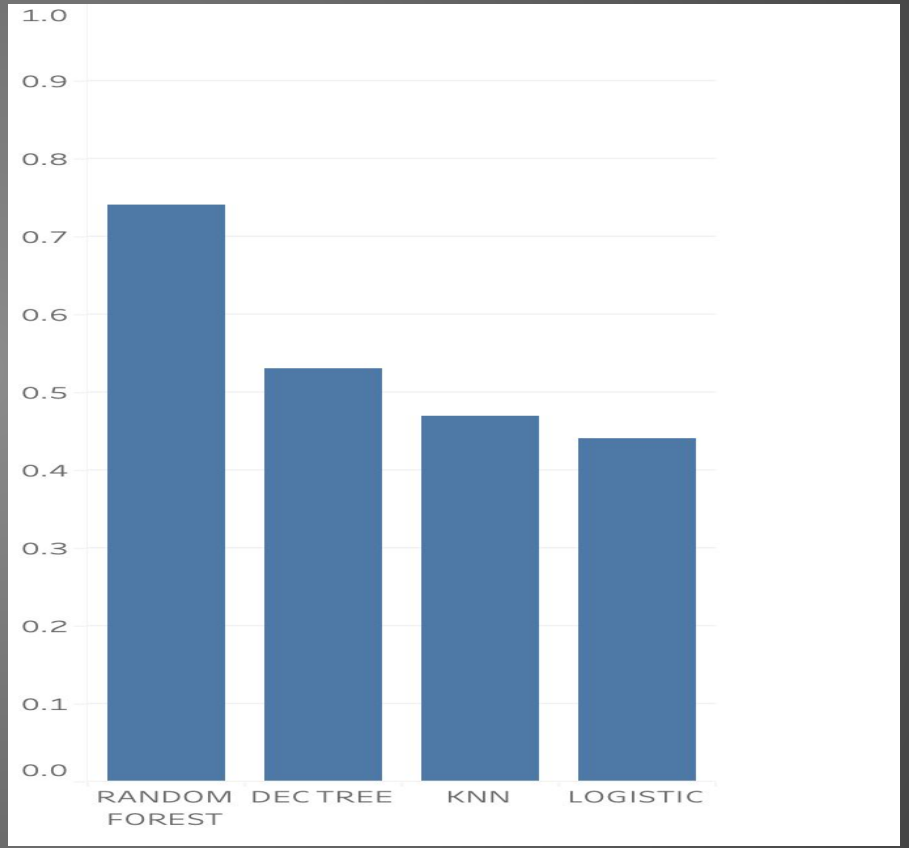
COMPARATIVE PERFORMANCE

F-Beta Score

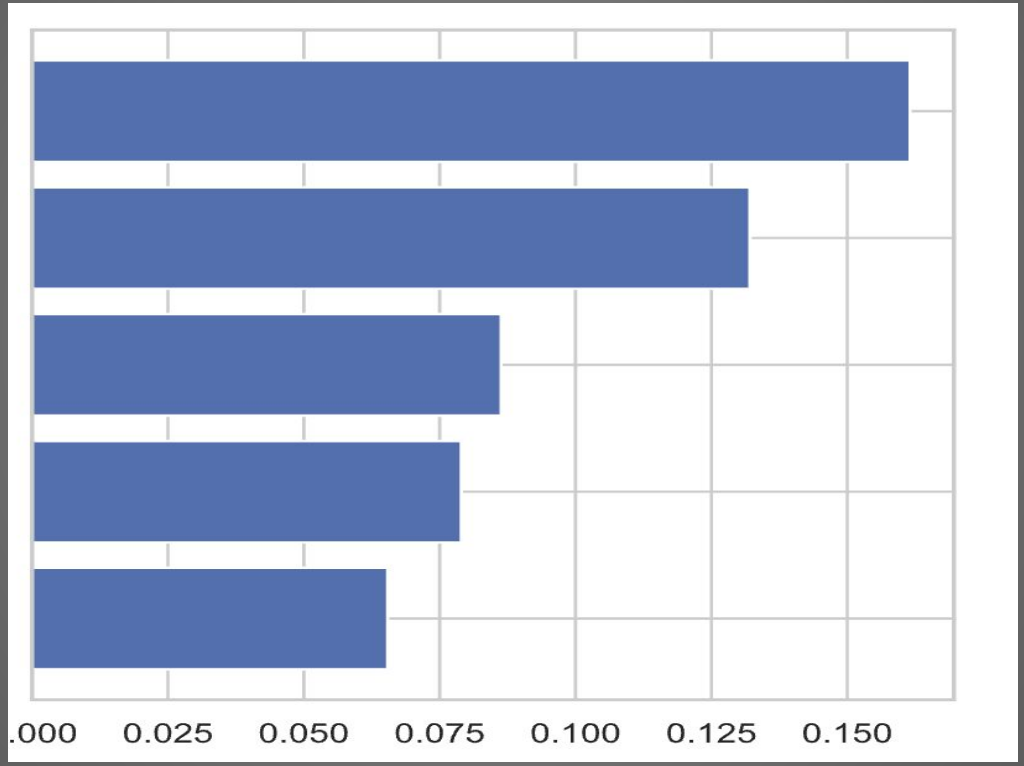
0.74

F-Beta Score

Beta = 3



Feature Importance



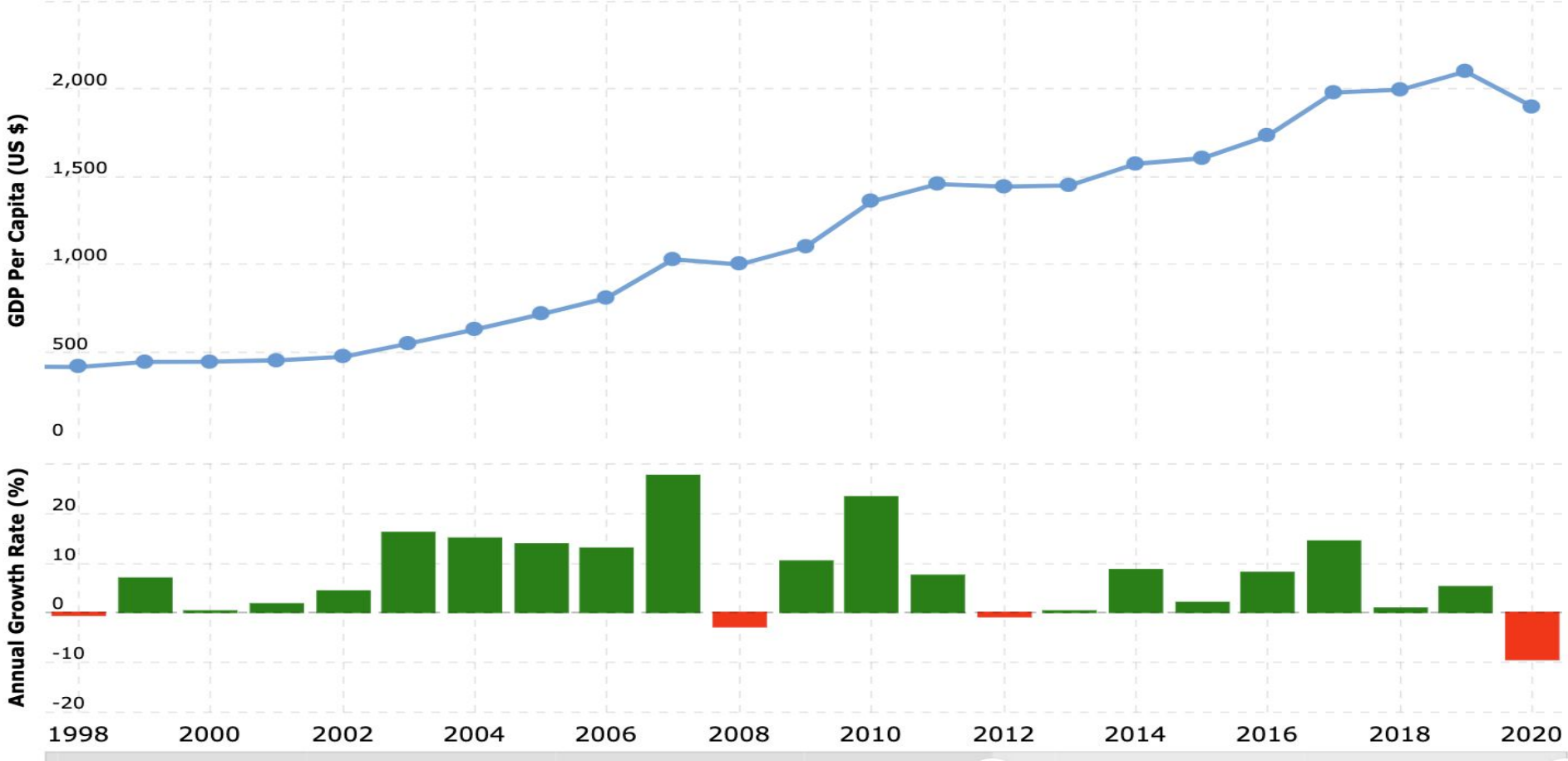
Future Work

- Add Rural v. Urban Feature
- Include dollar value of loans
- XGBoost


Thank you!

APPENDIX

Indian Per Capita GDP GROWTH



PAIR/HUE PLOT

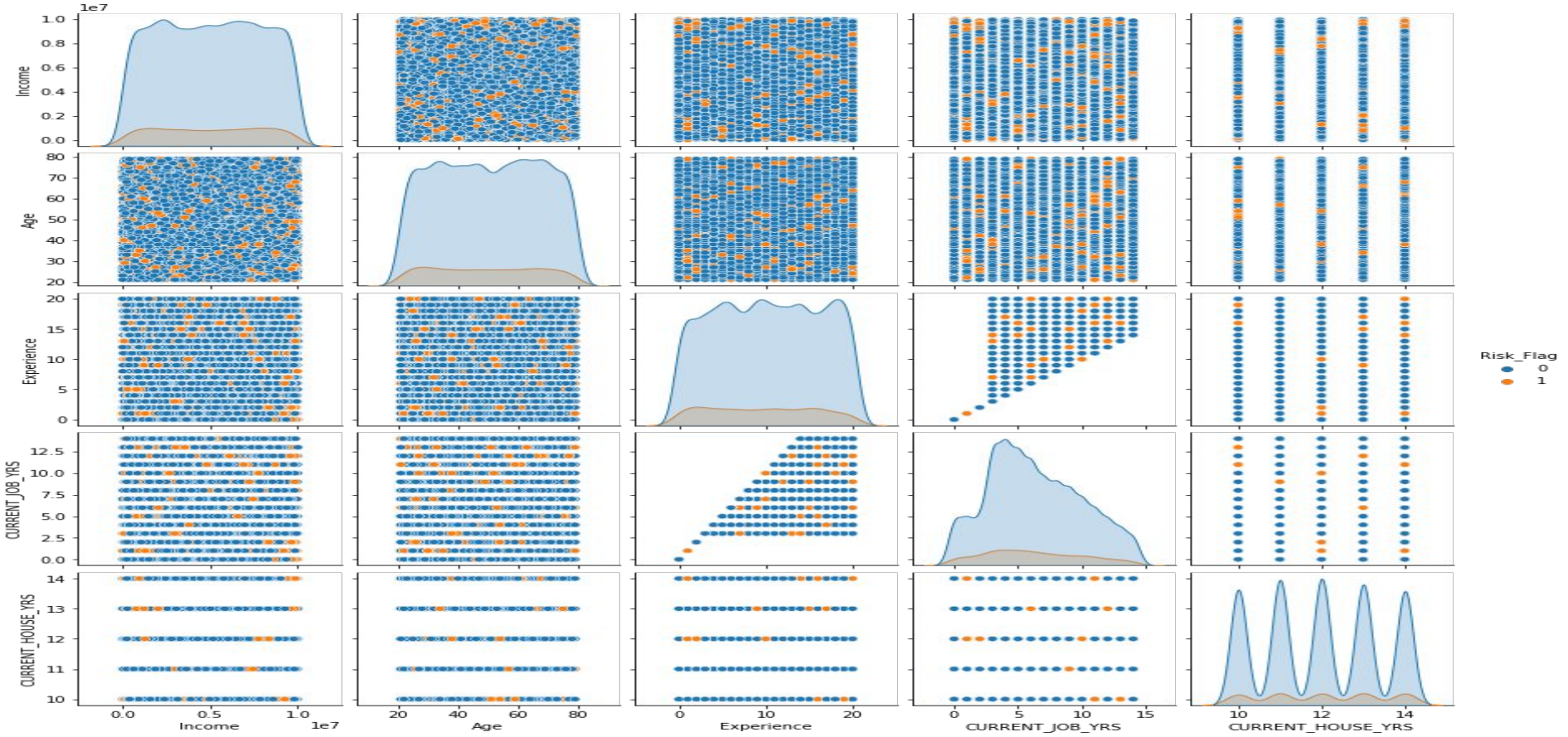


Figure 1. Retail financial services industry loans - India