



Module 4: Classification

CREDIT CARD CHURN

Objective

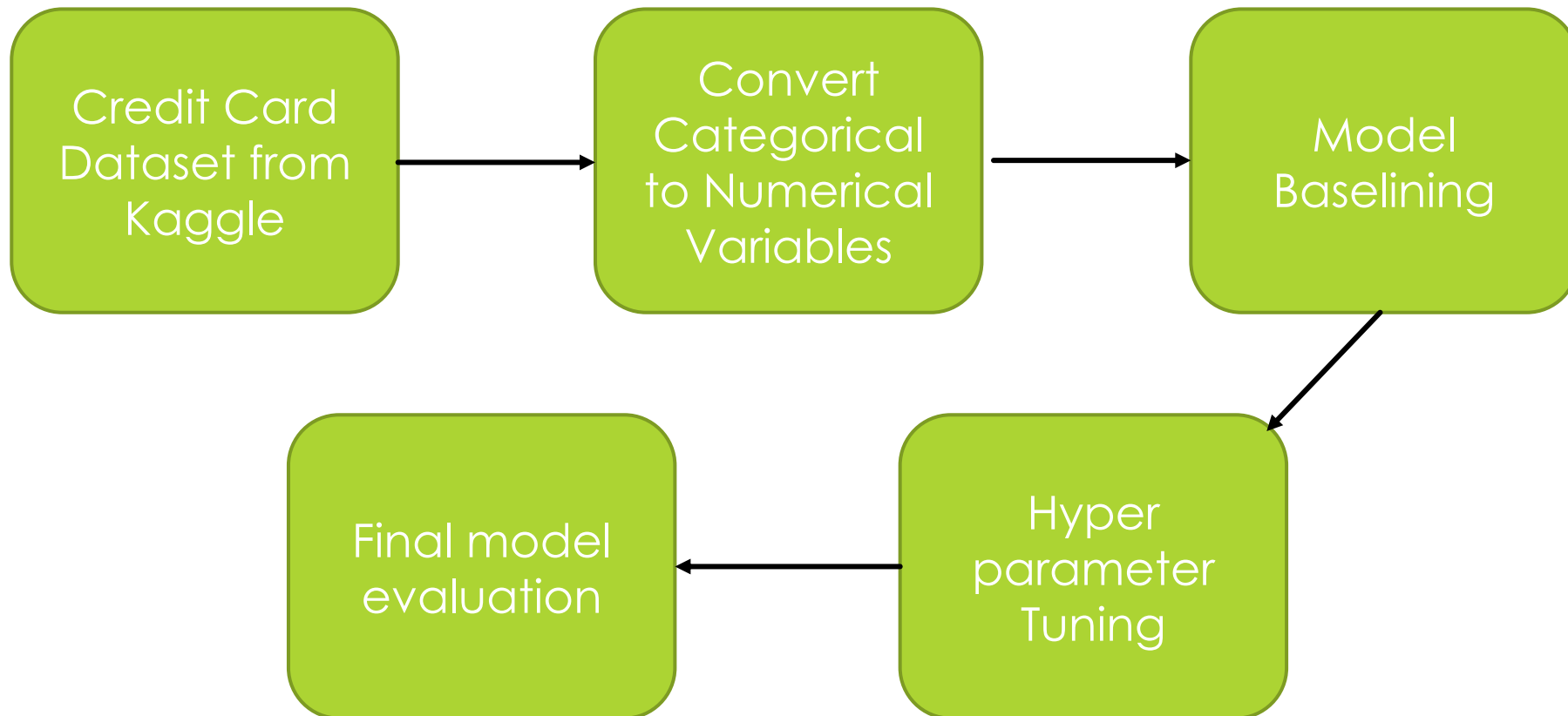
Credit card user churn is a big problem for credit card companies

In order to improve customer retainment, we must be able to predict what users are likely to “churn” or move on to other credit card companies

Let's use Machine Learning!



Process



Data

Demographic Info:

Gender,
Education,
Age,
Marital Status, etc.

Credit Card Behavior:

Total Transaction Amount,
Avg Utilization,
Avg change from Q4-Q1

Financial Status:

Card Category,
Income Category,
Credit limit,
Relationship count, etc

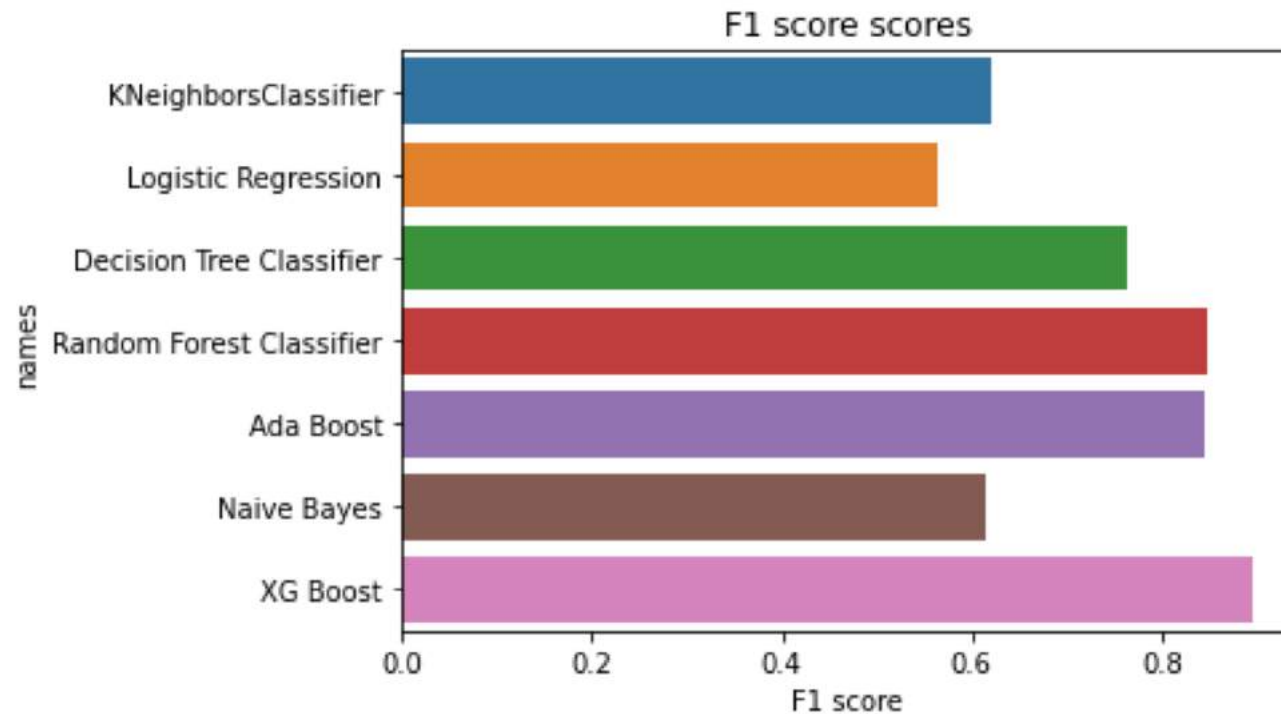
Features

Churn Status

0 : No churn
1 : churn

Target

Baselining



- ▶ 7 models tried out for initial screening
- ▶ F1 score is the most relevant metric
- ▶ **XG Boost is the best classifier**
- ▶ Naïve Bayes, Logistic Regression and Kneighbors are poor classifiers for this data

Hyperparameter Tuning

Tuned across 4 key
hyperparameters for XGBoost:

n_estimators: 50, 100, 150, 200, 300

Learning_rate: 0.2, 0.25, 0.3, 0.45, 0.6, 0.7

Reg_alpha: 0.0, 0.3, 0.5, 0.7

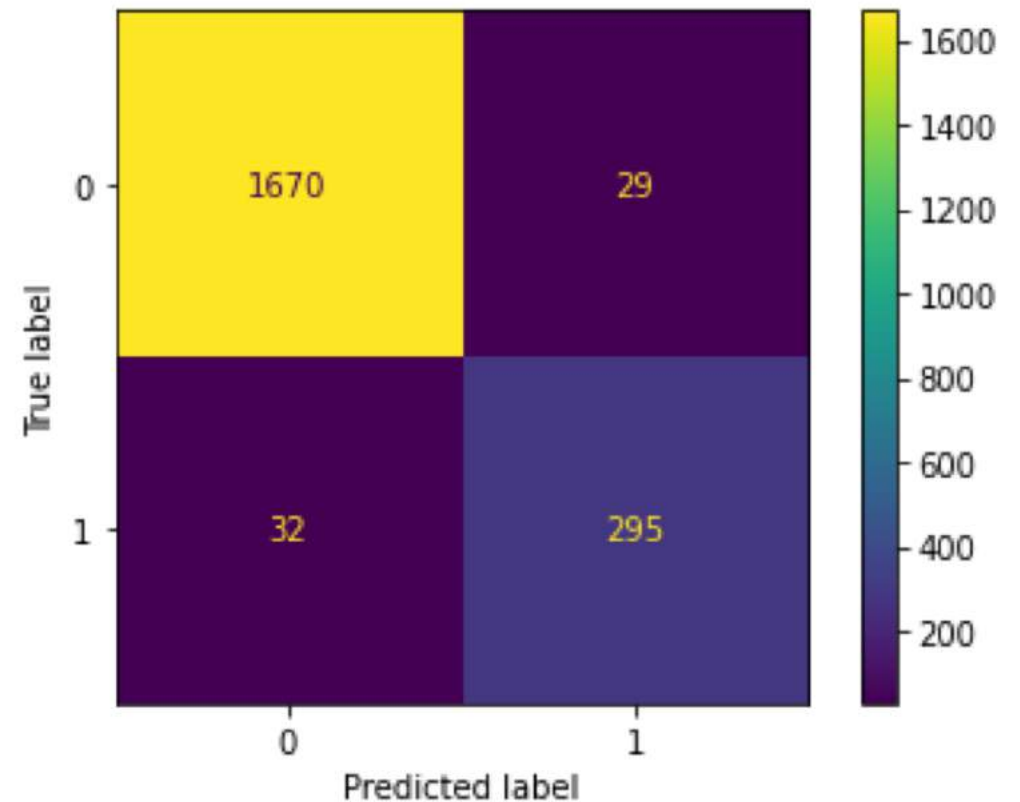
Scale_pos_weight: 0.5, 0.7, 0.85, 1

Screened 480 hyperparameter combinations

Final Optimized XGBoost Model

| | Precision | Recall | F-1 Score |
|---------------------|-----------|--------|---------------|
| Before Optimization | 91.40% | 87.77% | 89.55% |
| After Optimization | 91.05% | 90.21% | 90.63% |

The final tuned model is better in overall performance and is balanced



Conclusion

- ▶ XGBoost was the ideal model for our complex credit card churn problem
- ▶ After tuning hyperparameters, our best model has a F-1 score of 90.63%, exceeding the performance of our baseline model
- ▶ Using our model, we can target customers likely to churn, and give them targeted services and promotions

Next Steps

- ▶ Determine if the current churn rate is acceptable and if not, devise mitigation strategies targeted at customers with high probabilities of churning
- ▶ Continue to improve the current model with more granular hyperparameter optimizing, gathering different data from customers, and updating the model as more users are added to the database.