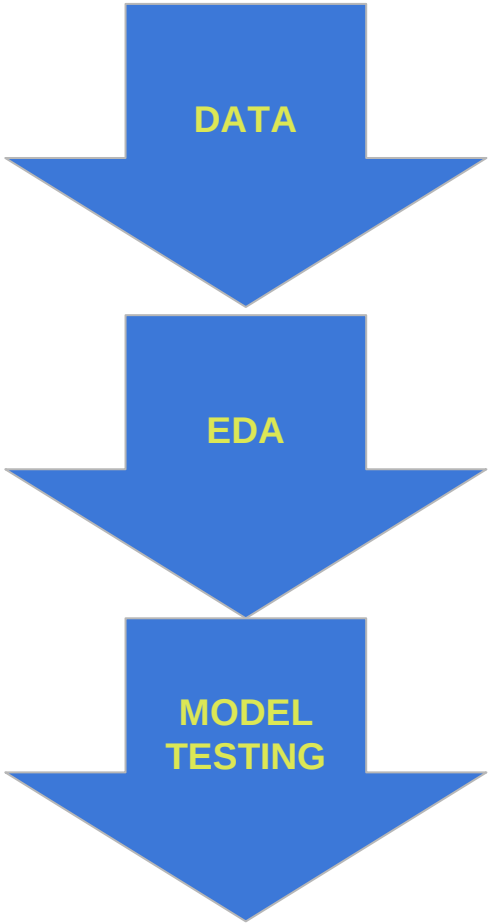
Predicting Customer Churn

By: SHIVAM CHAUDHARY

Motivation

- Maintain Customer Base
- Telecommunication Sector
- Focus retention efforts
- Predict customers at risk of churning.

Process



kaggle

(7043 x 21)

tableau SOFTWARE

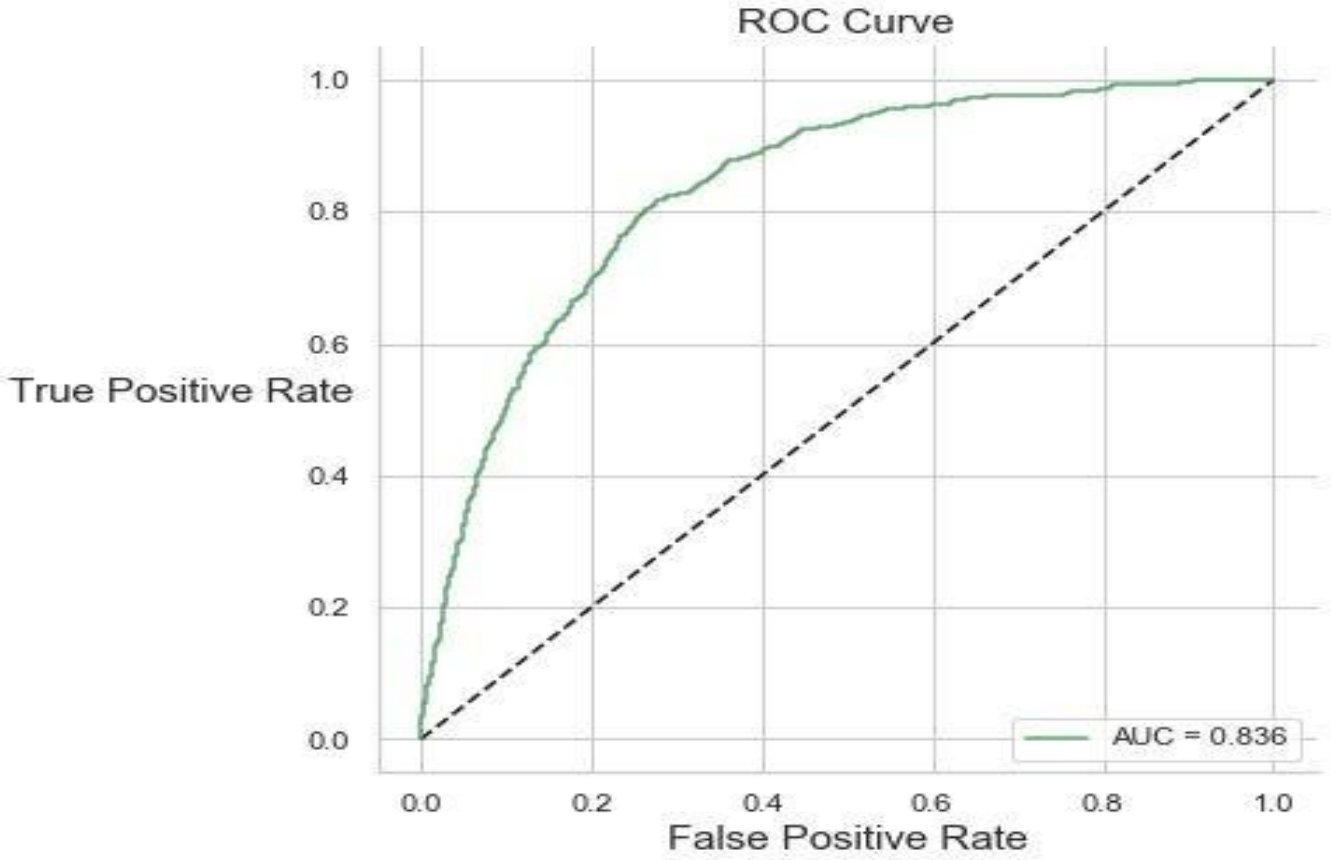
pandas

matplotlib

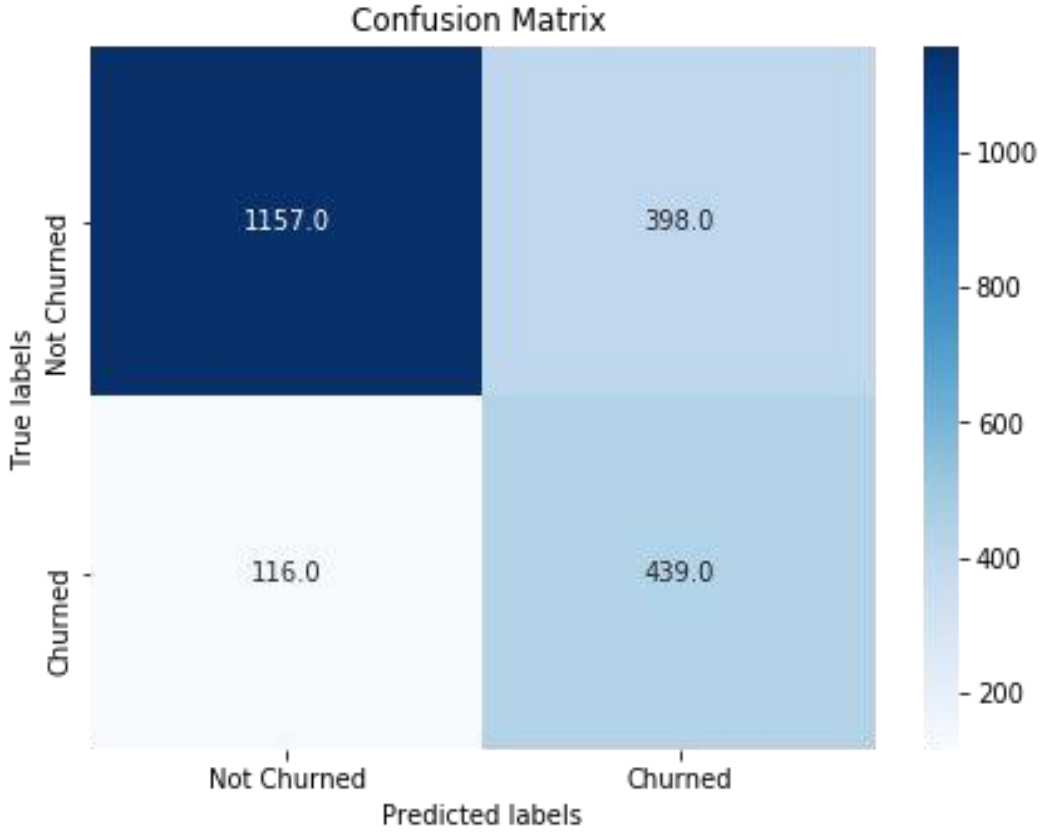
scikit learn

LOGISTIC REGRESSION

Final Model (Logistic Regression)



What does the Model Say?

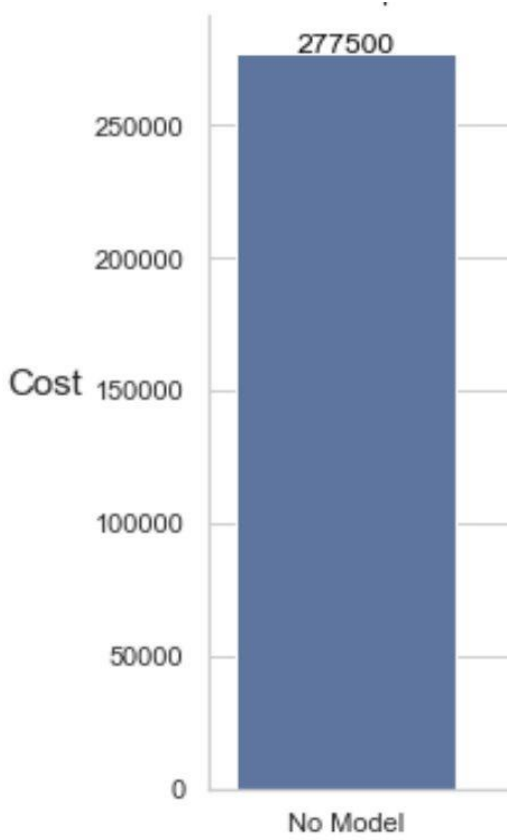


For Churned Customers
Recall:0.79

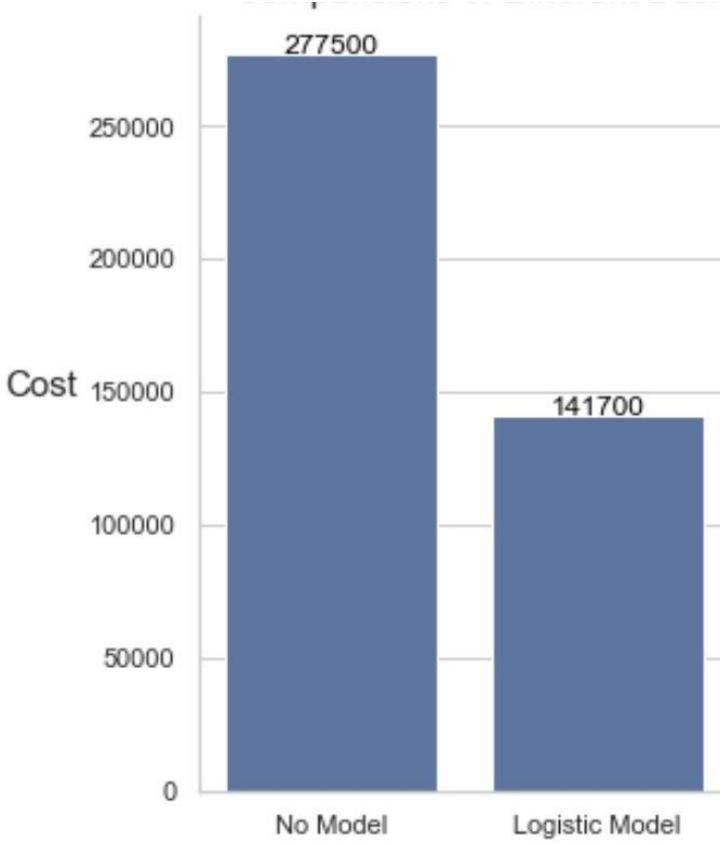
Cost Evaluation (Assumptions)

- Customer Acquisition Cost =5 x Customer retention costs
- Customer Acquisition Cost: USD 500
- Customer Retention Cost: USD 100
- Total Cost to maintain Customer Base= FN(500) + FP(100)+TP(100)+TN(0)

Cost Evaluation (Without Predictive Model)

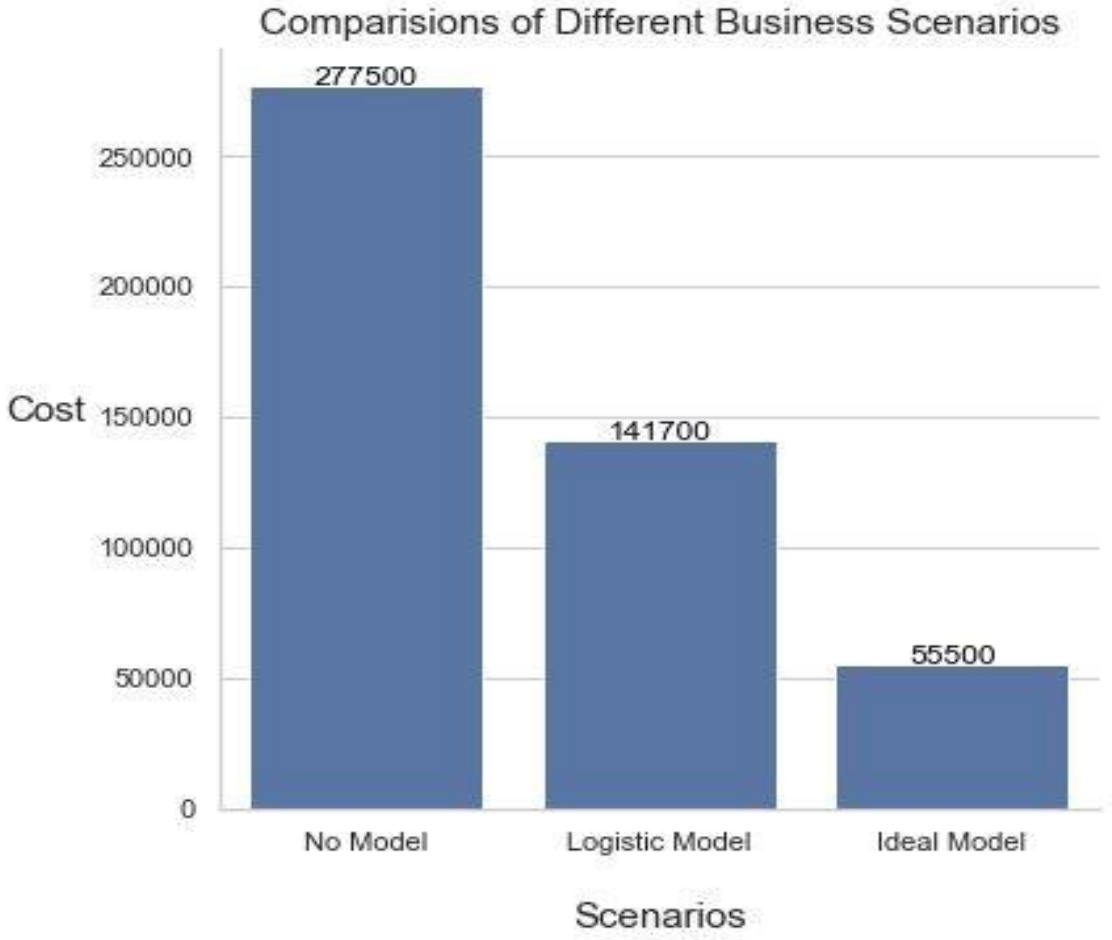


Cost Evaluation (With Current Model)



Save USD 135,800 in a month

Cost Evaluation (With an Ideal Model)



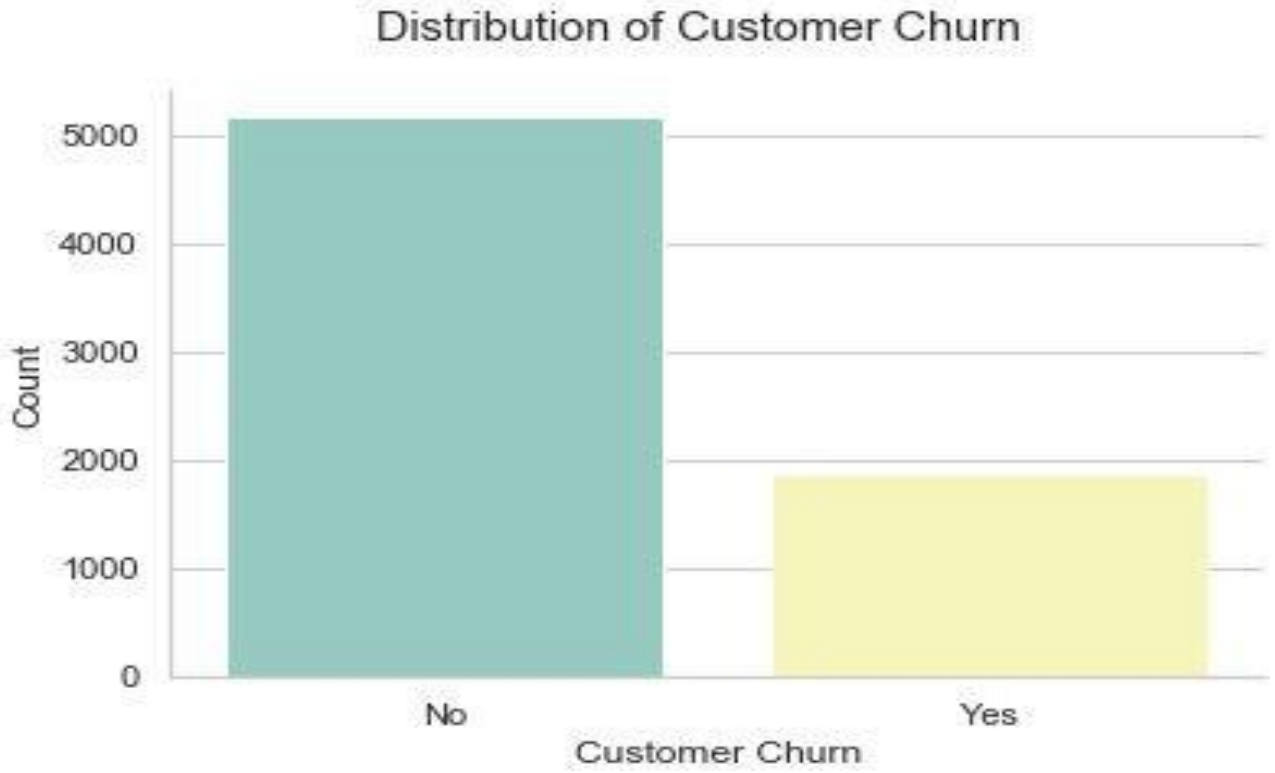
Future Work

- Optimize Model by ;
  - ★ enriching the dataset
  - ★ including costs in determining the optimal threshold
  - ★ Distinguish between voluntary & involuntary churners
- Make current web app better

Thank you

Appendix

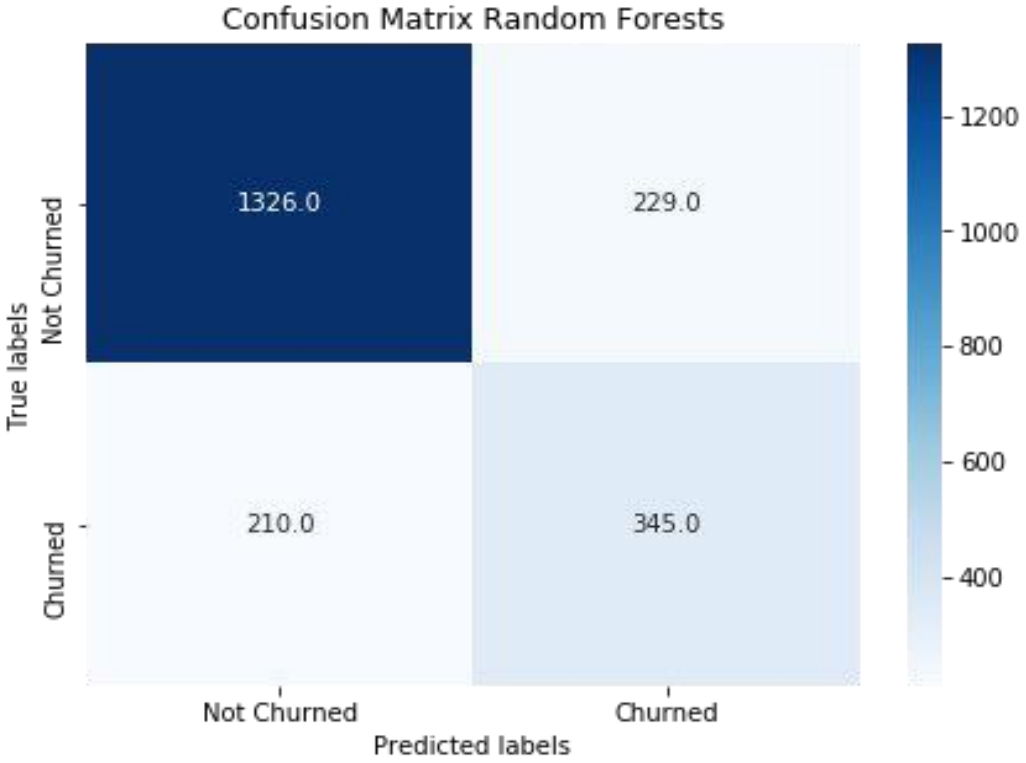
Distribution of the Target Variable



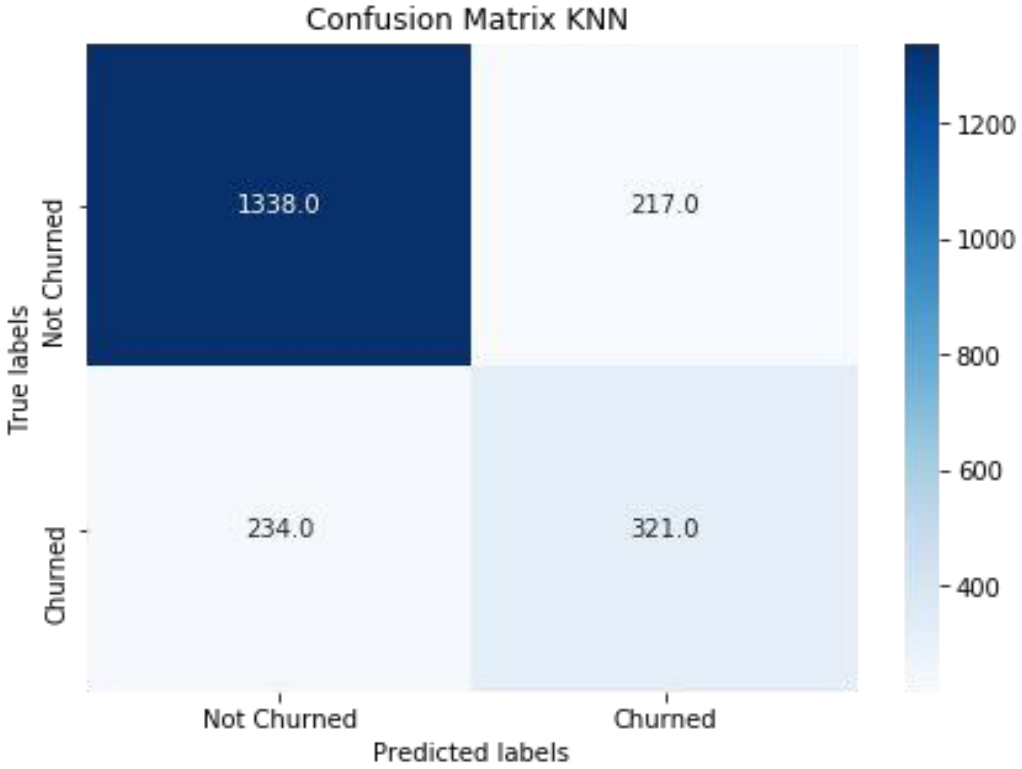
Logistic Regression

	precision	recall
0	0.91	0.74
1	0.52	0.79

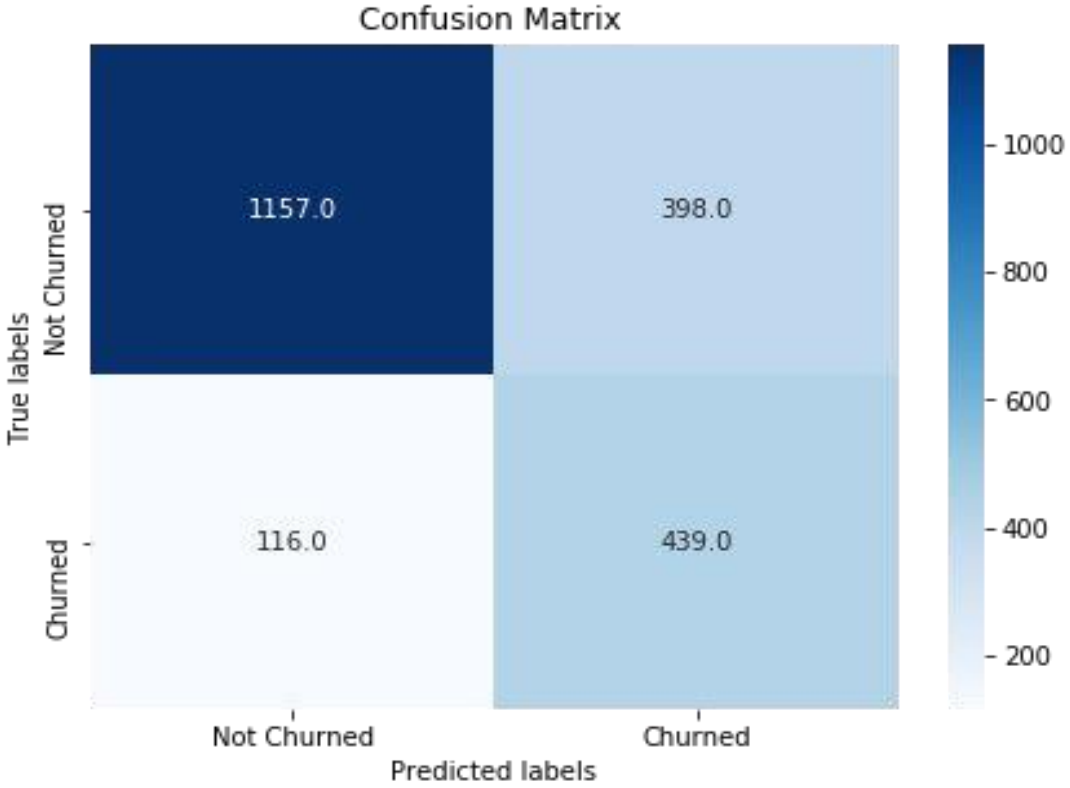
Random Forests Confusion Matrix



KNN Confusion Matrix



Confusion Matrix:Logistic Regression



Cost Calculations

- **Worst Case Scenario:** We assume that no customer will churn.But in reality, 555 customers ended up churning.So to maintain the customer base, the business will have to spend ($500*555) **USD 277,500** .
- **Applying Model:** The models identifies, FNs, TPs & FPs.So to maintain the current customer base, the business will have to spend (116* 500)+(837* 100) **USD 141700.** So using a model will **save USD 135,800** in a month.
- **Best Case Scenario:** We have a model that correctly identifies customers about to churn (FN=0 & FP=0).So to maintain the customer base, the business will have to spend (555*100) **USD 55,500**.

Current Web App