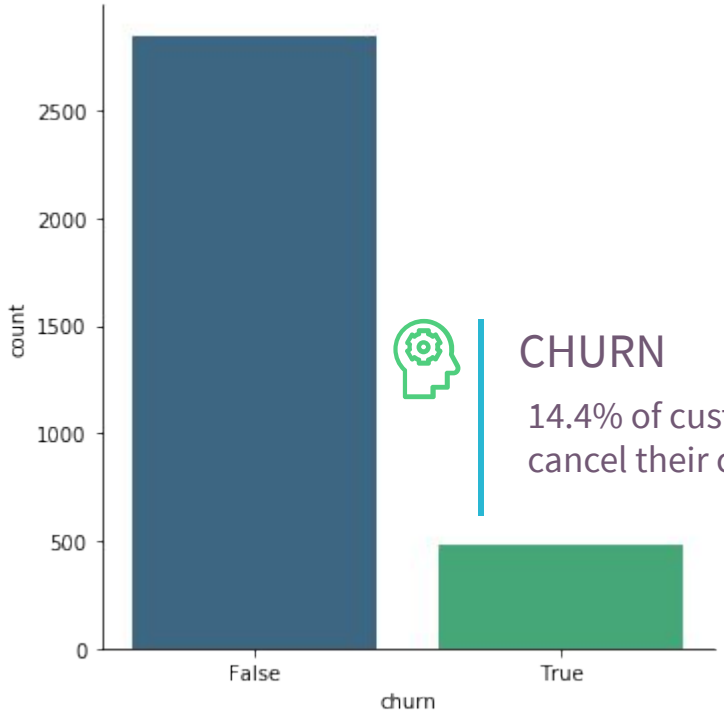
Predict Customer Churn for a Telecom Company

George Paskalev

FLATIRON SCHOOL

THE PROBLEM



CHURN

14.4% of customers cancel their contract

- Hired by a telecom company to assist with churn prediction.
- It costs $315 for a telecom company to acquire new customers.
- What impacts a customer's decision to switch to a competitor?

THE DATA

SOURCE

Dataset from Kaggle

SHAPE

Over 3300 rows, 20 columns/predictors

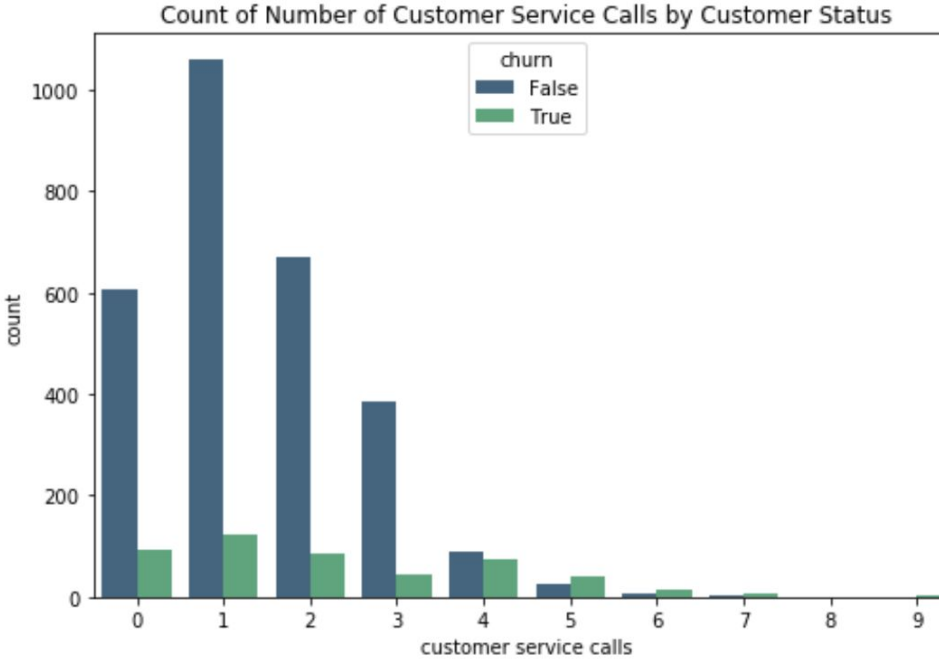
ISSUES

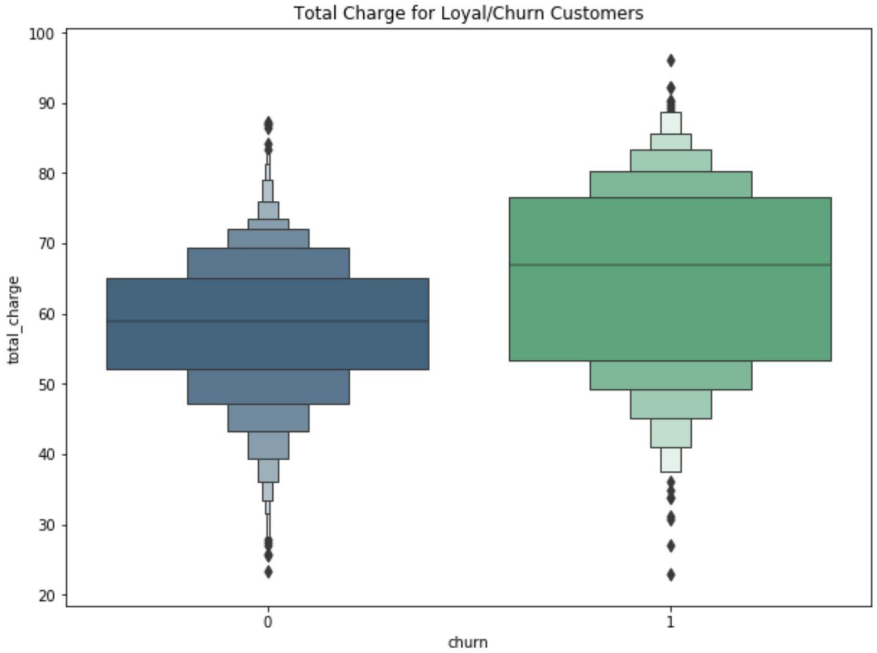
Pretty clean set, only issue is class imbalance

OUTLIERS

No major outliers that could skew the models

INTERESTING FEATURES





VANILLA MODELS

	Model	Recall Score	Accuracy(Test)	F1 Score
0	Log Regression	0.768	0.779376	0.510638
1	KNN	0.568	0.718225	0.376658
2	Naive Bayes	0.632	0.533573	0.288848
3	Decision Tree	0.880	0.928058	0.785714
4	Random Forest	0.792	0.947242	0.818182
5	Bagging Classifier	0.880	0.966427	0.887097
6	AdaBoost	0.784	0.906475	0.715328
7	Gradient Boosting	0.880	0.979616	0.928270
8	XGB	0.880	0.971223	0.901639
9	SVM	0.624	0.836930	0.534247

MAIN METRIC:
Recall

FINAL MODEL

A CLASSIFIER USING GRADIENT BOOSTING

93.28%
RECALL (TRAIN)

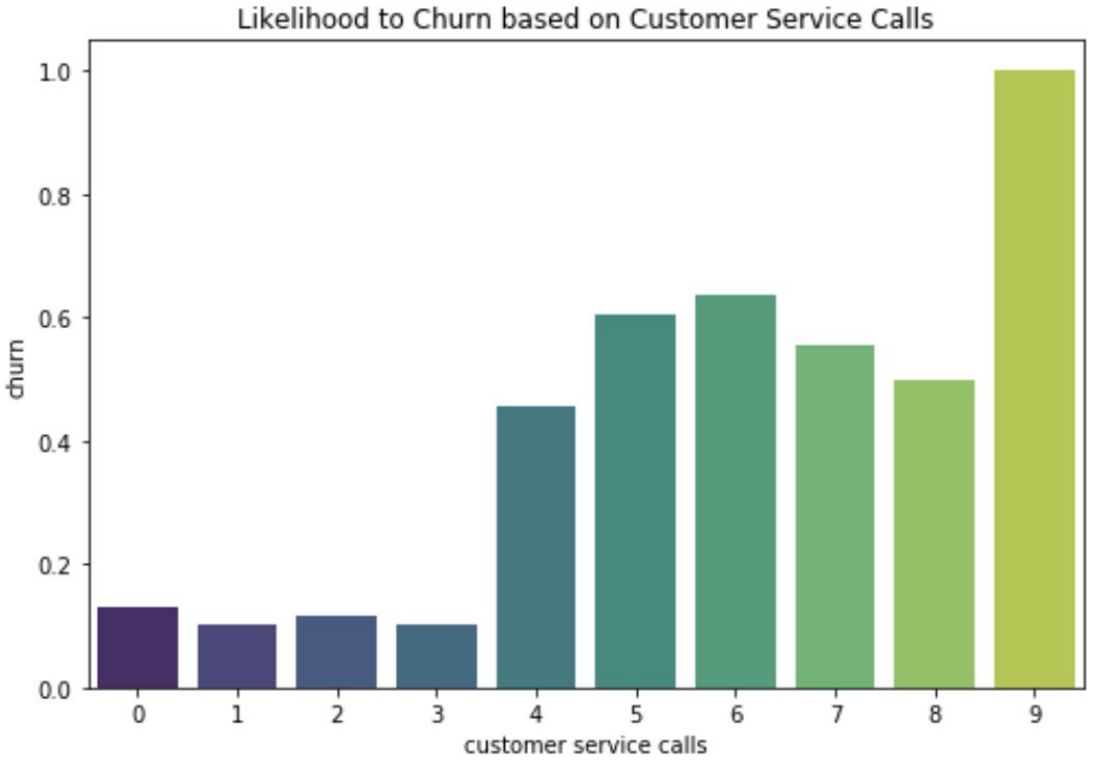
88%
RECALL (TEST)

98.2%
ACCURACY

KEY PREDICTORS

Customer Support Calls:
Any customer making 4 or more calls to your support team is at risk.

Total Charge:
Customers paying over $40 are at higher risk of leaving.



NEXT STEPS:

- Implement Customer Retention Strategy around Key Predictors
- Collect data on competitors in each state and add it to the model
- Analyze coverage in cities/states and see if that has an impact on customer churn

THANK YOU!!!

For more details, visit my GitHub:

https://github.com/georgepask