Flatiron School

Data Science Course

Maria Antonietta Ricci

Module 3

SyriaTel
Churn Rate Predictive Classifier

Business Problem

- This analysis aims to:
  - investigate the presence and significance of patterns of the Customer Churn Rate
  - build a classifier predicting whether a customer will churn

- What is the Churn Rate?
  - phenomenon where customers of a business no longer purchase or interact with the business for products or services

- Why is the Churn Rate important?
  - The customer satisfaction keeps the business running and thriving

Hypothesis

Is it possible to build a model to predict the customer based on this set of features given in the dataset?

Binary Classification Problem

Dataset

- SyriaTel Company
- Details:
  - Small dataset: 3300 entries per 21 columns
  - Columns dropping and dummy variables used
  - No missing values

Results

- CRITERIA:
  - Metrics
  - Business Scenarios

Results

- Metrics (Evaluation and Comparison):
  - Precision
  - Recall
  - Accuracy
  - F1 score
  - ROC and AUC

Results

- Most Relevant Business Scenario:
  - Customers incorrectly identified as not intending to churn

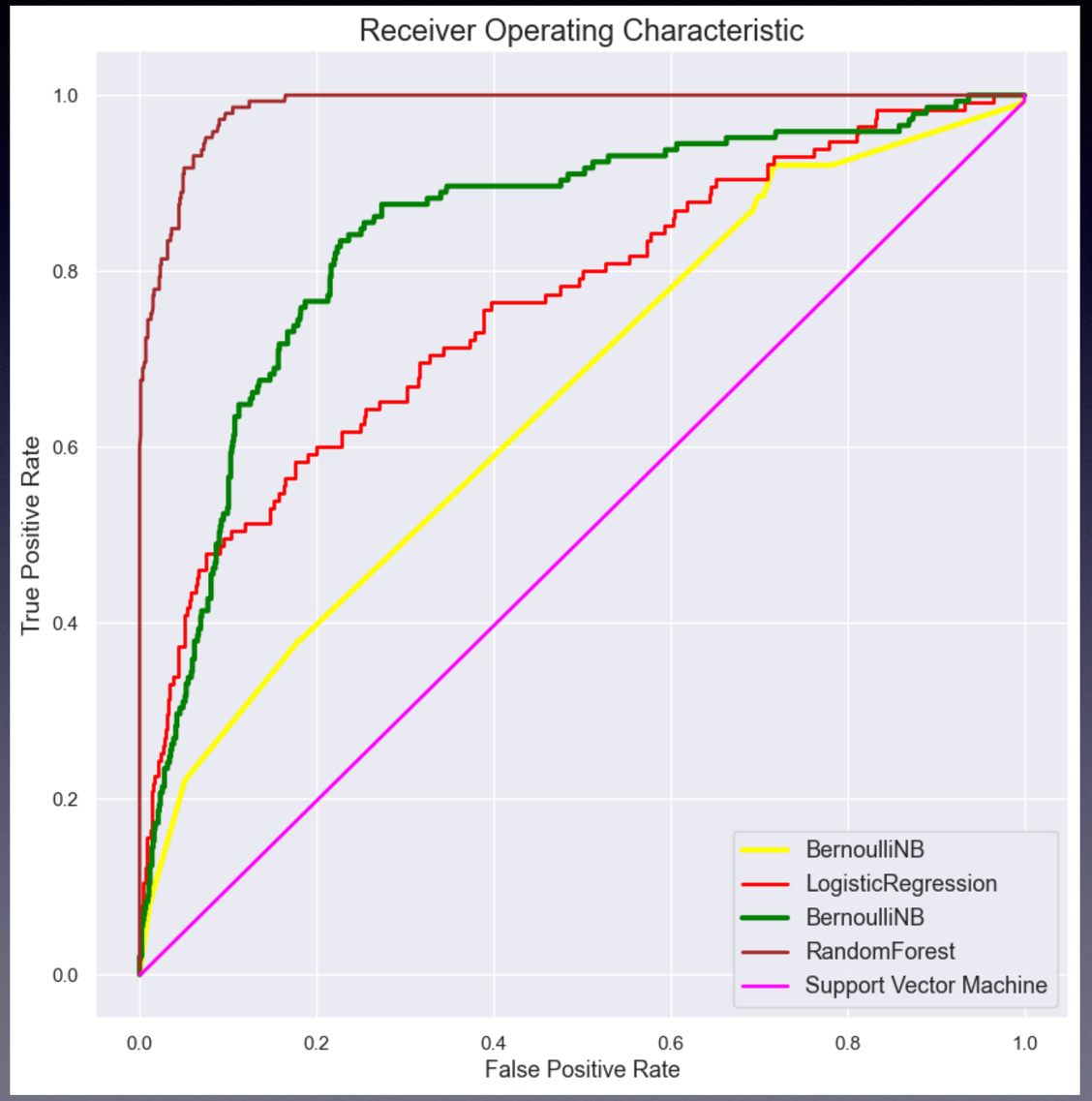
Models Results

- Best Results are given by the Random Forest model

	precision	recall	f1-score	support
0	0.941566	0.998830	0.969353	855.000
1	0.989247	0.634483	0.773109	145.000
accuracy	0.946000	0.946000	0.946000	0.946
macro avg	0.965406	0.816657	0.871231	1000.000
weighted avg	0.948479	0.946000	0.940898	1000.000

Models Results

- The results is confirmed by the AUC score: Random Forest, 0.98



Conclusions

- Data Analysis sees the Random Forest as best performing classifier for this dataset.
- Improvements in the three catagories analized by the model: focusing on offering a better customer service and more favorable rates for day calls.

Further Research

- More models testing
- Parameters manipulation

Q&A

Thanks for the attention!

Contact

mariaantoniettaricci1983@gmail.com