Devin Duran

Russ Steil

**Enhancing Customer Experience: Leveraging Predictive Analytics for Payment Channel Optimization**

**Introduction**

We analyzed a dataset containing customers for an insurance company who have a bill due in the next five days. We created a predictive model for this insurance company to see how likely a customer is to make a service payment call in the next five days. The predictive model we will use is random forests, and we will perform data wrangling and frequently adjust the hyperparameters of the model.

**Data Cleaning**

We imported the packages we needed to load in our dataset containing 130,086 observations and 29 features. Our target variable is ‘Call\_Flag’ which contains values of 0 or 1, representing if a customer made a service payment call or not. We dropped redundant features and take care of missing values. We removed ‘EVENT1\_30\_FLAG’ because all of the values are zero. We removed ‘DATE\_FOR’ since there are only five dates that don’t really seem to match.When cleansing the data, we printed the value counts of the ‘Call\_Flag’ feature only to notice massive imbalance between the two classes. Of the 130,086 customers, 125,322 did not make a service payment call while only 4764 customers made a service payment call.We can combat this using techniques of undersampling and oversampling.

**Model**

We used a random forest model with gini impurity and trained on 100 random forests because it did not show any metric improvement after 100 random forests. The evaluation is on the next page.

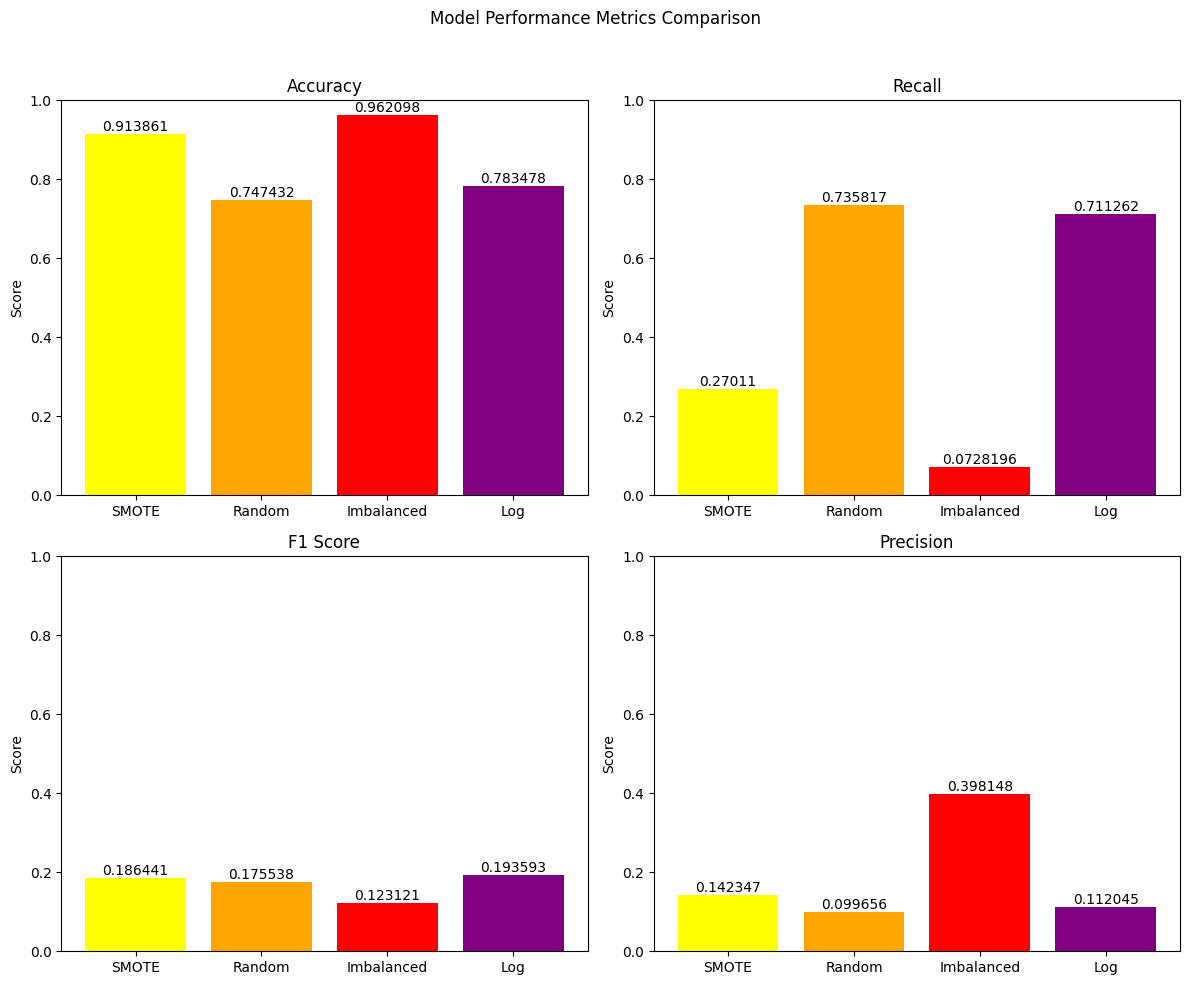
**Observations**

The observations were that the F1 score was low while the accuracy and recall were high for the chosen model: the undersampling model.

**Take Home Message for Customer**

The model needs to be trained with more balanced data. This model can be put into production and will reduce costs somewhat because it will send an email to most of the people it needs to, but if you get more data about when a customer needs to get sent an email it will perform better.

**Model Evaluation**

****