SyriaTel: How to keep the customer churn?

Guofa Shou Self-paced

Outline

- The aim: to provide suggestions about how to keep the customers churn to SyriaTel company
- Data analysis: A binary classification model is built to identify important features closely related to customer churn
- Suggestions

Data

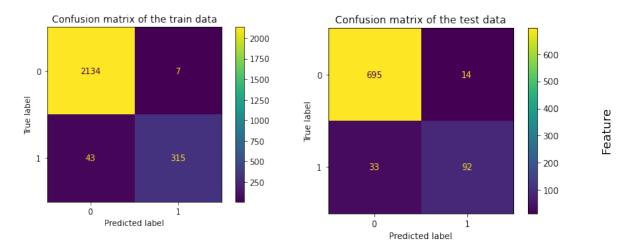
- Data source:
 - bigml_59c28831336c6604c800002a.csv
- Data summary:
 - 20 features
 - 1 target: churn with false or true
 - 3333 customers

RangeIndex: 3333 entries, 0 to 3332 Data columns (total 21 columns):

```
Non-Null Count Dtype
    Column
                            -----
     state
                            3333 non-null
                                            object
    account length
                            3333 non-null
                                            int64
    area code
                            3333 non-null
                                            int64
     phone number
                            3333 non-null
                                            object
    international plan
                            3333 non-null
                                            object
    voice mail plan
                                            object
                            3333 non-null
    number vmail messages
                            3333 non-null
                                            int64
    total day minutes
                            3333 non-null
                                            float64
    total day calls
                                            int64
                            3333 non-null
    total day charge
                                            float64
                            3333 non-null
    total eve minutes
                                            float64
                            3333 non-null
    total eve calls
                            3333 non-null
                                            int64
12 total eve charge
                            3333 non-null
                                            float64
    total night minutes
                            3333 non-null
                                            float64
   total night calls
                            3333 non-null
                                            int64
    total night charge
                            3333 non-null
                                            float64
    total intl minutes
                            3333 non-null
                                            float64
17 total intl calls
                            3333 non-null
                                            int64
    total intl charge
                            3333 non-null
                                            float64
    customer service calls 3333 non-null
                                            int64
20 churn
                            3333 non-null
                                            bool
dtypes: bool(1), float64(8), int64(8), object(4)
```

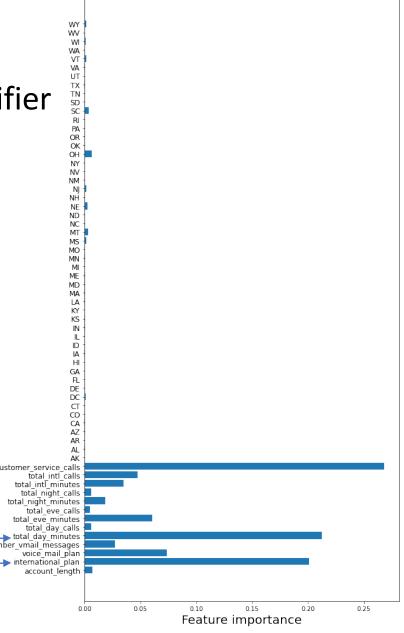
Data Analysis

- Classification model: GradientBoostingClassifier
- Classification performance



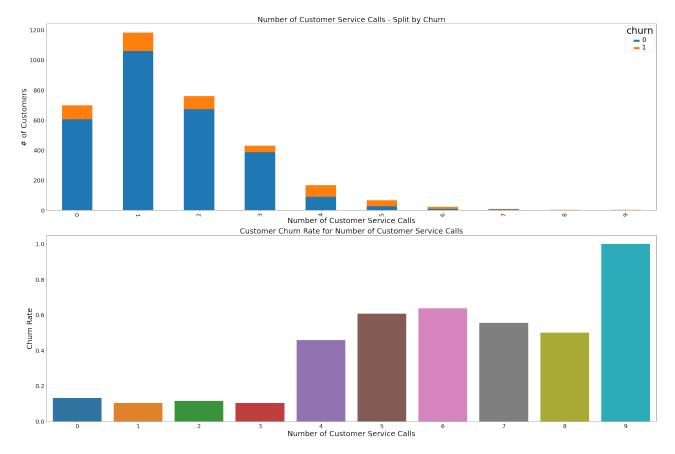


- Customer_service_calls
- Total_day_minutes
- International_plan



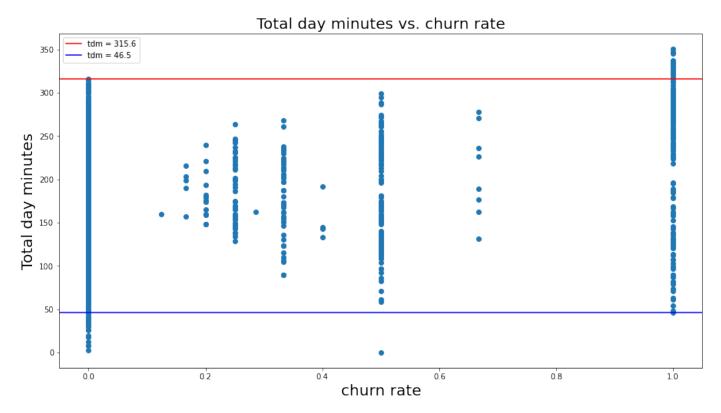
Final Model Feature Importance

Suggestion based on customer service calls



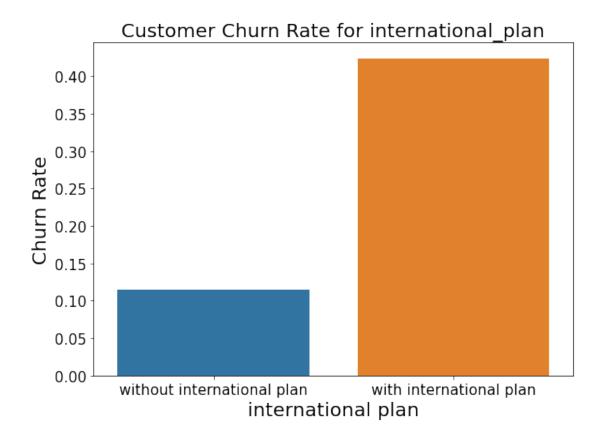
Increase the service calls will keep the customers churn with the service

Suggestion based on total day minutes



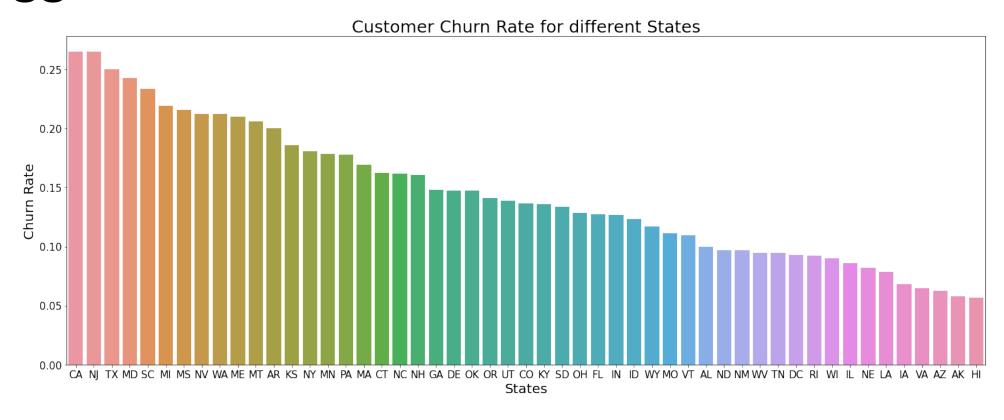
The company needs to work more on the customers with the total day minutes between 46.5 to 316 mins

Suggestion based on international plan



The company could promote the international plan to keep customers churn with the service

Suggestion based on state



The company needs to pay more attention on customers from AZ, AK and HI states