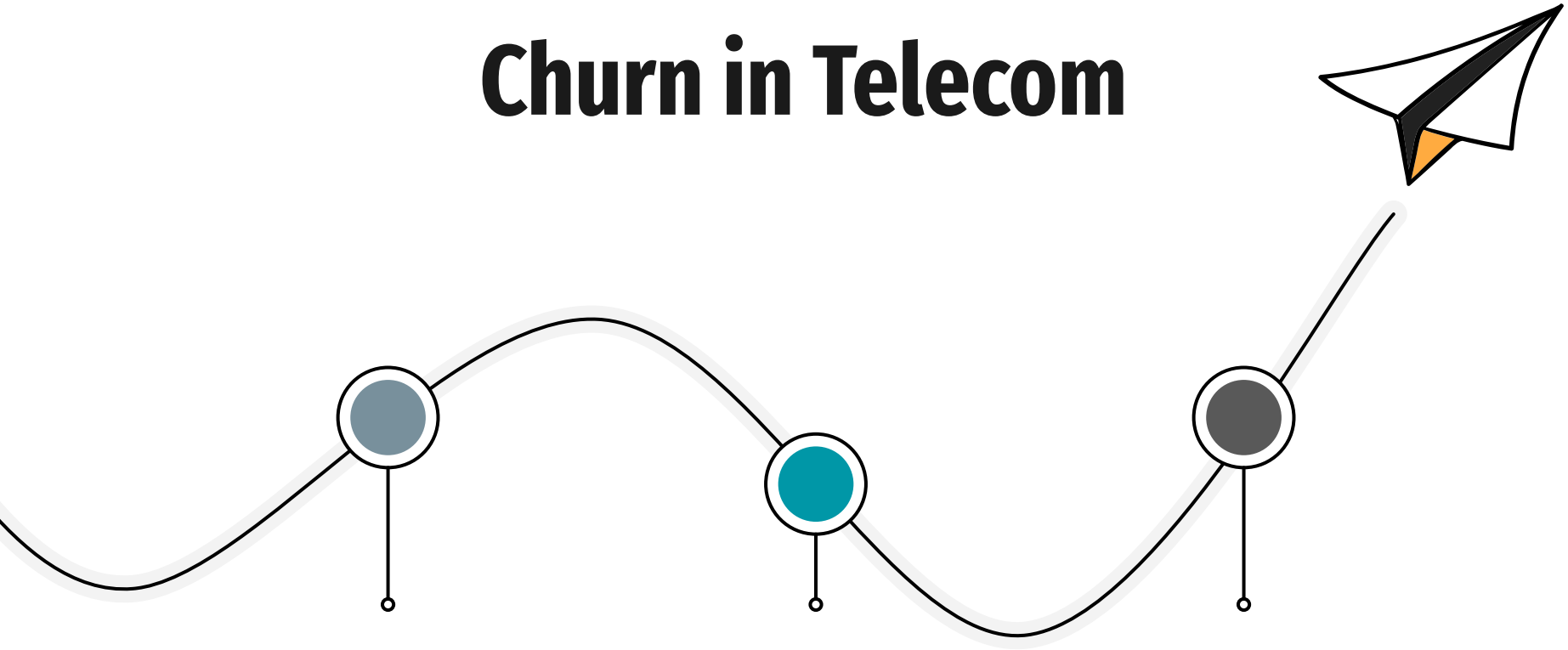
Churn in Telecom



DSI-14 Capstone

Nayara Reis

Agenda


Business Case
Data and Variables
EDA
Models and Evaluation
Conclusions
Future Development

Business Case

Definition:

$$\frac{\text{USERS AT BEGINNING OF PERIOD - USERS AT END OF PERIOD}}{\text{USERS AT BEGINNING OF PERIOD}} = \text{CHURN RATE}$$

Business Case

Potential Audience

Telecom companies

Goal

Predict which customers might Churn and define them as target for retention strategies

Success Metrics

Precision and recall

Data and Variables

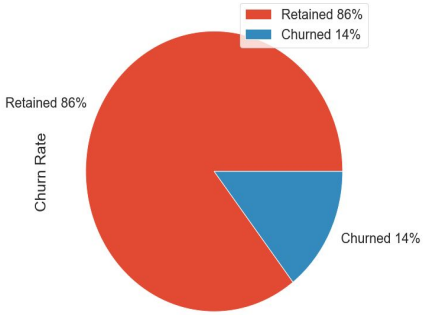
Columns:
State
Account length
Area code
International plan
Voice mail plan
Number vmail messages
Total day minutes
Total day calls
Total day charge
Total eve minutes
Total eve calls
Total eve charge
Total night minutes
Total night calls
Total night charge
Total intl minutes
Total intl calls
Total intl charge
Customer service calls
Churn

Source:
https://bml-data.s3.amazonaws.com/churn-bigml-80.csv

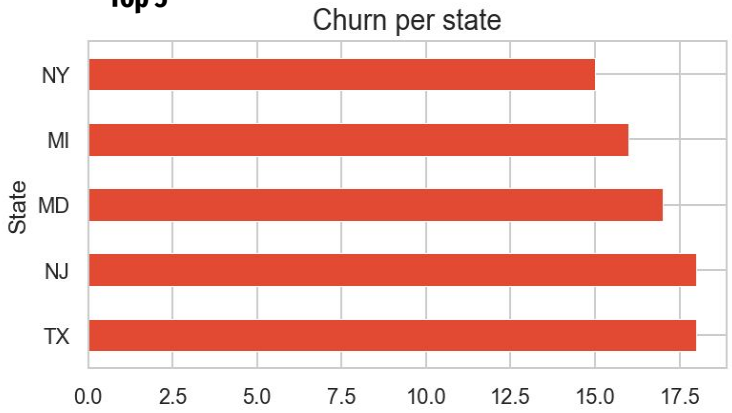
Dataset:
Orange Telecom for USA, 2019

Shape: 3333, 20

EDA

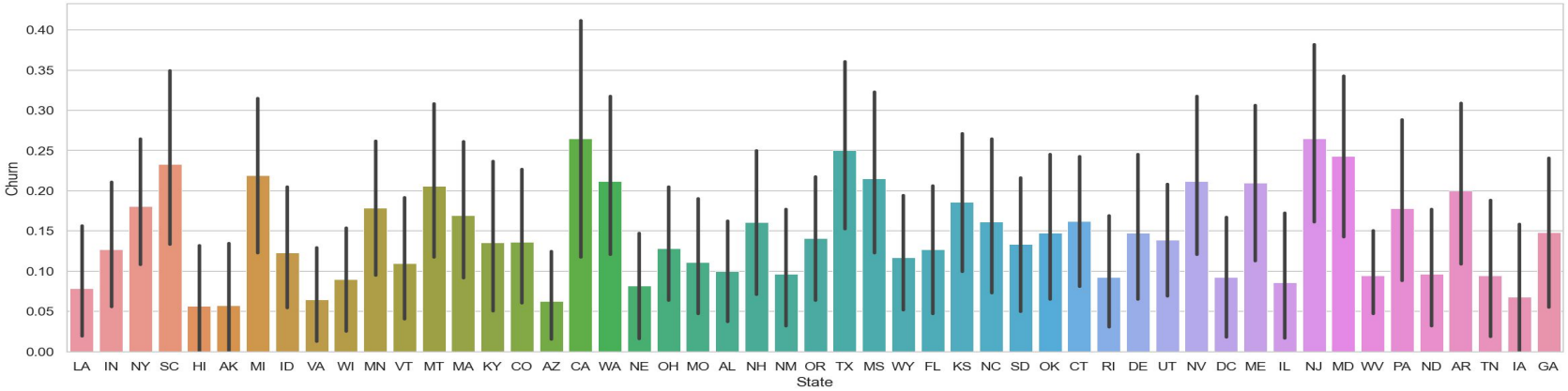


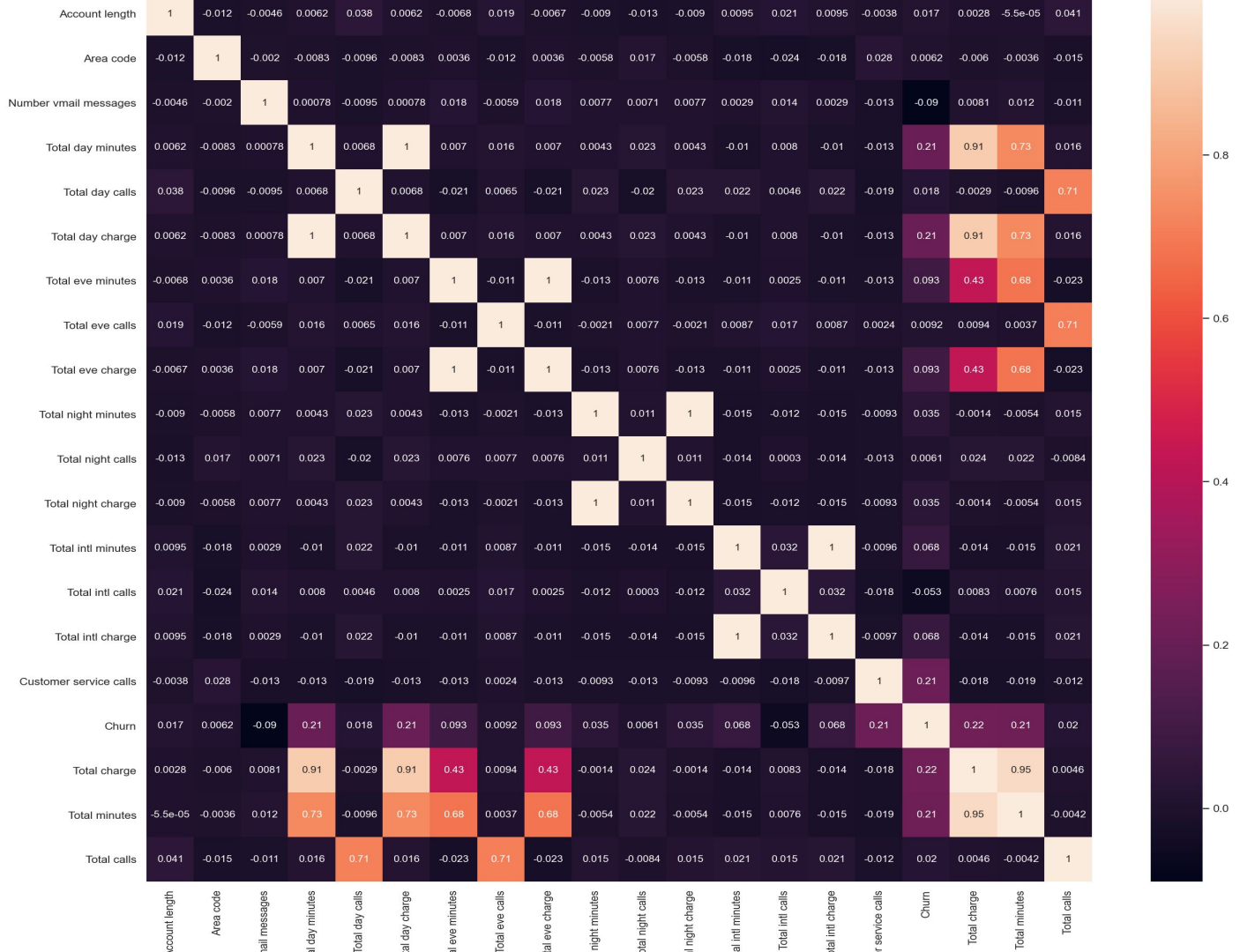
Top 5



Relevant variables

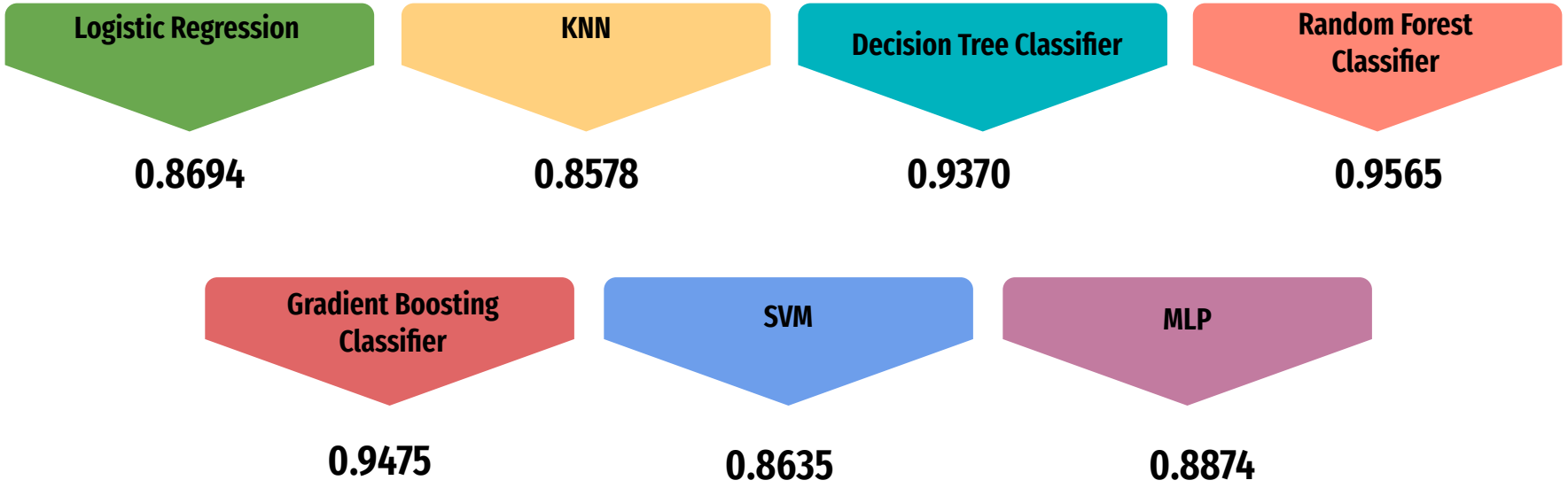
- **Customer Service Calls**
- **Total Charges**
- **Total minutes**




Account length 1 -0.012 -0.0046 0.0062 0.038 0.0062 -0.0068 0.019 -0.0067 -0.009 -0.013 -0.009 0.0095 0.021 0.0095 -0.0038 0.017 0.0028 -5.5e-05 0.041
Area code -0.012 1 -0.002 -0.0083 -0.0096 -0.0083 0.0036 -0.012 0.0036 -0.0058 0.017 -0.0058 -0.018 -0.024 -0.018 0.028 0.0062 -0.006 -0.0036 -0.015
Number vmail messages -0.0046 -0.002 1 0.00078 -0.0095 0.00078 0.018 -0.0059 0.018 0.0077 0.0071 0.0077 0.0029 0.014 0.0029 -0.013 -0.09 0.0081 0.012 -0.011
Total day minutes 0.0062 -0.0083 0.00078 1 0.0068 1 0.007 0.016 0.007 0.0043 0.023 0.0043 -0.01 0.008 -0.01 -0.013 0.21 0.91 0.73 0.016
Total day calls 0.038 -0.0096 -0.0095 0.0068 1 0.0068 -0.021 0.0065 -0.021 0.023 -0.02 0.023 0.022 0.0046 0.022 -0.019 0.018 -0.0029 -0.0096 0.71
Total day charge 0.0062 -0.0083 0.00078 1 0.0068 1 0.007 0.016 0.007 0.0043 0.023 0.0043 -0.01 0.008 -0.01 -0.013 0.21 0.91 0.73 0.016
Total eve minutes -0.0068 0.0036 0.018 0.007 -0.021 0.007 1 -0.011 1 -0.013 0.0076 -0.013 -0.011 0.0025 -0.011 -0.013 0.093 0.43 0.68 -0.023
Total eve calls 0.019 -0.012 -0.0059 0.016 0.0065 0.016 -0.011 1 -0.011 -0.0021 0.0077 -0.0021 0.0087 0.017 0.0087 0.0024 0.0092 0.0094 0.0037 0.71
Total eve charge -0.0067 0.0036 0.018 0.007 -0.021 0.007 1 -0.011 1 -0.013 0.0076 -0.013 -0.011 0.0025 -0.011 -0.013 0.093 0.43 0.68 -0.023
Total night minutes -0.009 -0.0058 0.0077 0.0043 0.023 0.0043 -0.013 -0.0021 -0.013 1 0.011 1 -0.015 -0.012 -0.015 -0.0093 0.035 -0.0014 -0.0054 0.015
Total night calls -0.013 0.017 0.0071 0.023 -0.02 0.023 0.0076 0.0077 0.0076 0.011 1 0.011 -0.014 0.0003 -0.014 -0.013 0.0061 0.024 0.022 -0.0084
Total night charge -0.009 -0.0058 0.0077 0.0043 0.023 0.0043 -0.013 -0.0021 -0.013 1 0.011 1 -0.015 -0.012 -0.015 -0.0093 0.035 -0.0014 -0.0054 0.015
Total intl minutes 0.0095 -0.018 0.0029 -0.01 0.022 -0.01 -0.011 0.0087 -0.011 -0.015 -0.014 -0.015 1 0.032 1 -0.0096 0.068 -0.014 -0.015 0.021
Total intl calls 0.021 -0.024 0.014 0.008 0.0046 0.008 0.0025 0.017 0.0025 -0.012 0.0003 -0.012 0.032 1 0.032 -0.018 -0.053 0.0083 0.0076 0.015
Total intl charge 0.0095 -0.018 0.0029 -0.01 0.022 -0.01 -0.011 0.0087 -0.011 -0.015 -0.014 -0.015 1 0.032 1 -0.0097 0.068 -0.014 -0.015 0.021
Customer service calls -0.0038 0.028 -0.013 -0.013 -0.019 -0.013 -0.013 0.0024 -0.013 -0.0093 -0.013 -0.0093 -0.0096 -0.018 -0.0097 1 0.21 -0.018 -0.019 -0.012
Churn 0.017 0.0062 -0.09 0.21 0.018 0.21 0.093 0.0092 0.093 0.035 0.0061 0.035 0.068 -0.053 0.068 0.21 1 0.22 0.21 0.02
Total charge 0.0028 -0.006 0.0081 0.91 -0.0029 0.91 0.43 0.0094 0.43 -0.0014 0.024 -0.0014 -0.014 0.0083 -0.014 -0.018 0.22 1 0.95 0.0046
Total minutes -5.5e-05 -0.0036 0.012 0.73 -0.0096 0.73 0.68 0.0037 0.68 -0.0054 0.022 -0.0054 -0.015 0.0076 -0.015 -0.019 0.21 0.95 1 -0.0042
Total calls 0.041 -0.015 -0.011 0.016 0.71 0.016 -0.023 0.71 -0.023 0.015 -0.0084 0.015 0.021 0.015 0.021 -0.012 0.02 0.0046 -0.0042 1
Account length
Area code
Number vmail messages
Total day minutes
Total day calls
Total day charge
Total eve minutes
Total eve calls
Total eve charge
Total night minutes
Total night calls
Total night charge
Total intl minutes
Total intl calls
Total intl charge
Customer service calls
Churn
Total charge
Total minutes
Total calls
0.8
0.6
0.4
0.2
0.0

Models

Baseline: 0.8551



Random Forest Classifier

RF

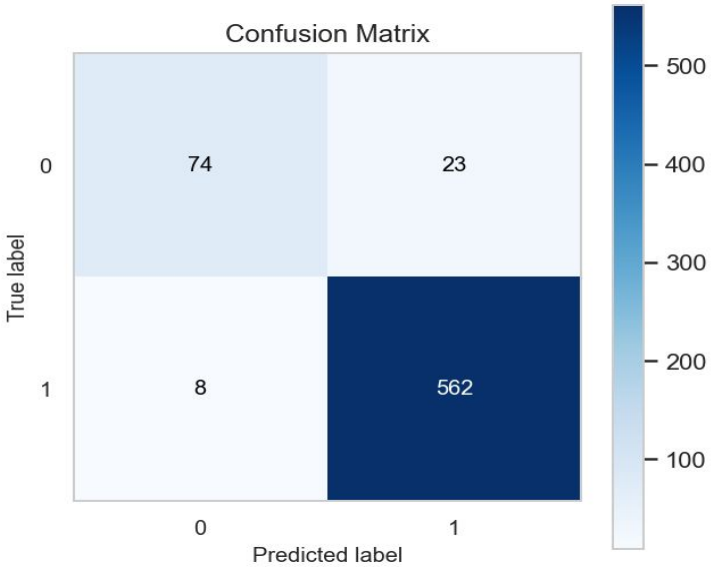
class_weight = 'balance'

scoring = 'recall'

Cross-validated score = 0.9565

Should I stay or should I churn?

Test set



Classification report

Test set	Precision	Recall	F1-scores	Support
0	0.90	0.76	0.83	97
1	0.96	0.99	0.97	570
Accuracy			0.95	667
Macro Avg	0.91	0.85	0.88	667
Weighted avg	0.94	0.94	0.94	667

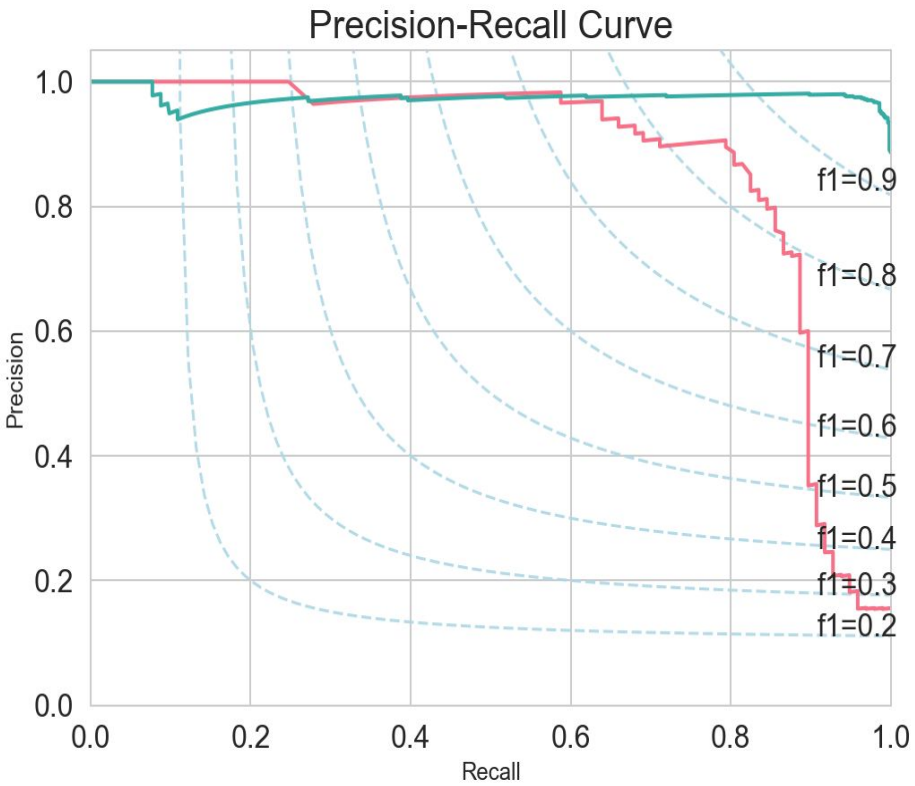
Summary

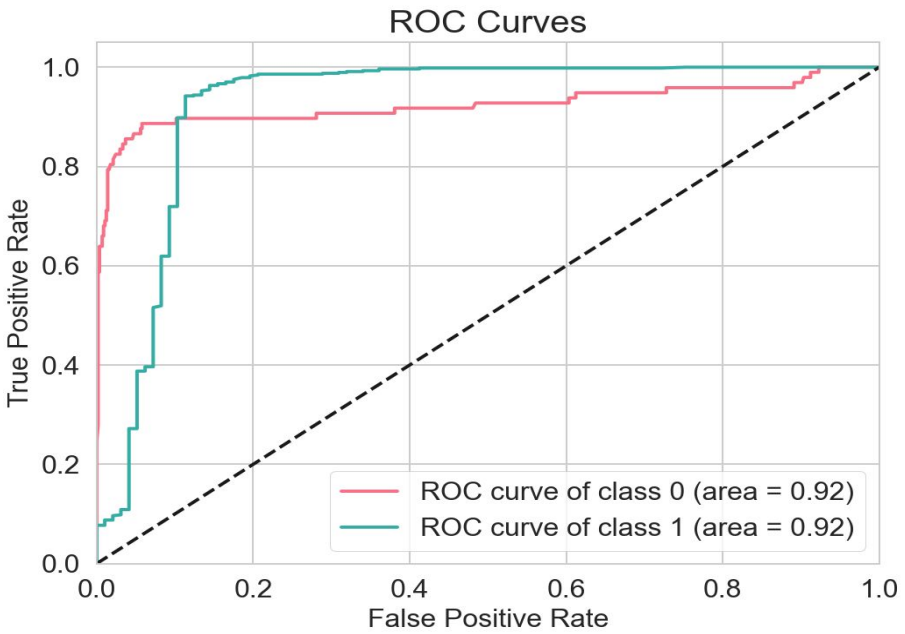
tp: 74
fp: 8
tn: 562
fn: 23
Number of correct classifications: 636
Number of classification errors: 31

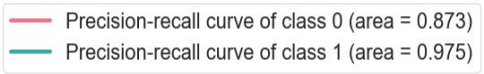
Classes

0: Churn
1: Retained

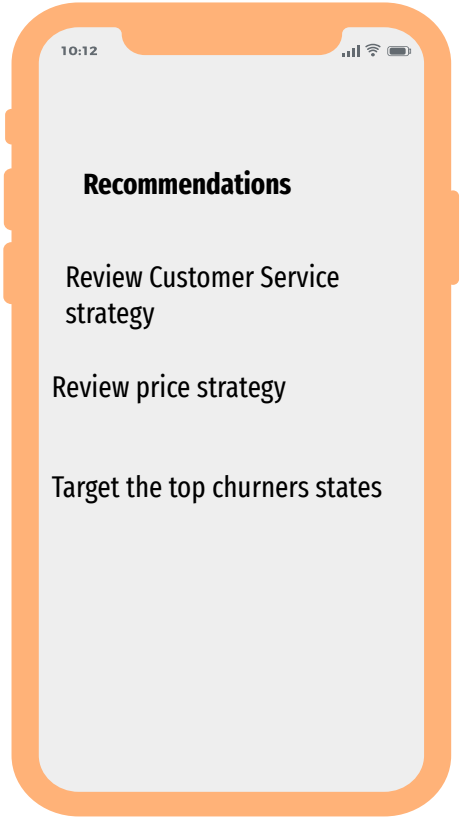
Precision x Recall and ROC Curves

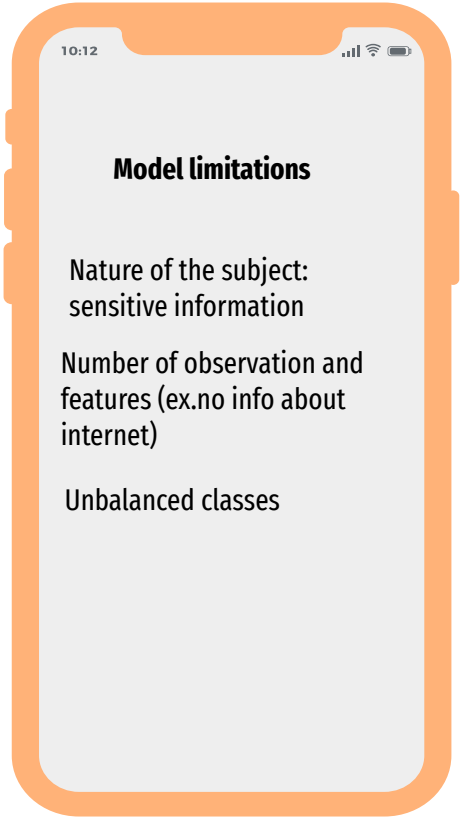






Conclusions





Thank you