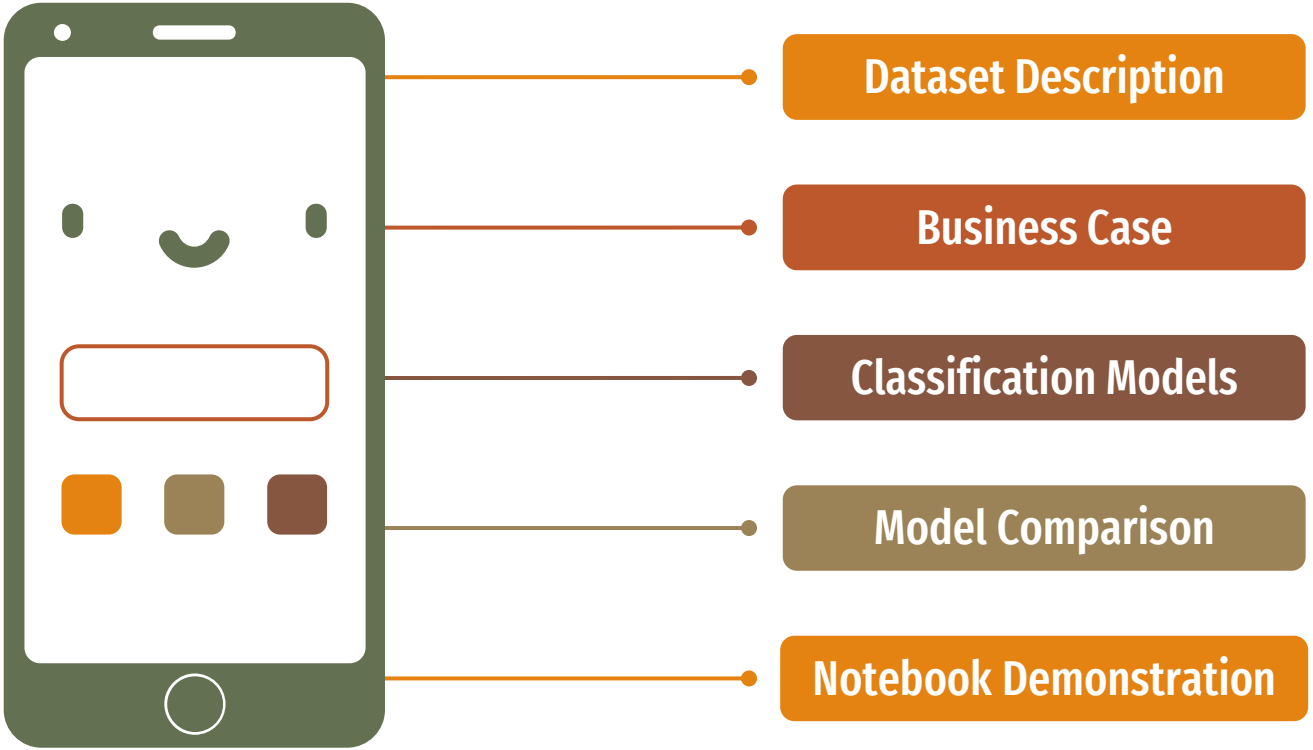
ADMN MOBILE SUBSCRIPTION CHURN PREDICTION

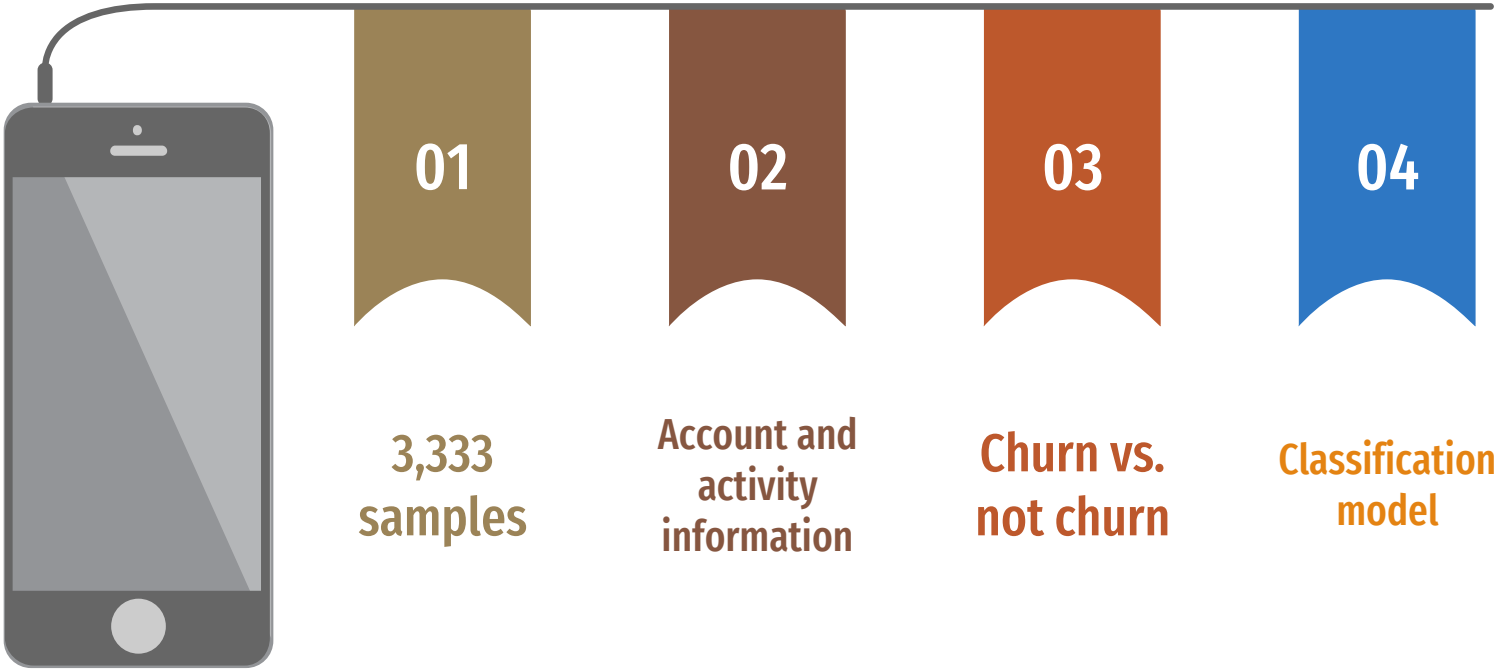
ADMN5016: APPLIED ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

MARIA ISABEL DIO GRAU
APRIL 20, 2022

Outline


Dataset Description
Business Case
Classification Models
Model Comparison
Notebook Demonstration

Dataset Description



Business Case

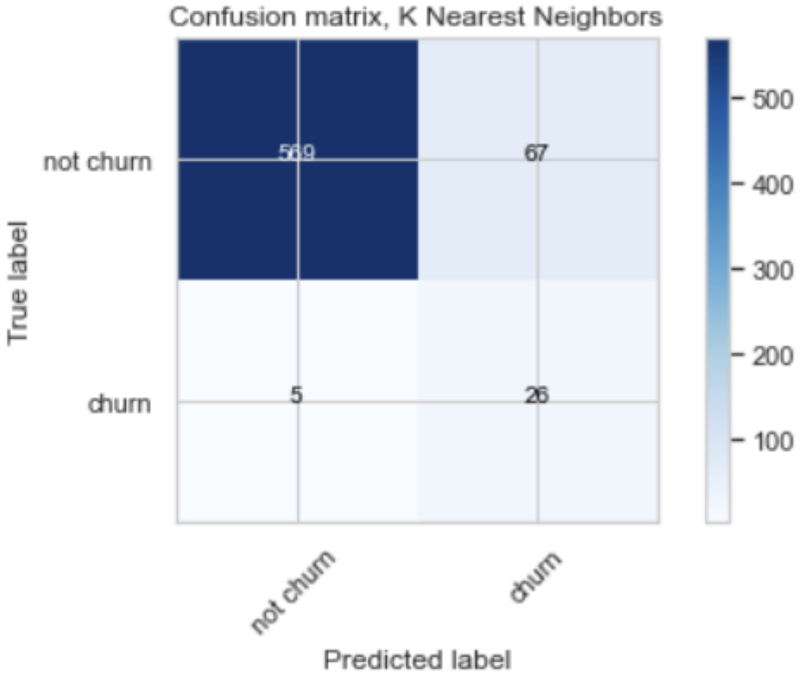
Costs less to retain customers than to acquire new customers
$425 to acquire a new customer*

If a customer is predicted as likely to churn, the company can take corrective actions to improve its services.

Improve brand loyalty and strengthen brand reputation

**US$340, T-Mobile 2017 data*

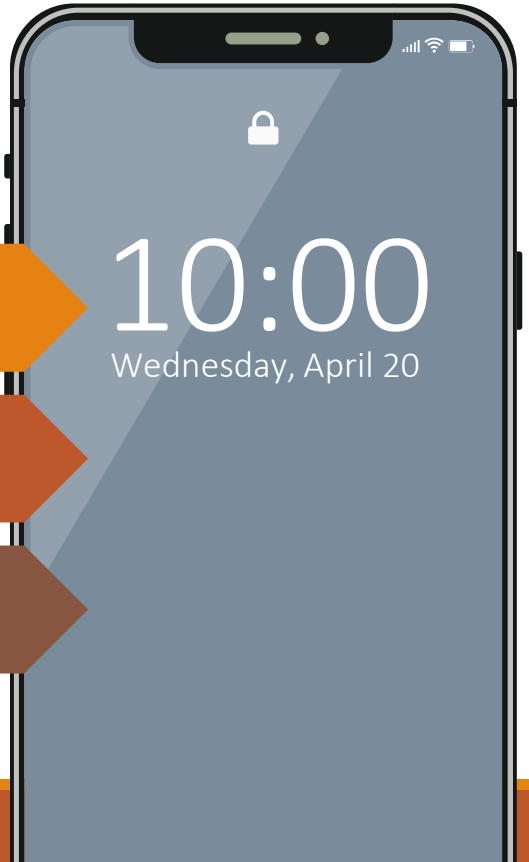
K Nearest Neighbors



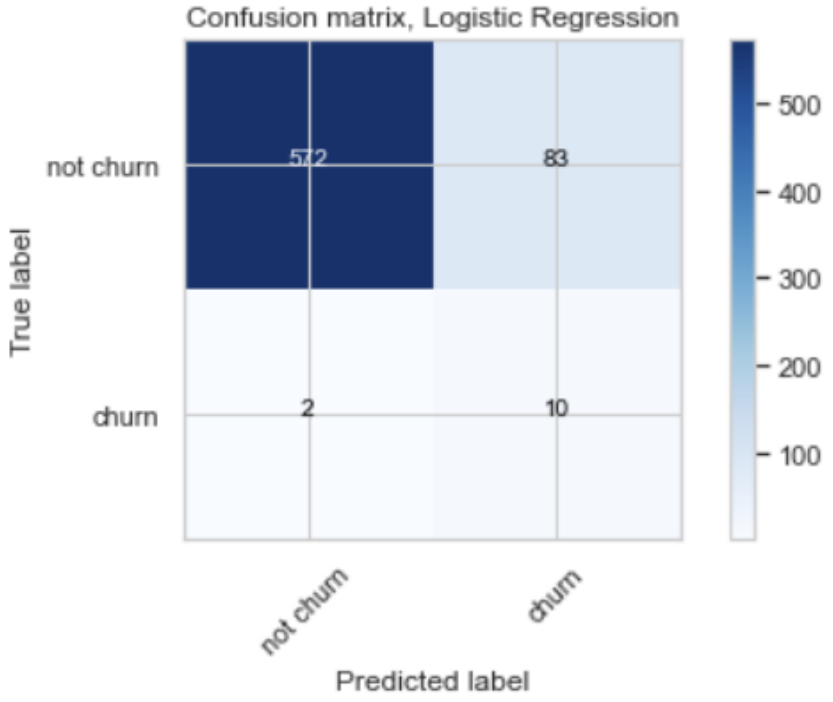
Accuracy: 89.2%

Recall: 28.0%

Precision: 83.9%



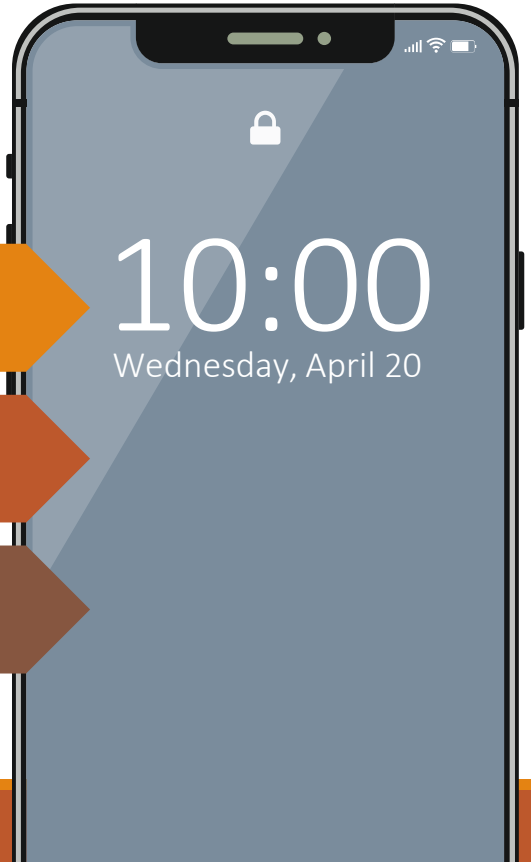
Logistic Regression



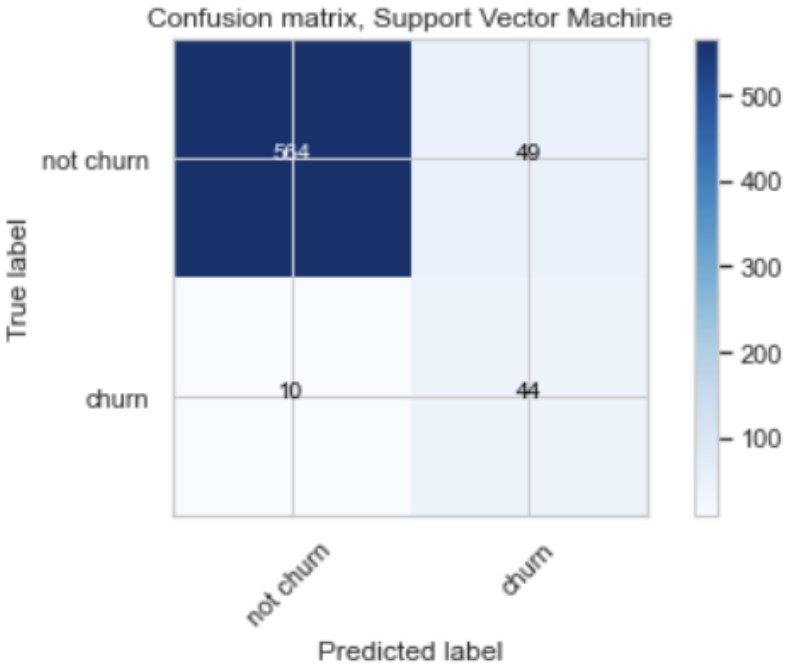
Accuracy: 87.3%

Recall: 10.8%

Precision: 83.3%



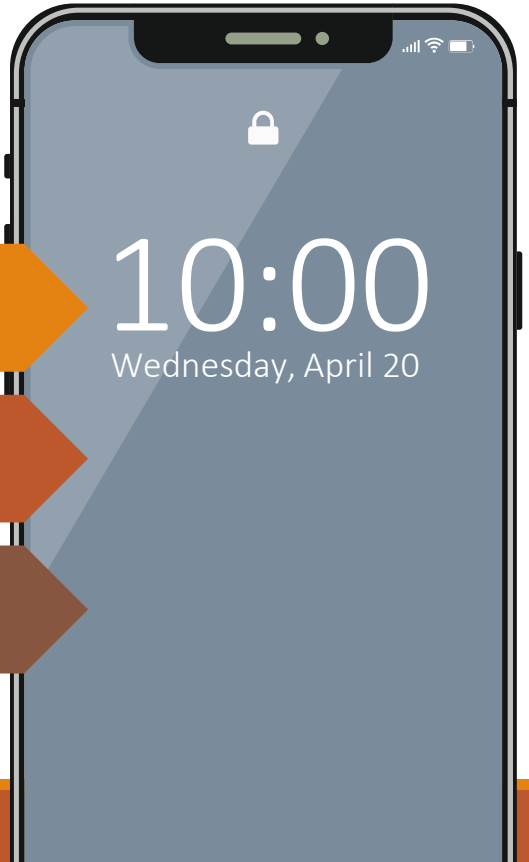
Support Vector Machine



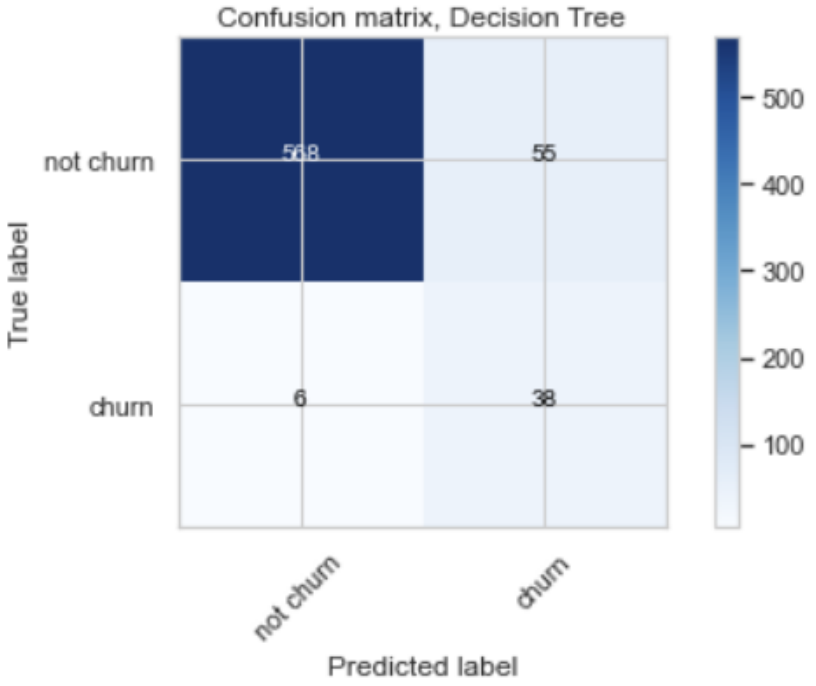
Accuracy: 91.2%

Recall: 47.3%

Precision: 81.5%



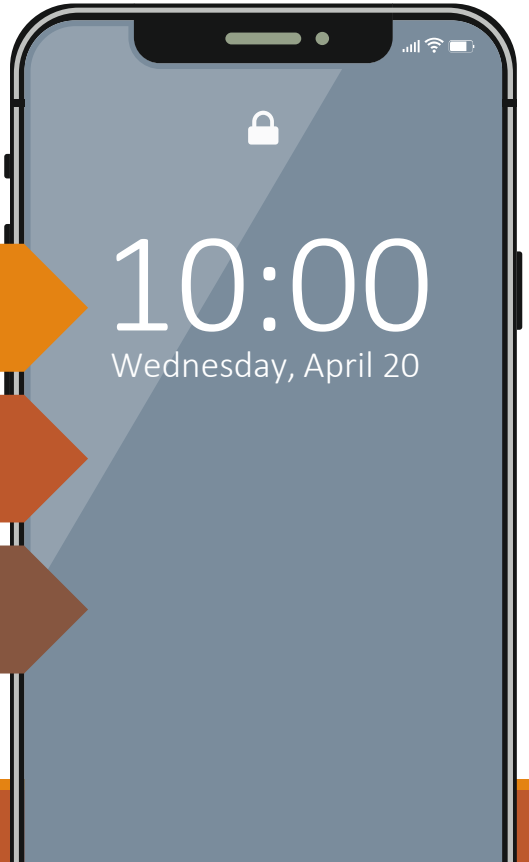
Decision Tree



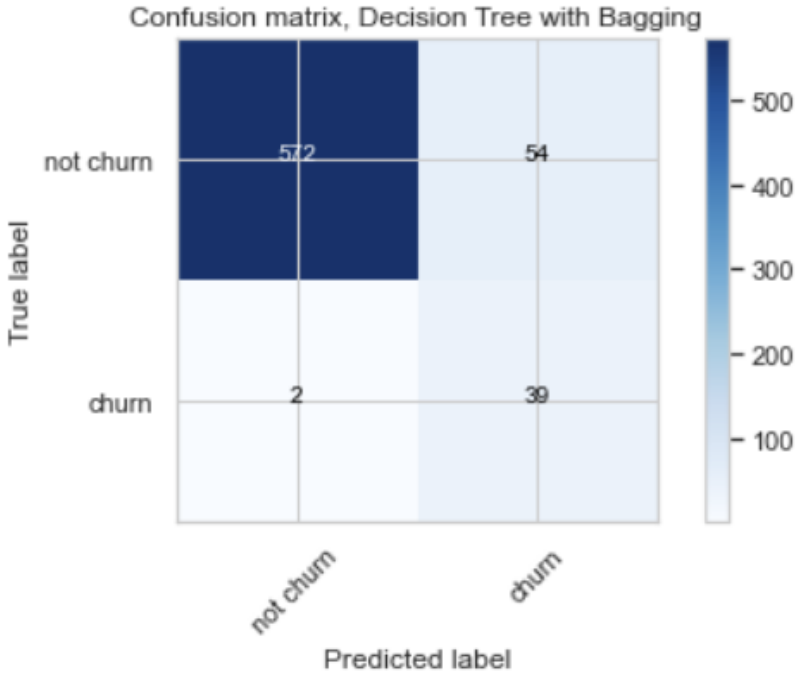
Accuracy: 90.9%

Recall: 40.9%

Precision: 86.4%



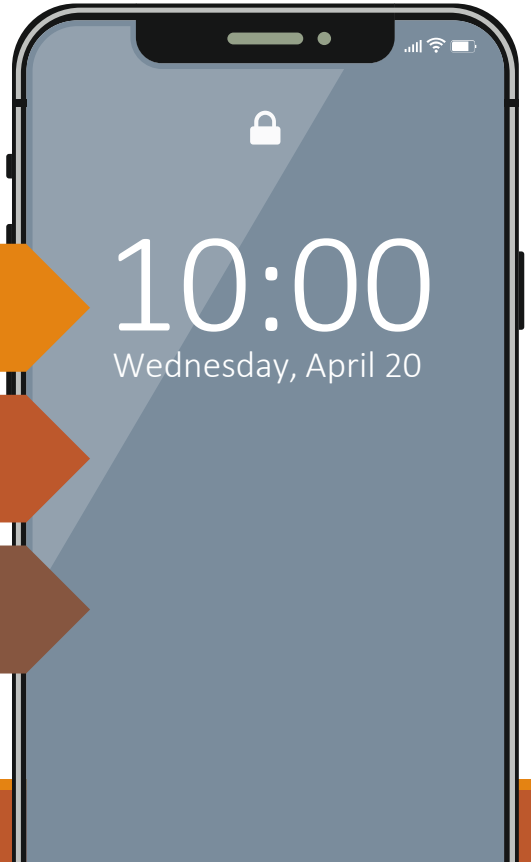
Decision Tree with Bagging



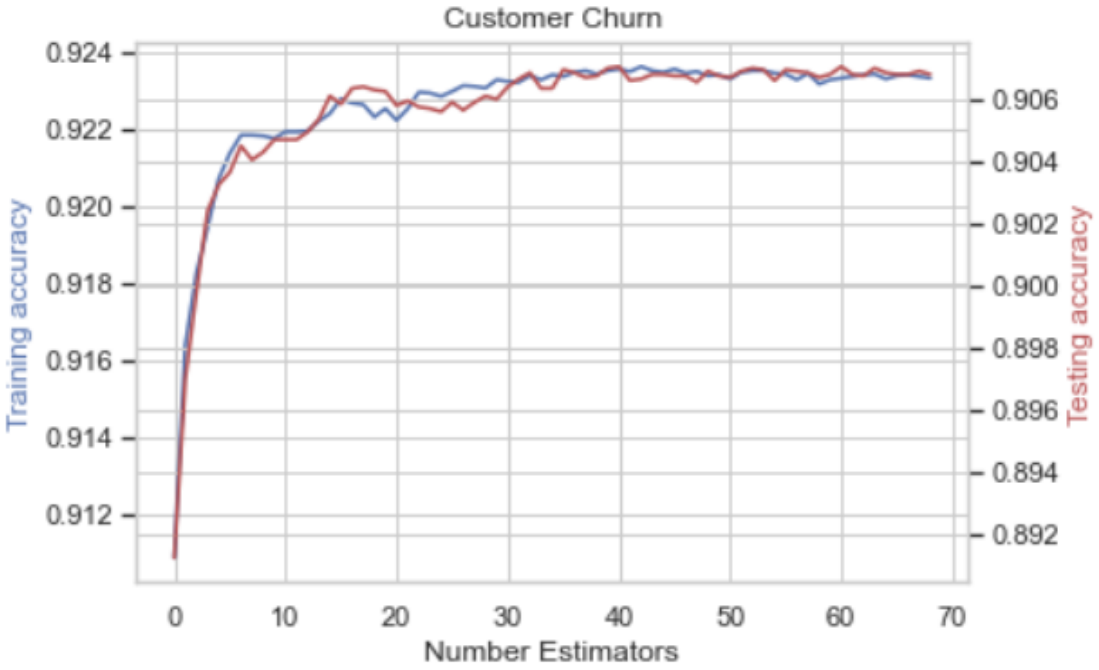
Accuracy: 91.6%

Recall: 41.9%

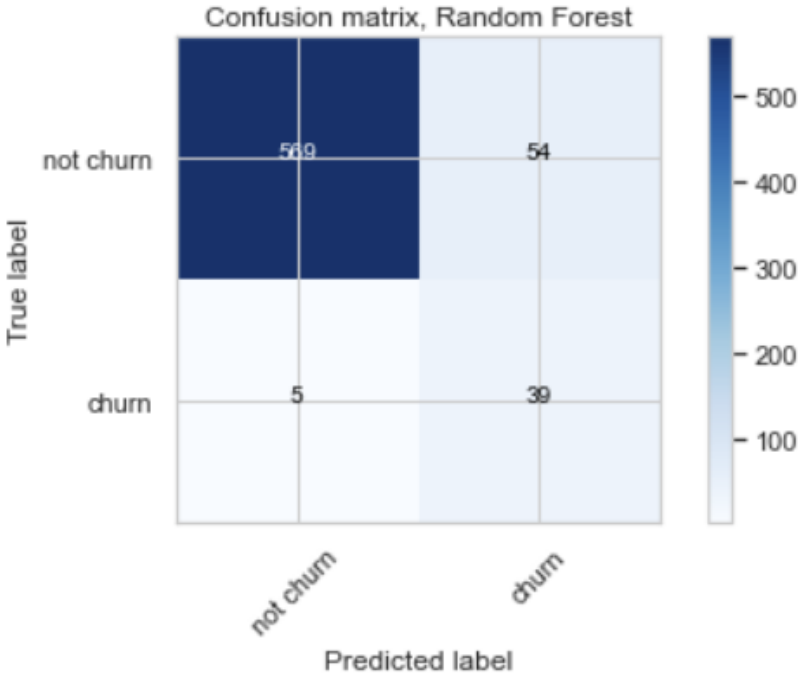
Precision: 95.1%



Decision Tree with Bagging



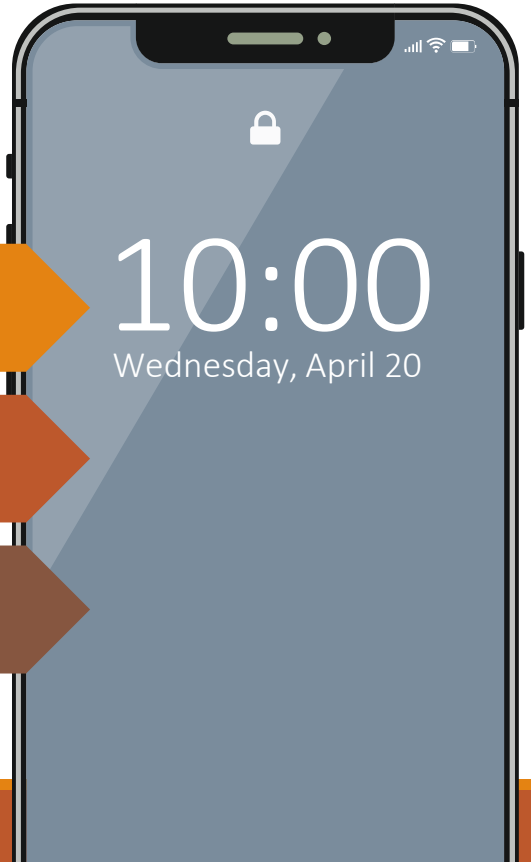
Random Forest



Accuracy: 91.2%

Recall: 41.9%

Precision: 88.6%



Model Comparison

	accuracy	precision	recall	F1
KNN	0.892054	0.838710	0.279570	0.419355
Logistic Regression	0.872564	0.833333	0.107527	0.190476
SVM	0.911544	0.814815	0.473118	0.598639
Decision Tree	0.908546	0.863636	0.408602	0.554745
Decision Tree with Bagging	0.916042	0.951220	0.419355	0.582090
Random Forest	0.911544	0.886364	0.419355	0.569343

Model Comparison: Further Analysis

RECALL

If the churn rate were higher, then recall would be a better metric to evaluate the models

Model Comparison

	accuracy	precision	recall	F1
KNN	0.892054	0.838710	0.279570	0.419355
Logistic Regression	0.872564	0.833333	0.107527	0.190476
SVM	0.911544	0.814815	0.473118	0.598639
Decision Tree	0.908546	0.863636	0.408602	0.554745
Decision Tree with Bagging	0.916042	0.951220	0.419355	0.582090
Random Forest	0.911544	0.886364	0.419355	0.569343

Model Comparison: Further Analysis

RECALL

If the churn rate were higher, then recall would be a better metric to evaluate the models

DIMENSION REDUCTION

Reduce complexity

BOOSTING

Reduce bias

BAGGING

Bootstrap all models

ADMN MOBILE SUBSCRIPTION CHURN PREDICTION

THANK YOU