Bank Customers and Churn

Prediction of whether bank's customers

Keep their account

Or

They close their account

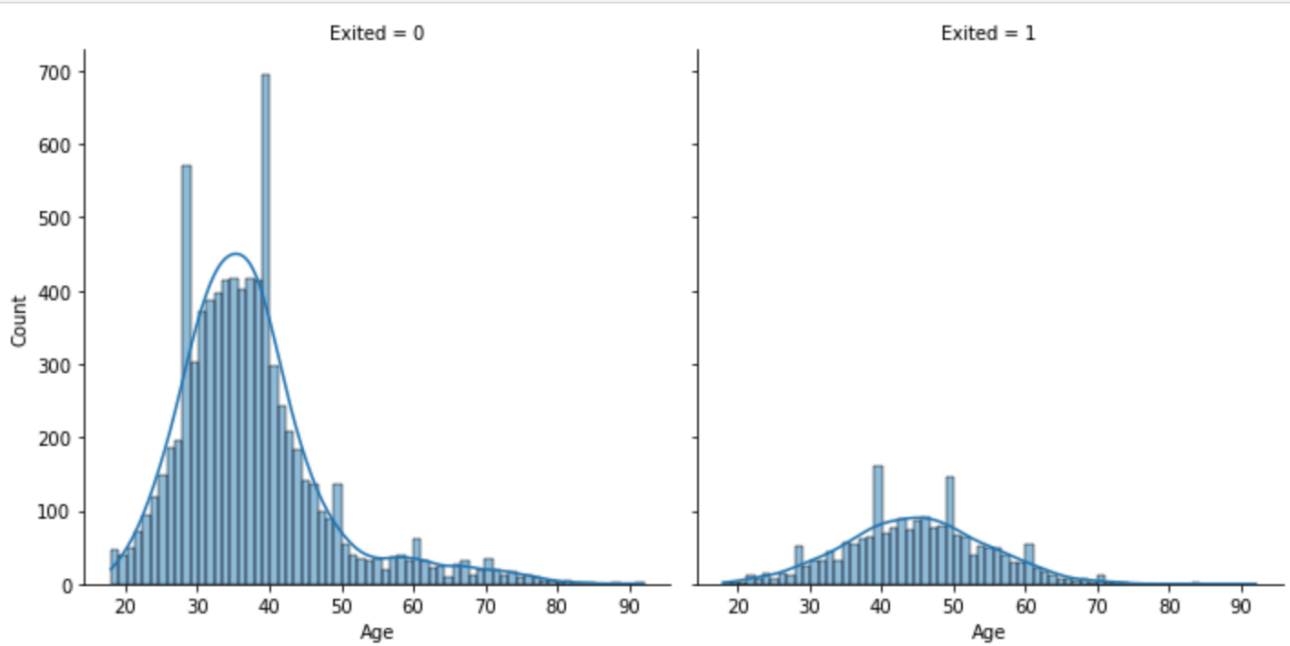
By
Alireza Karimi

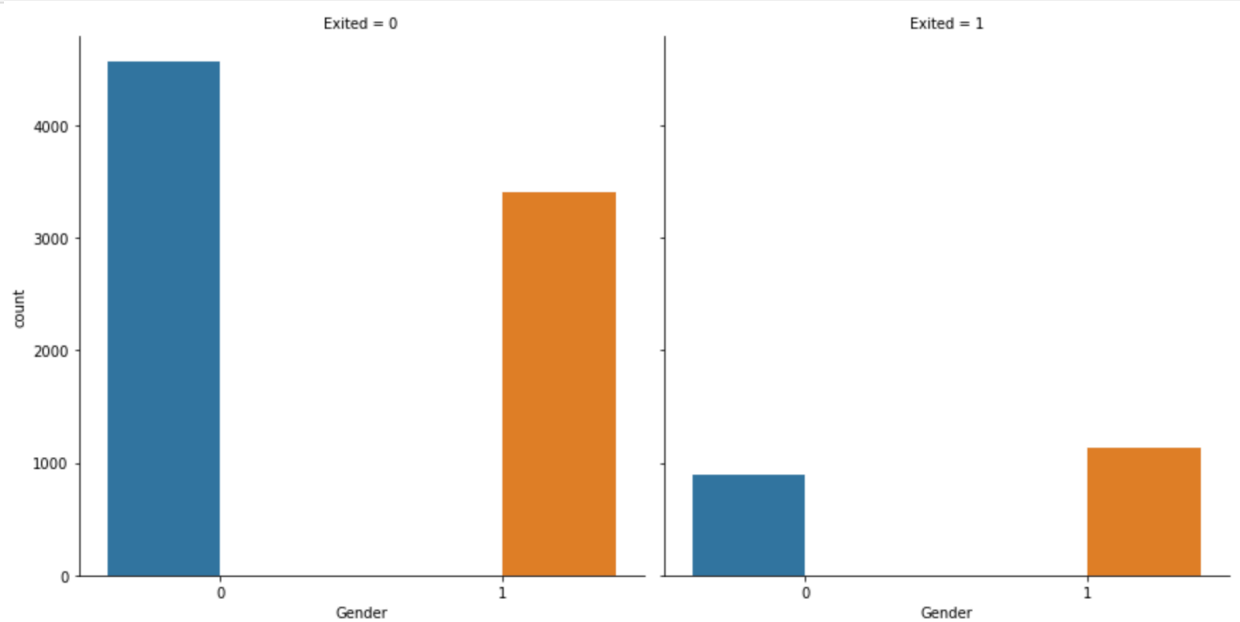
Data Description

- Downloaded from Kaggle
- 10000 records of customer's bank
- France, Germany and Spain
- 3 most correlated features
- Age, Balance and Region of banks

Business problem

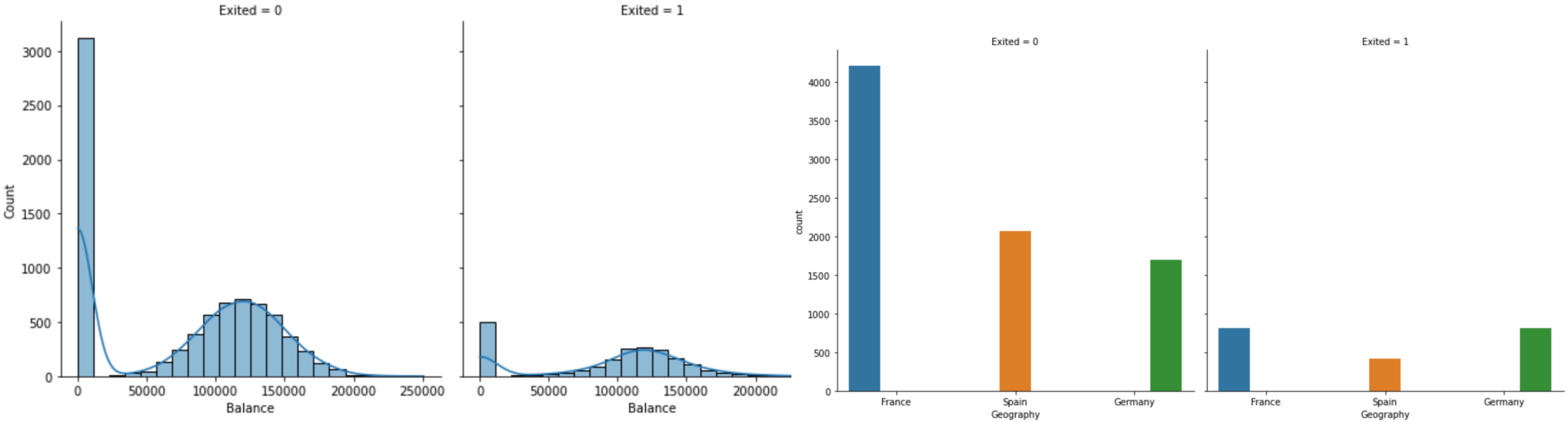
- Customer churn
- Identifying segments of these customers
- Price for acquiring a new customer
- Prevent it by special offers
- Make a model to predict customer’s behavior





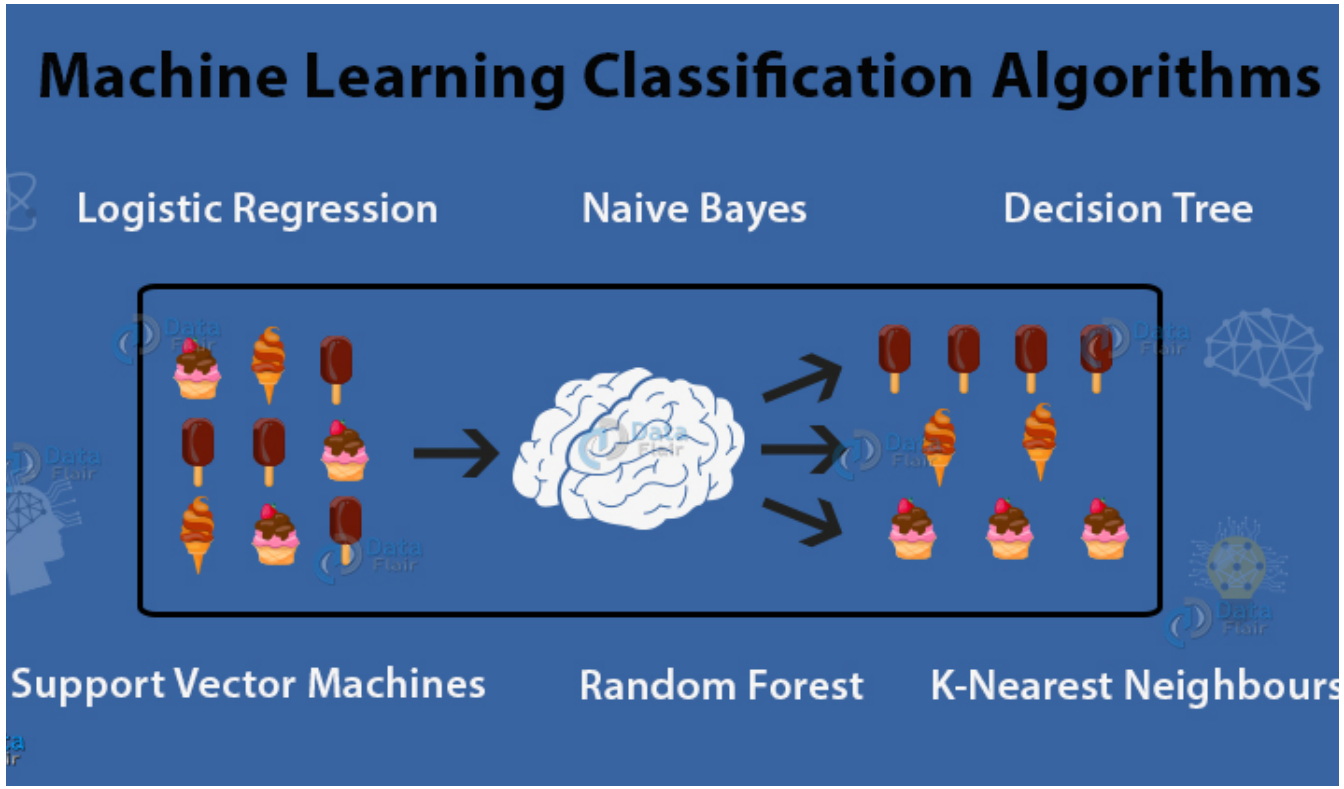
Class 0 = Stayed
Class 1 = Exited

Gender 0 = Male
Gender 1 = Female



Class 0 = Stayed
Class 1 = Exited

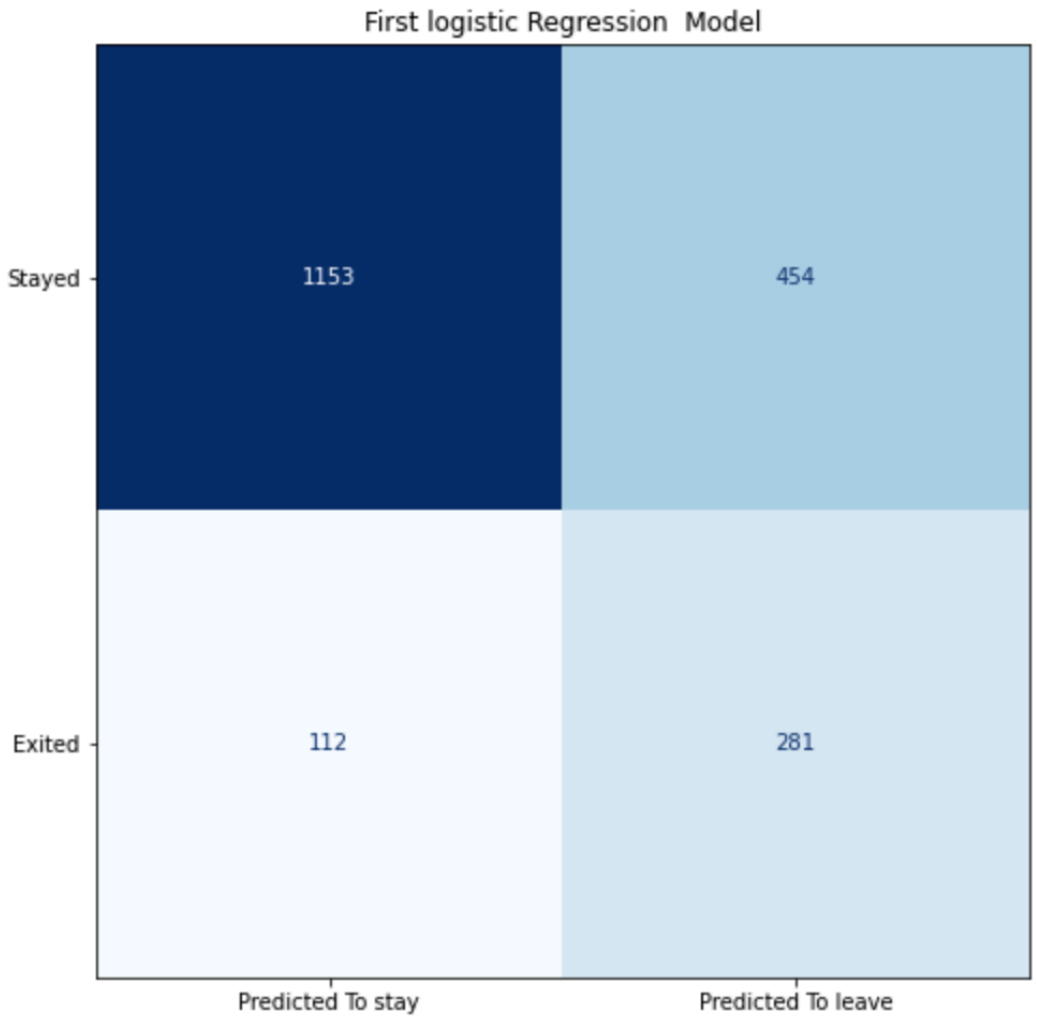
Modeling

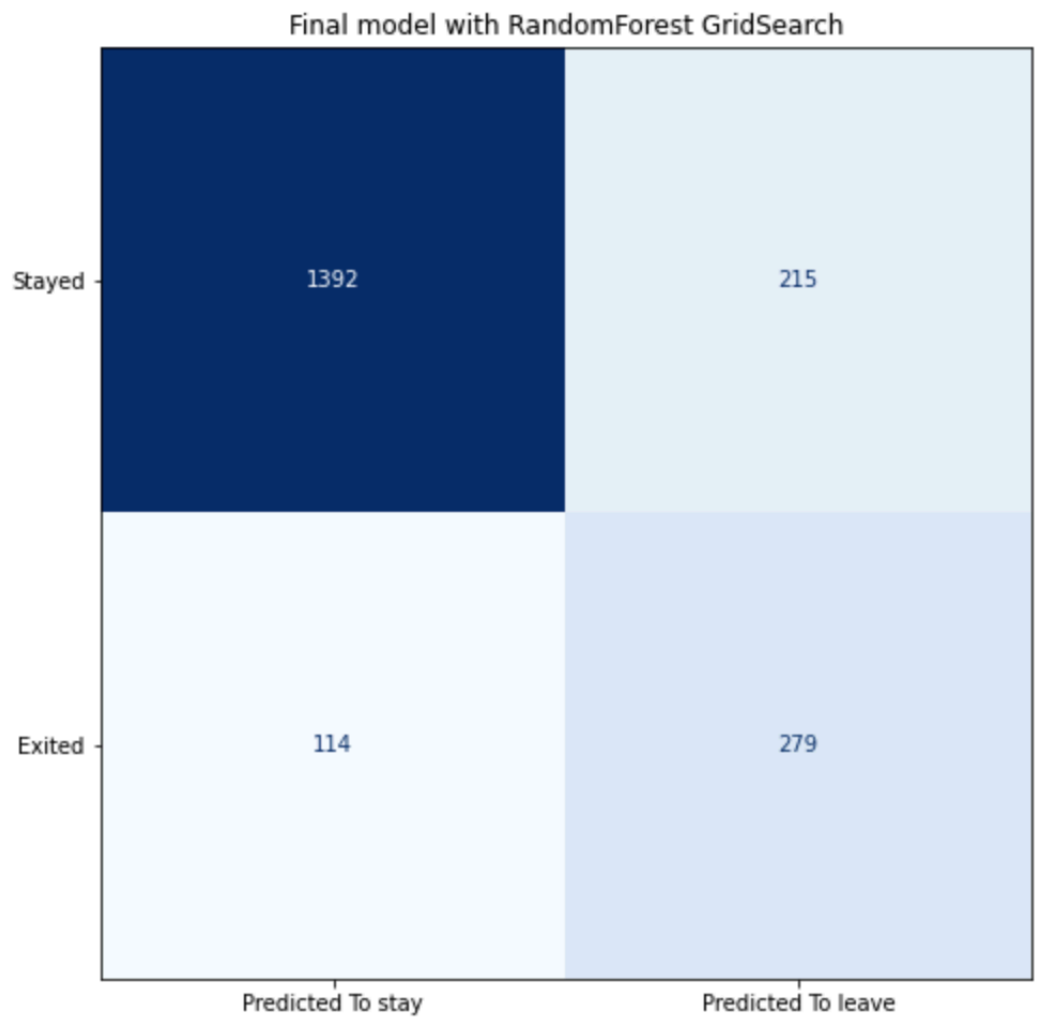


- Using Classification Algorithms
- Making models
- Focusing on f1 score as major metric

First simple model to Final model with Pipeline

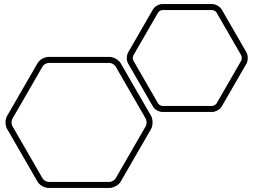
- LogisticRegresion, because it's simple, fast and easy to interpretation, F1 score ≈ 0.49
- KNeighborsClassifier with the average F1 score ≈ 0.48
- DecisionTreeClassifier and the average F1 score ≈ 0.50
- RandomForestClassifier F1 score ≈ 0.59
- RandomForestClassifier with GridSearch F1 score ≈ 0.63
- XGBClassifier with the average F1 score ≈ 0.60


First logistic Regression Model
Stayed
Exited
1153
454
112
281
Predicted To stay
Predicted To leave

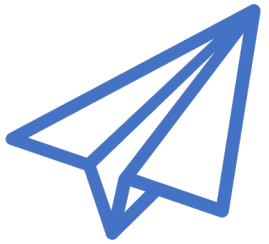

Final model with RandomForest GridSearch
Stayed
Exited
1392
215
114
279
Predicted To stay
Predicted To leave

Next Step and Recommendations

- keep track of age on customers 40 to 50 at higher risk
- Customers with balance between 100,000 and 150,000
- It is most likely to leave bank account from Germany



Thanks for your time



Alireza Karimi

GitHub: @aligodspeed

Email: alirezakarimi80@gmail.com