Client Attrition Analysis

Troy Reynolds

J.P. Morgan

December 17, 2021

BANKMEGA

Goal for Client Attrition Analysis

Business priorities

- Create a client churn model to identify clients that are likely to terminate services.
- The bank can use that model to focus customer retention strategies on clients with high attrition risk.

Error Metric

- Unbalanced Data: Recall
  - Assumption for recall: The opportunity loss from false negative outweighs the profit loss from a false positive.
  - Reasoning: The attrition of a false negative client eliminates all profits from the client while the promotions sent to a false positive client will reduce some profits.

Agenda

01. Process

Describe the dataset and changes the data undergoes as it flows through the pipeline.

02. Insights

Present discoveries found during Exploratory Data Analysis.

03. Models

Understand the models and reasoning for their candidacy in the model selection.

04. Results

Compare the cross-validation results of various metrics to select the best model.

05. Solutions

Introduce recommendations that work in tandem with the customer attrition model.


Process

Feature Characteristics

Column	Data type	Description
RowNumber	Integer	10K customers
CustomerID	Integer	Unique Identified for client
Surname	String	Client Surname
CreditScore	Integer	Ranging from 350 to 850
Geography	String	USWM Sales Division
Gender	String	
Age	Integer	
Tenure	Integer	Length of client relationship in years
Balance	Decimal	Investment balance snapshot
Number of Products	Integer	Number of products with JPMC
HasChckng	Integer	1 = Has a checking account 0 = No checking account
IsActiveMember	Integer	1 = Digitally Active 0 = Digitally Inactive
EstimatedSalary	Decimal	Salary

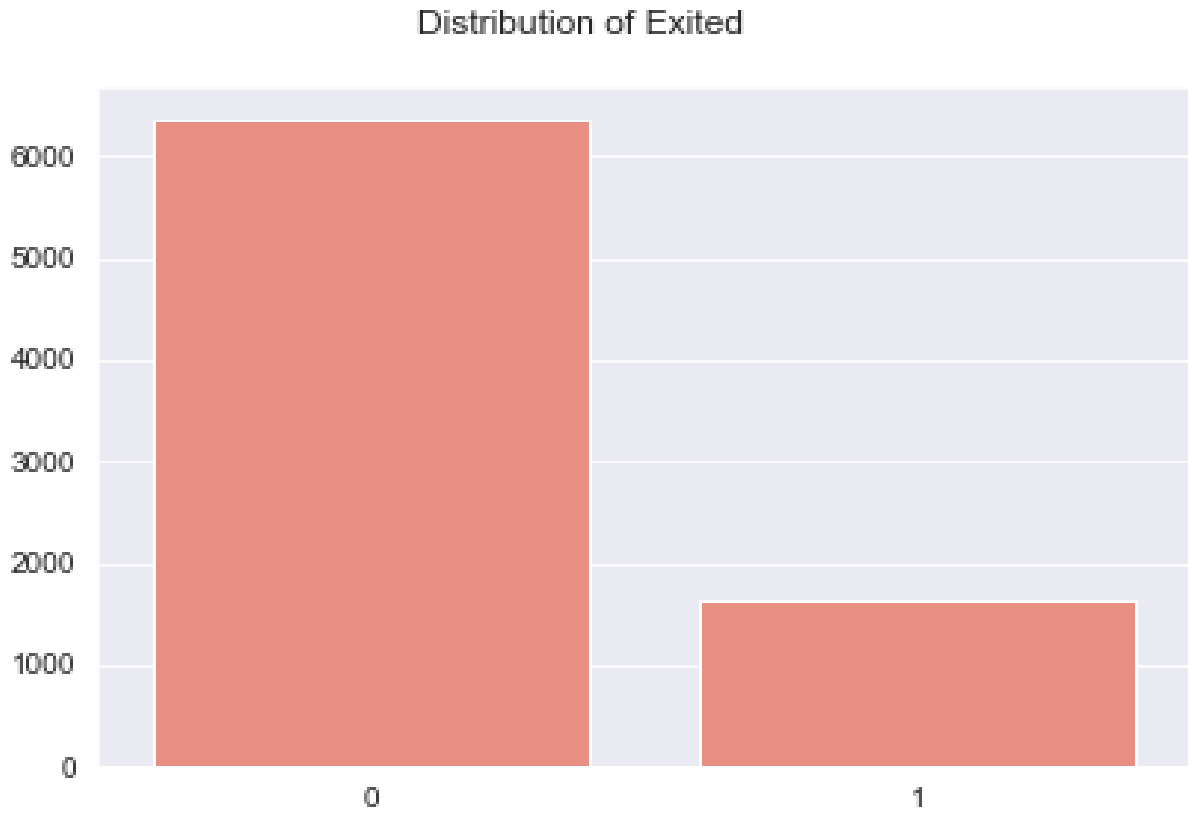
Things to Note:

1. Credit Score data was missing for some clients.
2. Max age of clients is 190.
3. There are no interactions amongst the features

Label Characteristics

Things to note:

1. Labels are unbalanced.
2. Recall is the better error metric.



Pipeline

1. Preprocessing

IDs are removed, age is caped at 100, and features were reformatted to appropriate data type.

2. Imputation

Missing values were replaced with the mode for categorical features or median for continuous features.

3. Transformation

Continuous features were standardized, categorical features were one-hot-encoded, and integer feature remained the same.

4. Modeling

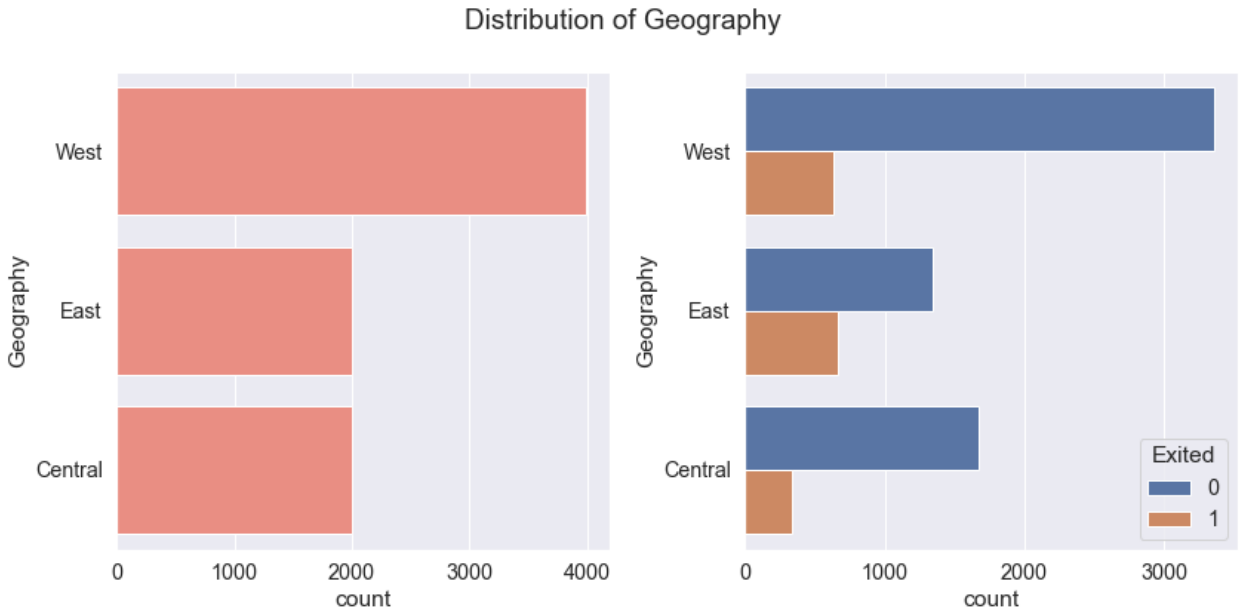
The processed data moves to the next step for model training and tuning.

Insights

During the Exploratory Data Analysis process, key differences were identified between churned and non-churned clients.

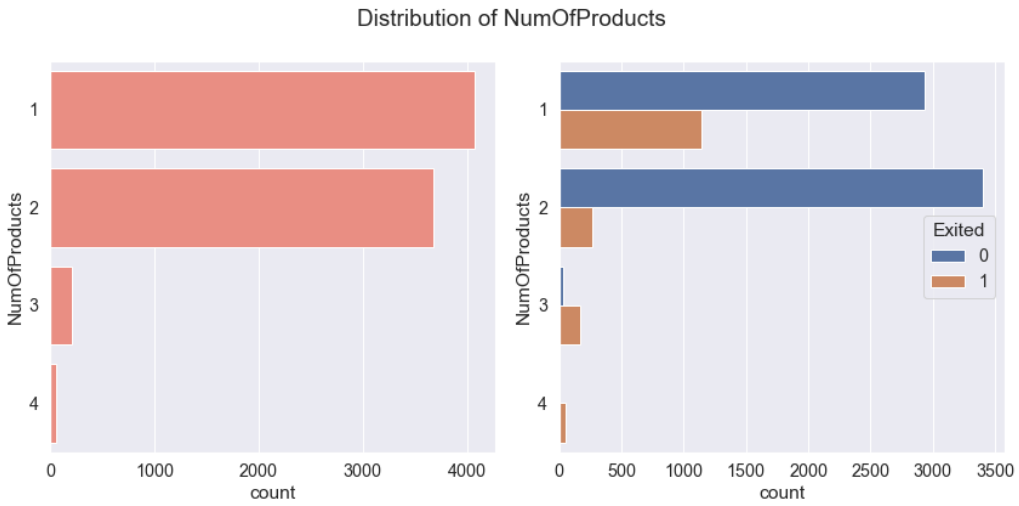
Key Insights

1. Clients in the West churn less frequently
2. Clients in the East churn more frequently

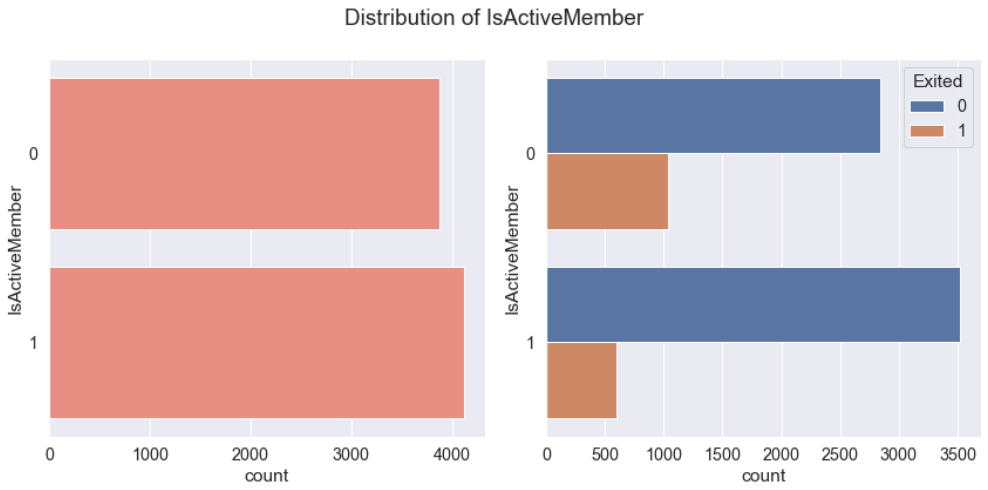


Key Insights

3. Clients with 2 products churn the least

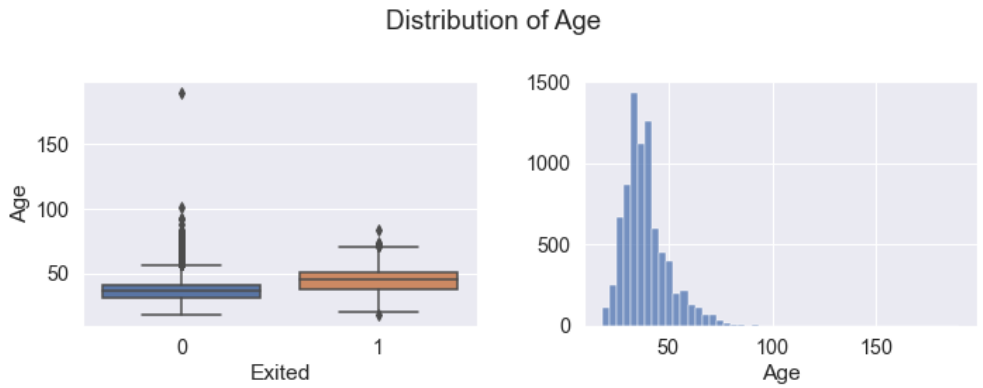


4. Active Client churn less

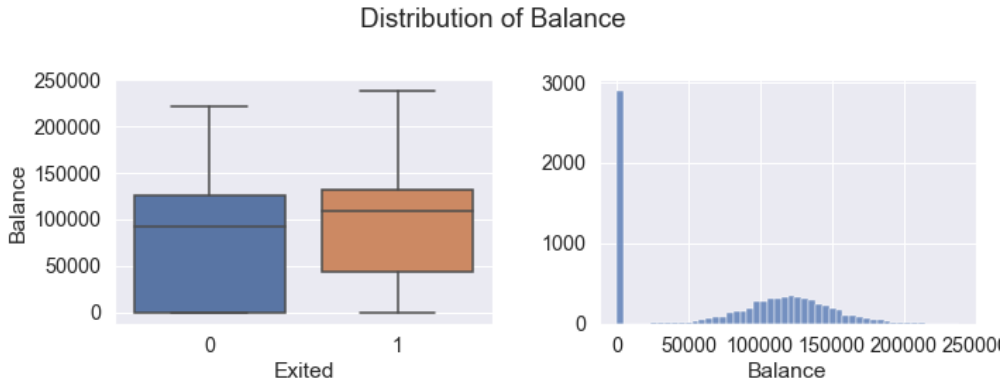


Key Insights

5. Older clients churn more frequently



6. Clients with a larger balance churn more frequently



Models

Models

Baseline Model

Majority Class Predictor (Should achieve 80% accuracy)

Logistic Regression

Linearly Separating Predictor

K-Nearest Neighbors (KNN)

Distance-based Predictor

Random Forest

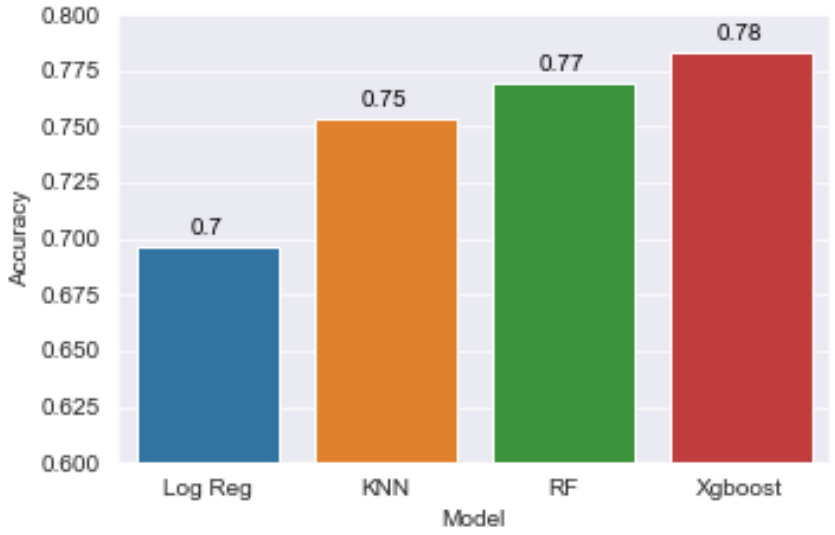
Grouping Purity-based Predictor

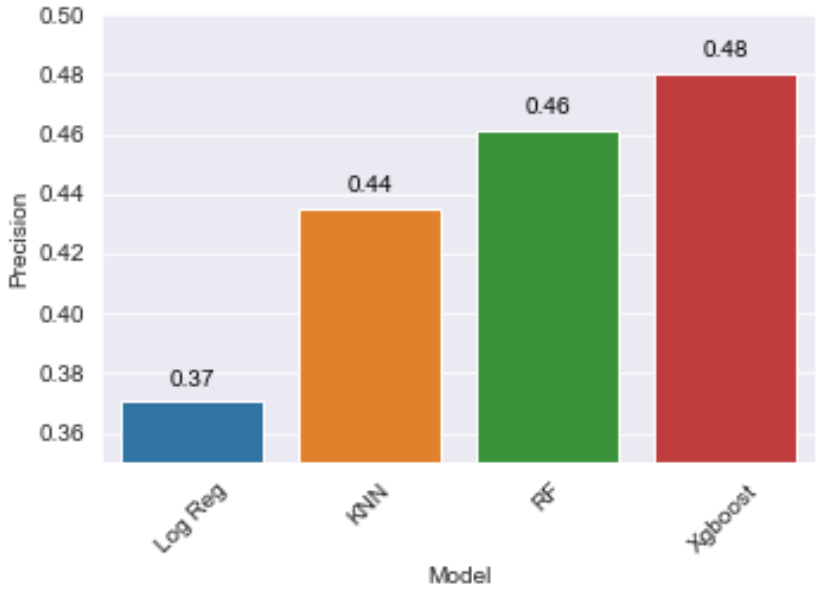
XgBoost Classifier

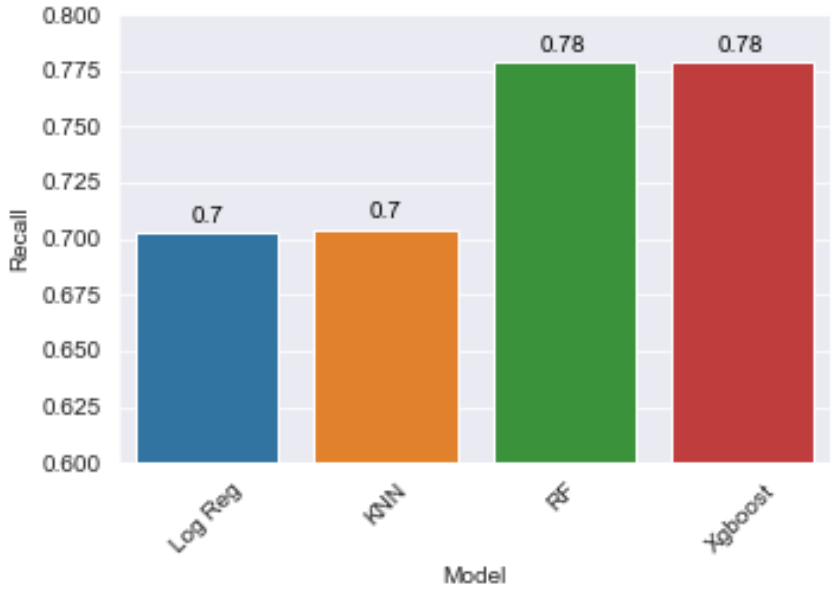
Gradient Boosted Decision Tree Predictor


Results

Model Results







Probability Threshold: 0.2
Best Model: XgBoost Classifier

Results on Test Data:

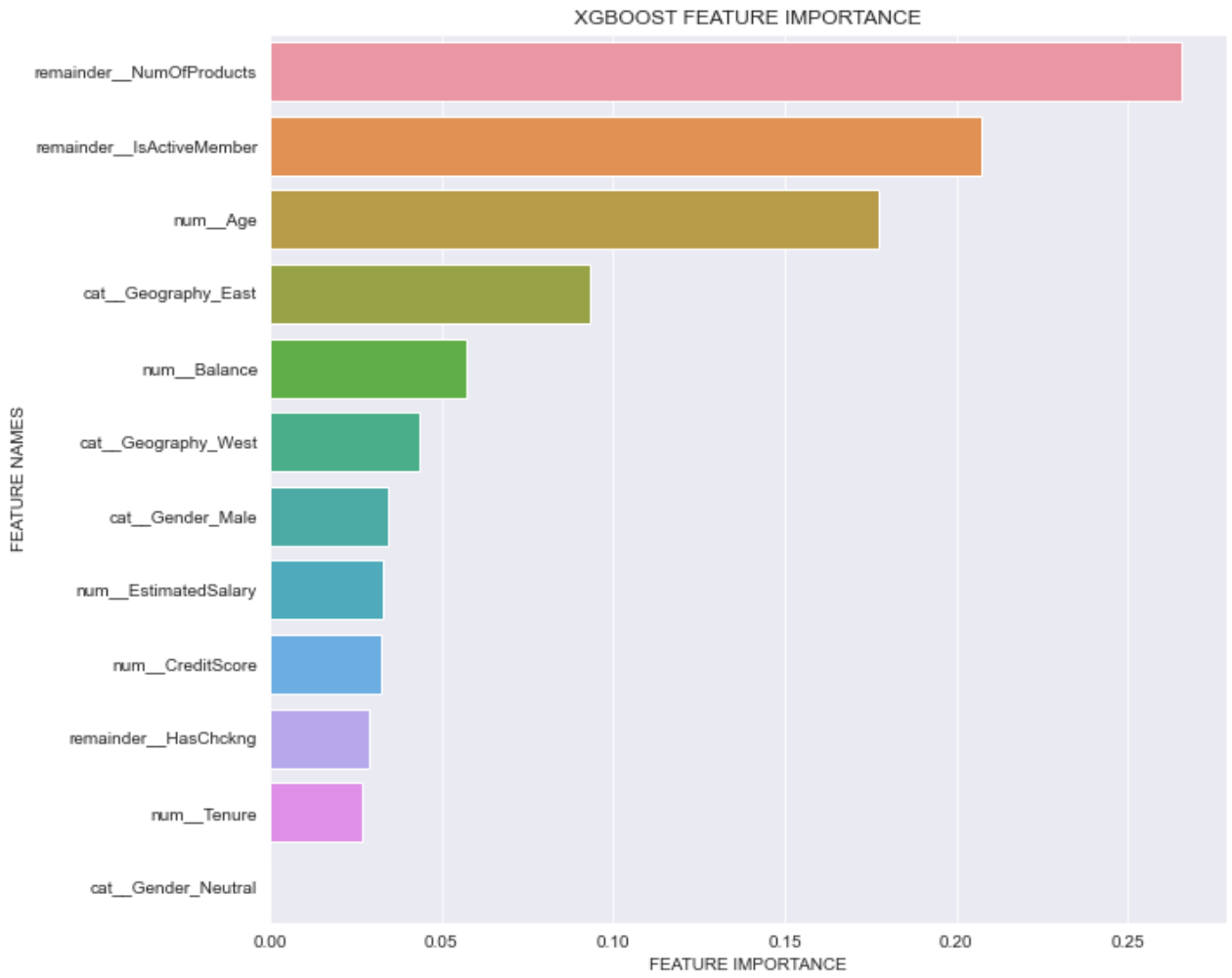
- Accuracy: 0.78
- Precision: 0.47
- Recall: 0.79


Solutions

Xgboost Feature Importance

The Feature Importance chart agrees with many of the insights that were discovered earlier:

1. 2 Products: Less likely to churn
2. Active Member: Less likely to churn
3. Older: More likely to churn
4. East: More likely to churn
5. Large Balance: More likely to churn
6. West: Less likely to churn



Recommendations

Use the Xgboost Classifier to Identify Possible Attrition

Depending on the characteristics of the client, below are some possible ways to prevent client churn.

Member Activity

Incentivize clients to be a more active member.

Geography

Focus client retention strategies on the east region.

Number of Products

If clients are only using 1 product, persuade clients to use an additional product.

Age

Reward older clients for being a customer.

Thank you

Thanks for your time and giving me the opportunity to present.

I look forward to working together in the future!

Troy Reynolds

Troy.Lloyd.Reynolds@gmail.com