PREDICTING BANK CUSTOMERS ATTRITION

Nada Alzahrani

Medium GitHub LinkedIn

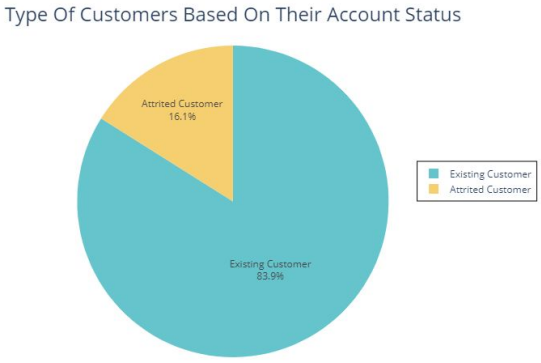
PROJECT MOTIVATION

Customer attrition which is also known as customer churn, customer turnover is defined as the loss of customers in a business, it is one of the biggest concerns especially in banking since customers are considered as the most valuable part of it. We are using the bank customer's data to predict possible attrited customers using Machine Learning to help prevent any possible attrition that may happen in the future.

PROCESS

1. Exploratory Data Analysis

- Class Balance.



- Customer's Demographic Profile
- Customer's Account Information.

2. Preprocessing

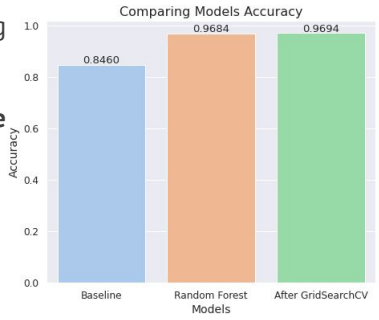
- Encoding categorical features ordinary like Education State, Credit Card Type etc.
- Split Data to 80% Train and 20% Test.

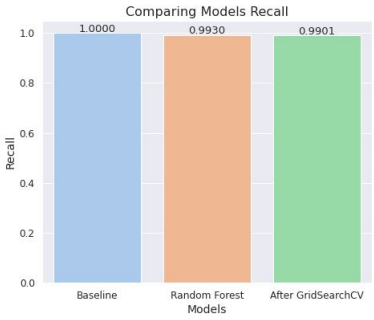
3. Modeling

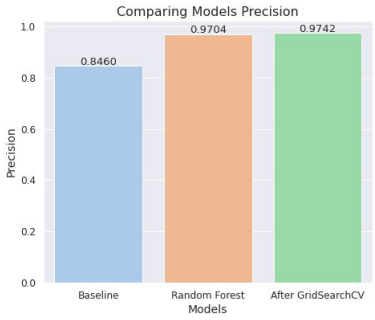
- Created a Baseline model that predict the most occurring class for all samples.
- used a machine learning ensemble method that uses multiple machine learning algorithms at once to obtain a better predictive performance.
- Picked the highest scoring model from the ensemble method then tuned it with GridSearchCV.

RESULTS

- The Best attrition predicting model was the **Random Forest after the GridSearchCV** with:
  - Accuracy: 96.94%
  - Recall: 99.01%
  - Precision: 97.42%







- **The Random Forest After GridSearch** got a higher false-positive rate than the normal Random Forest Model between 0 and 0.1 on the false positive rate axes, which means that the positive class (the existing customer) was identified falsely in the model after GridSearchCV.