Author: Chandramouli Das
Department: Product - Learning and Development

Core ML - Bank Customer Churn Prediction

Problem Statement :

Customer churn or customer attrition is a tendency of clients or customers to abandon a brand and stop being a paying client of a particular business or organization. The percentage of customers that discontinue using a company's services or products during a specific period is called a customer churn rate. Several bad experiences (or just one) are enough, and a customer may quit. And if a large chunk of unsatisfied customers churn at a time interval, both material losses and damage to reputation would be enormous.

A reputed bank "ABC BANK" wants to predict the Churn rate. Create a model by using different machine learning approaches that can predict the best result.

Dataset Description :

This is a public dataset, The dataset format is given below.

Inside the dataset, there are **10000 rows** and **14 different columns**.

The target column here is **Exited** here.

Download the dataset - **Churn Dataset**

The details about all the columns are given in the following data dictionary -

Variable	Definition
RowNumber	Unique Row Number
CustomerId	Unique Customer Id
Surname	Surname of a customer
CreditScore	Credit Score of each Customer
Geography	Geographical Location of Customers
City_Category	Category of the City (A,B,C)
Gender	Sex of Customers
Age	Age of Each Customer
Tenure	Number of years
Balance	Current Balance of Customers
NumOfProducts	Number of Products
HasCrCard	If a customer has a credit card or not
IsActiveMember	If a customer is active or not
EstimatedSalary	Estimated Salary of each Customer
Exited	**Customer left the bank or Not (Target Variable)**

Working Flow :

In order to create a model these are the following procedure -

- The Dataset need to be splitted in 70% of Train set and 30% of Test Set
- Proper feature engineering steps need to be followed.
- During Evaluation, need to check the accuracy score for both Training and Test Set
- Need to compare the accuracies for both Training and Test set, in order to check for the overfitting issues

Evaluation :

Your model performance will be evaluated on the basis of your prediction of the target column. The Evaluation will be based on the full Classification Report. The classification report should contain the following -

- Precision Value
- Recall Value
- F1 score
- Accuracy

Submission Process :

Need to submit the complete python notebook with proper details and comments.