Project Brief

Classification – Prediction of Bank Churn Customer

Bank Churn

One of the greatest challenges faced by service companies is to retain its customers. More customers means more business and revenue to the company. The cost involved in targeting new customers is a lot more than the effort to retain the existing ones. Keeping this in mind, companies (that also includes banks) try to get some insights into the customer data to identify those customers who are likely to leave (churn).

Companies employ lots of techniques to retain their existing customers – viz: loyalty programs, better discount rates, priority services, etc. The advantage of using these techniques is to determine what causes the churn and how it can be reduced in today's highly competitive market.

Business Objective

- Build a predictive model to determine which customers are likely to churn (leave the bank)
- Given a set of customer record, determine the probabilities of churn for each customer

Data Dictionary: BankChurn Data

The Customer churn dataset is of a bank that contains 10000 records in a CSV format. Click here to download the dataset for this project.

S.No	Feature	Data Type	Description
1	customerId	Numeric	Unique Customer ID
2	Surname	Character	Surname of the customer
3	Creditscore	Numeric	Credit Score of the customer (cannot be nulls or 0)
4	Geography	Categorical	Country of origin (this dataset contains only 3 countries)
5	Gender	Categorical	Gender of customer
6	Age	Numeric	Customer present age (in years, cannot be nulls or 0)
7	Tenure	Numeric	Tenure with the bank
8	Balance	Numeric	Current bank balance of customer

Data Dictionary: BankChurn Data (contd.)

S.No	Feature	Data Type	Description
9	Numofproducts	Numeric / Categorical	Number of bank products currently held/purchased
10	Hascrcard	Numeric / categorical	Does the customer own a credit card ?
11	Isactivemember	Numeric / categorical	Is the customer still active or not ?
12	estimatedSalary	Numeric	Salary of the customer
13	Exited	Numeric / categorical	Has the customer churned or not ? (1 -> yes, 0 -> no)

Technical Goals

- Understand the data very well. Do all transformations / data engineering / etc. wherever applicable
- Perform Exploratory Data Analysis (EDA)
- Carry out all the Data mining tasks
- Identify the salient features that will determine the best results
- Perform the model evaluation to select the appropriate algorithms