Project proposal

For

Customer Churn Analyzing Dashboard

Version 1.0 approved

Prepared by Ayesh Vininda

190649F

2020/04/01

Contents

1. Problem definition

A. PROBLEM

One of the most pressing challenges for business organizations is the acquisition and retention of new customers. On the other hand, organizations always face customer churn, which losing customers due to poor service of them, prices, etc. Because obtaining new customers is more expensive than keeping existing ones, firms must analyze customer turnover and develop strategies to decrease it.

As above-described higher customer churn is bad for Telco companies such as Chatter Box. A higher customer churn rate is positively influenced and proposition to the company's sustainability. Find ways to solve that kind of problem using data science which helps get a better analytical idea for finding solutions to the problem.

B. MOTIVATION

The goal of this project is to develop an end-to-end business strategic analysis solution for Chatter Box Telco Pvt.Ltd in the Banana Republic to identify and reduce customer churn. As the solution CEO of Chatterbox Telco, Mr. William suggests that building a web-based interactive dashboard provides valuable insights for him to make decisions.

The business solution provides a web-based application that can use them anywhere because it is deployed on the cloud platform. The dashboard shows a lot of features that help to identify reasons why customers attrition happens. Identifying the reasons helps avoid customer attrition which reduces losing customers. The system provides a way to check each customer separately displaying his details.

C. OBJECTIVES

1. Analyze the given dataset in Chatterbox Telco, to do the Customer Churn Prediction using data handling tools.
2. Do the data preprocessing on the dataset. Such as,
    a. Data cleaning
    b. Feature Engineering
    c. Feature selecting
    d. Data transformation

3. Training different models and comparing the performance and selecting the best model to evaluate.

4. Creating an Interactive dashboard using graphing tools.

5. Deploy the dashboard on the web-based application.

2. Proposed solution (step-by-step)

There are a lot of objectives to achieve, but slightly going through the above objectives we can categorize t all of them into two main objects.

1) Analyzing dataset and creating machine learning model to predict the churn.

2) Create a web-based interactive dashboard.

A. ANALYZING DATASET AND CREATING MACHINE LEARNING MODEL TO PREDICT THE CHURN.

❖ As same as in every project initially we need a collection of data. In this project Chatter Box Telco PVT.LTD already provides it.

❖ Data analysis is another major part of the data science project. Data analysis is happening any time before the get perfect model.

  o When initially data collection gets.

  o Before and after every data preprocessing step.

  o Before and after training model

We can categorize data analysis into major 3 categories. Such as descriptive, exploratory, and predictive.

❖ Data preprocessing is another important step in a data science project. Here we have to do

  o data cleaning

    ▪ to handle duplicates, outliers, and missing values.

  o feature selection

    ▪ selects the most essential features to model creation

  o data transformation

    ▪ transform data into different structures.

❖ Train different models, rather than creating one model train a few models for this problem. Since this is supervised classification we can use LogisticRegression, k -nearest neighbors, decision trees, support vector machine, Naïve Bayes, and also try boosting models such as CatBoost, LightGBM, XGBoost classifiers. To increase the performance of the model do some analytical and preprocessing techniques.

❖ Evaluate models and select the best model to get the final result.

FOR IMPLEMENTATION TECHNIQUES

❖ For the implementation and analysis models, I mainly use python 3.9 programming language because it is easy to highly used in machine learning and data science areas and it has a lot of built-in methods for machine learning and data science and it has a good community.

❖ For the coding, I used Jupiter notebooks (ipython files .ipynb) and python files (.py)

❖ Use google collab, Jet brains Data spell, and vs code as development environment software.

B.   CREATE A WEB-BASED INTERACTIVE DASHBOARD.

❖ Firstly need to add the dataset to the power bi.

❖ This web dashboard has a few pages such as home, overall analysis, customer analysis, churner analysis, and the prediction for the given input. So, first, create these pages using the power bi tool.

❖ Add features in which related overall to the overall analysis page, only customer's features add to the analysis page, add churner details to churner analysis.

❖ During adding features add the machine learning model described above, and put it into the power bi as a python script.

❖ Create an interactive dashboard And publish it to the web.

❖ Getting the embedded code and putting it in the simple web application is the final step.

FOR IMPLEMENTATION TECHNIQUES

❖ For mainly to implement an interactive dashboard, I hope to use the power bi desktop tool (which works with .pbix file format)

❖ For the publishing dashboard as embedded code use power bi web platform.

❖ For the web development just need HTML.

3. Proposed features of the dashboard

❖ In this dashboard, I just talk about some pages. Home page, Overall analysis, customer analysis, churner analysis.

❖ The home page is just a web page for user interaction.

❖ The overall analysis of both churners and customers. On the Overall page, I hope to add,

  o Total charges

- Total calls
- Account length

Features with interactions with location code international plan and voice mail plan. Above charges and calls.

❖ Churners and customers page can view features along with charges, calls, account length, the number of churners and customers these things also interact with the location code, international plan, and voice mail plan. In this way, we can compare both churner and current customers and easily identify the reasons for the customer attenuation.

❖ For these interactive web bases, I hope to use power bi graphical components because it helps the decision-maker to take quick decisions. Power bi provides a lot of graphs and charts with color and many more effects which help to understand the situations by just looking at the dashboard.

❖ Not only visualization there are some filtering and sorting techniques such as slicers, selectors, lists, dropdowns, and filters which help to quickly filter and sort specific things.

❖ And there is also a prediction page which discission makers can enter the current customer details to the application and get that customer to be churn or not.

4. Project timeline

No.	Project task	Start	End	Time (Days)
1	Get data collection and research domain	2022/03/25	2022/03/31	6
2	Analyzing dataset	2022/03/31	2022/04/06	6
3	Preprocessing	2022/04/06	2022/04/15	9
4	Model development	2022/04/15	2022/04/23	8
5	Finalizing model	2022/04/23	2022/04/30	7
6	Create a power bi dashboard	2022/04/30	2022/05/15	15
7	Web application	2022/05/15	2022/05/30	15
8	Final report	2022/05/30	2022/06/04	5