**SENIOR PROJECT PROPOSAL**

**Title of the Project:** Customer Behavior Recommender System using Machine Learning

**Project Team:** Hüseyin Arda Demirci, Umut Arıcı

**Advisor’s Name:** Assist. Prof. Dindar Öz

**Introduction / Background / Overview**

Data Science has become the backbone for service and product businesses when it comes to deciding strategies to maximize customer satisfaction and profits. For long, businesses were already taking advantage of their customer and sales data by using data analysts. However, when datasets grow exponentially with the increase in the number of different customer types and variety of products, profiling customer behavior has become impossible for data analysts. Today, data analysts are using the outcomes of machine learning models to analyze to decide on strategies in the best possible way.

What we want to create is a Recommender System desktop application(Windows, Linux) that is based on machine learning methods for E-Commerce businesses. The application will take datasets that contains data on customers with their purchases and turn these data’s into useful information using Clustering and Regression machine learning techniques. The application will consist of both frameworks and our own implementation techniques. Then the application will perform data analytics on the information created to profile customers to detect their behaviors using Random Forest machine learning technique. Because it uses machine learning methodologies this application will train itself continuously with every new given data. Aside from the data analytics part, the whole source code of this application can be reused in the future projects with similar problems.

For the user part, this application will be used by the data analyst of the company that purchases our application. After creating a profile for the customer behavior on the background, the application will show which customers are likely to purchase which and what kind of products in the future. The application will provide this guidance through Histograms, Cluster Analysis Mapping, and 3D Scatter Plot along with some other visualization techniques for the data analyst to observe. Through their observations, data analysts can easily decide on which customers are going to be advertised on which and what kind of products. And through their decisions they can give advertisement departments of the company very precise and on point instructions on what products to advertise on which customers.

Therefore, usage of our application will increase the speed and efficiency of the analysis and marketing processes while saving time and money on advertisement and feasibility. On the customer side, it will lead to high customer satisfaction since they will be exposed to things they relate to. Moreover, this will allow to identify at-risk customers and take proactive steps to retain them.

On training stage of the application, we will use a sample dataset that will be provided to us by the customer company and that dataset will have no vital importance and pose no threats for security. After the application is completed and delivered to the customer, they are going to run the application on a machine that is not connected to any network for security reasons, and they will feed the application their own datasets and use it themselves. Considering they are already going to use the most secure setups for this since they will also use it for their own datasets, cyber security will be customer companies’ responsibility.

To make things clear let us consider this example of how our Recommender System will work. Consider our client company as a mainstream E-Commerce shopping site (such as Trendyol, Amazon, etc.) and consider a middle-aged man as a customer that has purchased a ‘camping tent’ from their site. In the end, our application will profile his purchase behavior and recommend items that he is likely to purchase later such as sleeping mats, flashlights, outdoor clothing etc. Both the sales and customer data will be used in this process.

As a conclusion, our aim is to help customer company to manage their marketing budget efficiently and decide on feasible decisions by analytical data as well as providing enhanced customer experience which leads to higher sales numbers. In the basic terms, the goal is to help our customer company to achieve higher profits by satisfying their customers better.

**Aims and objectives**

**Aim:**

* To turn businesses’ raw data into useful information and to use that information to reveal customer behavior.
* To represent those profiling and recommendations visually to help understand easily.
* And help businesses increase their profits and customer satisfaction.

**Objectives:**

* To filter and sort out raw data to identify and categorize customers respectfully on their behavior.
* Identify and implement necessary Machine Learning principles and algorithms on the model to reveal patterns of customer behaviors.
* Visualizing end results using frameworks using graph, maps and 3D scatter plots in order to make it easy to use.

**Expected Outcomes/Deliverables**

* At the end we will have a desktop application that successfully profiles and reveals customers’ behaviors with any given additional relevant dataset.
* By doing so, it will achieve to turn the raw data of the businesses into enhanced and personalized customer experience by creating recommendations.

**Use-case Point Complexity**

|  |  |  |
| --- | --- | --- |
| **Factor Number** | **Description** | **Is its importance average or above for the project? (Write YES if it is)** |
| **T1** | Distributed system |  |
| **T2** | Response time/performance objectives |  |
| **T3** | End-user efficiency | **YES** |
| **T4** | Internal processing complexity |  |
| **T5** | Code reusability | **YES** |
| **T6** | Easy to install |  |
| **T7** | Easy to use | **YES** |
| **T8** | Portability to other platforms | **YES** |
| **T9** | System maintenance |  |
| **T10** | Concurrent/parallel processing |  |
| **T11** | Security features |  |
| **T12** | Access for third parties |  |
| **T13** | End-user training |  |

**Type of Project**

Data Processing and Data Analysis, Application Development Project, Recommender System

**Keywords**

Data Science, Machine Learning, Data Analysis, Customer Behavior, Recommender System

**Related Research**

Customer Behavior Prediction using ML, is a systematic process of collecting and preparing business data, selection of the right model and training it. Though it is called ‘customer behavior prediction’ on a variety of sources, we can refer it as ‘revealing customer behavior’ to prevent confusion since our application is not about predicting future. It is a very powerful tool for businesses which plan on making data-driven decisions in order to personalize and enhance customer satisfaction therefore growth of the business.

Understanding consumer behavior is crucial for businesses to adapt to changing preferences and preferences. Analyzing consumer behavior requires current analytical methods and innovative marketing strategies. Understanding consumer preferences is essential for identifying effective marketing strategies and ensuring the success of businesses in the modern technological world. Machine learning techniques are used to predict the behavior of the customer which give good accuracy. (Valecha, Varma, Khare, Sachdeva and Goyal, 2018).

<https://hrcak.srce.hr/file/332434>

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8597070>

**Research questions and hypotheses**

**Methods**

Rapid Application Development (RAD), Machine Learning Techniques such as Clustering methods, Random Forest methods, Phyton

**Resource Requirements**

Relevant Sample Datasets, PyCharm IDE, GitHub, Pandas, Anaconda, Jupyter, NumPy, Scikit-Learn, COMP 4350 and SE 4488 Course materials.