**Customer Churn Analysis**

Weekly Report

Ahmedabad University

4rier Series

Prof. Mehul Raval

CSE523 - Machine Learning

The initial project chosen by our group was customer segmentation and product recommendation. We chose to initially segment customers and then run specific product recommendations based on segmented customers in order to present a better and more accurate recommendation. The following flow diagram illustrates the same.



We intended to use the following to achieve the said recommendation. We were going to run the segmentation based on the following parameters:

* Demographic
* Behavioural
* Geographical
* Psychological

The same would give us different clusters to work upon. Additionally, we explored collaborative and content-based filtering, finding the applicability of both in the project.

We ran into a problem wherein we could not find datasets which could coordinate with the data required. A combined data set was required wherein we needed the sale of specific products as well as the customers who bought the specific products. This caused us to explore other problems post an internet-wide dataset search. We could find a few datasets which were dummy datasets and thus were not satisfactory for the implementation.

**University Prediction based on Marks**

The revised project chosen was a prediction of university admission based on various factors such as GRE score, IELTS score, TOEFL score, internships and projects done etc. We planned first to predict all the possible universities where a student might get admission and then recommend them to the student based on their specific requirements and the university credentials.

For the same, we planned to use two datasets: admission prediction and university rankings.

Although the dataset chosen by us had all the possible factors that might affect a student’s admission, it was not from a credible source. Additionally, no such parameterised datasets were available, which were from a ratified source.

The initial approach for this project was using collaborative filtering and algorithms such as k-means clustering and k-nearest neighbours to predict possible student clusters. Additionally, the Support Vector Machine algorithm and linear regression can be used to predict the possibility of admission to a specific university.

**Customer Churn Analysis**

Customer Churn refers to the level of attrition that can be seen across companies. It refers to customer loyalty and market retention. Our project focuses on customer churn rates in telecommunication sector. This churn can happen due to bad customer service, financial issues or it can be when customers switch to another competitors. We intend to see the churn rate as low as possible at all times.Our main motive is to identify “at-risk” customers and methods to lower churn.

Features can be subdivided into following categories:

**Demographic customer information:** gender, age, marital status, location

**Services that each customer has signed up for:** Phone Service, Multiple Lines, Internet Service, Online Security, Online Backup, Device Protection, Tech Support, Streaming TV, Streaming Movies

**Customer account information:** Tenure, Contract, Paperless Billing, Payment Method, Monthly Charges, Total Charges

Algorithms that can be used to implement customer churn are:

* Logistic regression
* Naive Bayes
* Random Forest
* Multivariate Gaussian Distribution
* Clustering

**Scope for Next Week**

We intend to analyse and clean the datasets and format them based on our requirements. We also intend to explore the different algorithms which could be used to predict customer churn. Furthermore, we would like to predict the probability of churn of a customer based on the various parameters and thus we intend to formulate a way of doing the same.

At the end of next week, we should be able to see the following:

1. Analyzed and Clean Dataset with categorical-numerical mappings
2. At least one model made to explain customer churn
3. An exploratory analysis of the available algorithms which solve this problem.