NLP - Assignment 1

Instructor: Dr Pooja Agarwal

Date: 08-Feb-2022

Assignment can be access through Edmodo class link fjppqj

Dataset Link

The Excel sheet for the same is attached

Dataset Description

The data provides historical customer transactions for a retailer, for purchases made on their e-commerce platform between Dec 2010 to Dec 2011. Customers are uniquely identified by their Customer ID (Col G).

Problem Statement

Assume that the company has been experiencing high customer churn and a reduction in repeat customers. As part of this case study, **you are required to build a predictive model to predict the likelihood of a customer churning**

Churn Definition (Target Variable definition)

For this problem, the churn should be defined as follows:

The training data should include all customers who have made a transaction between 1st Dec 2010 to 31st Aug 2011. Customers from this set who have NOT made any subsequent purchase in the period Sep 2011 to Dec 2011 should be labeled as **"Churn".** All other customers with purchases made during this follow-up period should be labelled as **"Not Churn"**

Submission Guidelines only through Edmodo

Submit a google drive link with edit persmission (Drive Name should be NLP_BATCH_BatchNo.) containing

1. **A folder with**

   A Jupyter Notebook (Python program) with the case study solution. The solution should incorporate the following four task:

   A. Target variable creation (based on churn definition provided)
   B. Customer level feature creation, from transaction dataset
   C. EDA on the features created
   D. Model training, validation and evaluation

2. **A presentation** (.ppt) containing the name of the students, SRN, Section, Abstract, Flowchart, packages or libraries used, output screenshots