DS 422 Business Analytics Lab Assignments (Last Updated: 12/08/2020)

Reference - *Larose & Larose, Data Mining and Predictive Analytics, Wiley, Second Edition. (Part I - Data Preparation; Chapters 1, 2, 3, and 4.)*

Case Study 1A – Telco Churn

1. Introduction – The client is a telecom company. They have an issue with customer churn. Currently, they want to improve their customer retention efforts. For this, they have engaged on a churn prediction exercise.

2. Challenge – The customer has stated that the objective of the churn prediction exercise is to predict churn. That is, the model should be able to identify (based on past data), who is going to churn in the next time period.

3. Your role – You are the lead data scientist for the project. In this phase, you have to confirm that you understand the business requirements (business understanding) as well as the data (data understanding) for the project.

4. Details of the data – The customer has shared the data with your team. The data is in the form of a text file Telco Churn Data – Double Click. The description of the various fields are given and the customer has stated that the target variable is "churn".

5. (a) On the basis of the problem statement and the data provided, carry out the following steps for the client. The response for this exercise (Data Preprocessing and Exploratory Data Analysis) will be a report, with separate sections for the deliverables below.

   (i) Business Understanding – Goals & Success Criterion.

   (ii) Data Understanding – Data Exploration & Quality Report.

   (b) Perform initial analysis (Variable Selection & Dimension Reduction) to identify possible variables that could be impacting churn. Provide a report of the same.

	Name	Description
1	region	Geographic indicator
2	tenure	Months with service
3	age	Age in years
4	marital	Marital status
5	address	Years at current address
6	income	Household income in thousands
7	ed	Level of education
8	employ	Years with current employer
9	retire	Retired
10	gender	Gender
11	reside	Number of people in household
12	tollfree	Toll free service
13	equip	Equipment rental
14	callcard	Calling card service
15	wireless	Wireless service
16	longmon	Long distance last month
17	tollmon	Toll free last month
18	equipmon	Equipment last month
19	cardmon	Calling card last month
20	wiremon	Wireless last month
21	longten	Long distance over tenure
22	tollten	Toll free over tenure
23	equipten	Equipment over tenure
24	cardten	Calling card over tenure
25	wireten	Wireless over tenure
26	multline	Multiple lines
27	voice	Voice mail
28	pager	Paging service
29	internet	Internet
30	called	Caller ID
31	callwait	Call waiting
32	forward	Call forwarding
33	confer	3-way calling
34	ebill	Electronic billing
35	loglong	Log-long distance
36	logtoll	Log-toll free
37	logequi	Log-equipment
38	logcard	Log-calling card
39	logwire	Log-wireless
40	lninc	Log-income
41	custcat	Customer category
42	churn	Churn within last month