Advanced Statistical Modeling
AMPBA
Assignment
Group Assignment

This is the group assignment and has 40% weightage in the course.

Deliverables:

1. A .pdf/word file document. Include any material (table/figures) to support your answers.
2. Code files (R, python or excel) used to solve the assignment.

General Instructions:

1. Attach Assignment Submission Form on the first page of your submission.
2. Submissions without this form will not be considered.
3. This is a group assignment.
4. The word file should be named as "Group-nn" where nn is your group number.
5. Make sure that only one person from the group submits the assignment.
6. Type your group member names on the first page of the Assignment. Marks will not be awarded if the name(s) are missing.
7. Please do not copy the complete questions, just mention question number.
8. Please do not submit .zip files, the submission will not be considered.
9. Late submission penalties are applicable according to the course outline.
10. The honor code for this assignment is **2N-b.**
11. The only graded submission format will be PDF/.docx.
12. We will check the other files/excel based on the requirement.
13. Handwritten content will not be considered for evaluation.
14. Submission will not be considered if the instructions are not followed.

Instructions for Assignment:

Each group has been provided with a separate dataset. (Please work on the dataset that corresponds to your group number, for example : if your group number is 12 then choose file12) These datasets are unique for each group. The description of the variables is given in the appendix.

You are free to use the software of your choice to carry out the analysis.

Based on your dataset, carryout the following analyses. Your analyses and the interpretation should be submitted by the due date. Please note that there is no penalty for early submissions!

1. (5 marks) Build a prediction model for predicting churn, using both discriminant analysis and logistic regression. You may consider using the appropriate explanatory variables. If necessary, create additional variables using the existing variables.
2. (5 marks) Comment on the predictive accuracy and the impact of each of the explanatory variables on churn.
3. (10 marks) Divide all the customers into 3 categories namely Low, Medium and High using the variable "TotalCharges". Let us call them "Customer Value Segments". Build prediction models to predict the category/ Value Segment. Comment on the profile of the customers in each category/ Value Segment. Identify appropriate strategies to shift customers from each value segment to the next higher segment.
4. (5 marks) Create an overall survival curve using the Tenure variable. Use Kaplan-Meier method.
5. (5 marks) Create separate survival curves for different categories of customers (for example, Gender). Comment on the differences in these survival curves.
6. (10 marks) Build Cox's Hazard model using appropriate explanatory variables. Comment on the coefficients of the model.

Deadline: 9th April 2022, Saturday, 11:55 pm

Note: Please attach the Assignment submission form.

Continue to the next page for Appendix

Appendix

Variable Name	Variable Description
customerID	Unique customer Identifcation Number
gender	Male or Female (Categorical)
SeniorCitizen	Is the person above 60 years of age (Categorical)
Partner	Is customer living with partner/ married (Categorical)
Dependents	Are there any Dependents on customer (Categorical)
tenure	Duration in Months, for which they have used the services (numerical)
PhoneService	Do they have calling services (Categorical)
MultipleLines	Did customer opt for multiple line services (Categorical)
InternetService	Did customer opt for Internet services (Categorical)
OnlineSecurity	Did customer opt for Online Security services (Categorical)
OnlineBackup	Did customer opt for Online Backup Services (Categorical)
DeviceProtection	Is the device being protected from software malware (Categorical)
TechSupport	Is customer provided with technical support (Categorical)
StreamingTV	Is customer provided with Streaming TV services (Categorical)
StreamingMovies	Is customer provided with Streaming Movie services (Categorical)
Contract	If the services are subscribed on monthly or yearly basis (Categorical)
PaperlessBilling	If the billing is online or being delivered at home (Categorical)
PaymentMethod	If the bills are paid in e-transfers or in cheque format (Categorical)
MonthlyCharges	Monthly carged to customer based on the services provided (Numerical)
TotalCharges	The total cost incurred by the customer (Numerical)
Churn	Customers who left within the last month (Categorical)