THE UNIVERSITY OF CHICAGO

MODULE 7 ASSIGNMENT

For this assignment, you are provided with student GRE scores and GPA and a binary column indicating whether they were admitted to a certain university. Follow the steps below to perform a logistic regression to predict how GRE scores and GPA effect admission to this school. Answer the following questions and submit your work as a knitted R Markdown file in either Word or PDF format.

Pre-Processing:

a. Load in the data.
b. GRE and GPA are measure on significantly different scales. To allow us to interpret these variables on the same range, scale both variables using standardization. This means each variable will have a mean of 0 and a standard deviation of 1.
c. Set the dependent variable "admit" as a factor variable and perform logistic regression with two predictors: GRE and GPA.

Questions:

1. Provide an interpretation for the intercept coefficient. What does it mean if both predictors are equal to 0?

2a. Assuming an average value for GRE, calculate the effect of a one unit increase around the mean for GPA.

2b. Assuming an average value for GPA, calculate the effect of a one unit increase around the mean for GRE.

3a. With an average value for GRE, calculate the probability of being admitting under the following conditions for GPA: 3 SD below mean, 2.5 SD below mean, 2 SD below mean, 1.5 SD below mean, 1 SD below mean, 0.5 SD below mean, Mean score, 0.5 SD above mean, 1 SD above mean, 1.5 SD above mean, 2 SD above mean. What is the average marginal effect?

3b. With an average value for GPA, calculate the probability of being admitting under the following conditions for GRE: 3 SD below mean, 2.5 SD below mean, 2 SD below mean, 1.5 SD below mean, 1 SD below mean, 0.5 SD below mean, Mean score, 0.5 SD above mean, 1 SD above mean, 1.5 SD above mean, 2 SD above mean. What is the average marginal effect?

4. How many standard deviations above the mean should your GRE score be if your GPA is 0.5 standard deviations below the mean and you'd like a 75% chance of being admitted?

5. Multiply the intercept by -1. Divide this value by the sum of the two slope coefficients. Use this result as values for an observation of GRE and GPA and calculate the output from the model. What's your interpretation?

6. Generate predictions using a 50% classification boundary. Report overall accuracy and balance accuracy. Feel free to share any other metrics you find interesting. Are you satisfied with this

classification boundary? If yes, say why. If not, evaluate results when using another classification boundary.

7. Plot an ROC curve and report the area under the curve. Based on this and your classification predictions, how do you evaluate the ability of this model to use GRE score and GPA to differentiate between whether students will be admitted?