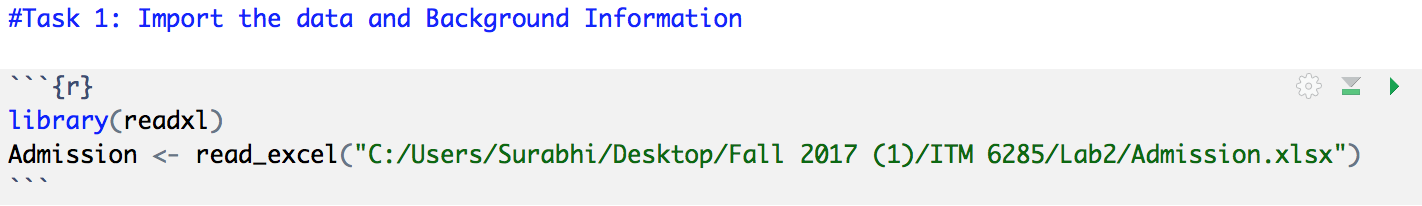
Surabhi Asati

NetID: va7892

ITM 6285 Data Mining Lab - Logistic Regression

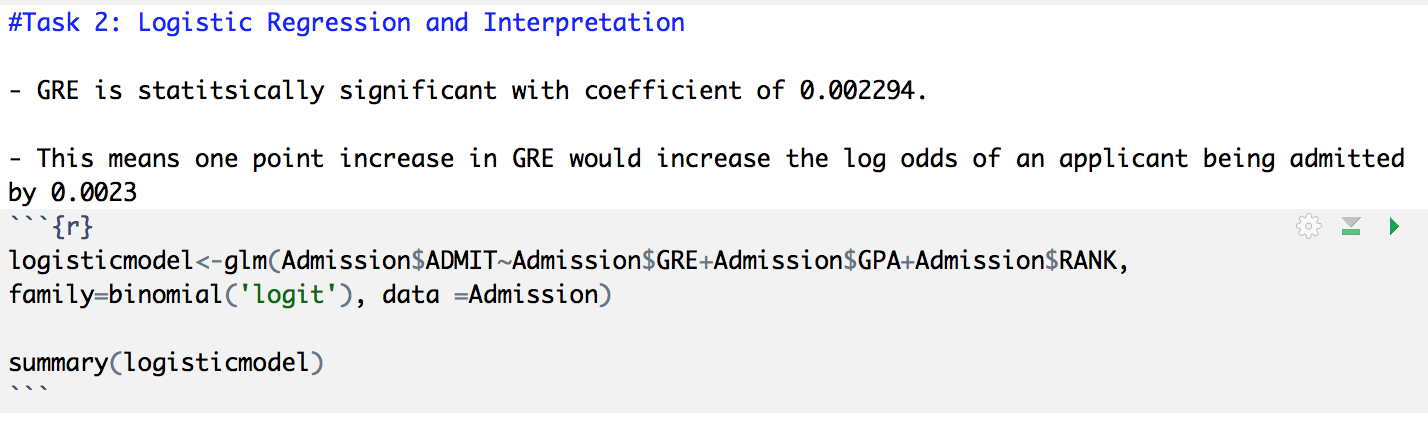
Logistic Regression, Confusion table, and Prediction (expected time - 1.5 hour)

**Task 1: Import the data and Background Information**



**Task 2: Logistic Regression and Interpretation**

Please run a logistic regression of Admit (the dummy variable being 1) onto GRE, GPA, and Rank. Please interpret the coefficients on GRE, GPA, and Rank: how changes in these variables affect the odds ratio of an applicant being admitted.



If the student’s GRE score is 640 and GPA score is 3.19 and he is applying to school that has rank 1, then probability of being admitted is 48.43%

**Task 3: Prediction and Confusion Table**

Please make predictions on the original sample. Use a probability of 0.50 as cutoff.

Then complete the following Confusion table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | Predicted Value | |  |
| 0 | 1 | Total |
| True Value  “Admit” | 0 | 253 | 20 | 273 |
| 1 | 98 | 29 | 127 |
|  | Total | 351 | 49 | 400 |

**Task 4: Calculate the Sensitivity and Specificity**

Calculate the sensitivity and specificity based on the previous confusion table

Sensitivity = 29/127 = 0.228 ~ 23%

Specificity = 253/273 = 0.9267 ~ 93%

**Task 5: Making Predictions for New Data Point**

Suppose we have a new student with GRE=640, GPA=3.19, Rank=1. Please calculate the probability of this students being admitted.

|  |  |  |
| --- | --- | --- |
| Intercept | -3.449549 | 1 |
| GRE | 0.002294 | 640 |
| GPA | 0.777014 | 3.19 |
| RANK | -0.560031 | 1 |
|  |  | -0.06274534 |
|  | Odds | 0.939182615 |
|  | Probability | 0.484318809 |

**PROBABILITY = 48.43%**