

IS 507 Assignment 8

Due by Sunday, December 4, 2022 11:59PM

Problem 1: The Excel spreadsheet Alzheimer.csv contains one sheet named Alzheimer, which is data attempting to explain whether a patient has Alzheimer's Disease. These are data from a sample of 336 employees and consists of 9 variables for each patient. These are:

- 1) Dementia-Outcome variable-patient diagnosis
- 2) Gender-Female=0 and Male=1
- 3) Age-Age of patient (in years)
- 4) Education-Years of Education
- 5) SES-Socioeconomic Status 1=Low and 5=High
- 6) MMSE-Mini mental state examination score
- 7) CDR-Clinical Dementia Rating
- 8) eTIV-estimated total intracranial volume
- 9) nWBV-Normalize whole brain volume
- 10) ASF-Atlas Scaling Factor

Develop a **Logistic Regression model** to classify the Dementia event from the other variables.

- a) Create a logistic regression model and explain the significant odds ratios in terms of Dementia.
- b) Create a confusion matrix and explain how well the model is classifying who has Dementia.
- c) Create an ROC curve and calculate the c-statistic (auc). What does this mean about the model?
- d) What are the differences between the information in part a and part b?
- e) How does this model differ from the linear discriminate analysis you ran in Assignment 7?