

Assignment 5: Part 2

Use the Diabetes Readmissions data (on Canvas)

Linear and Quadratic Discriminant Analysis

1. Use the training and test samples for the Diabetes data set that you used for the Logistic Regression analysis. Call them, say, Train and test.
2. Use the Train data set to build Linear Discriminant Analysis (LDA) and Quadratic Discriminant analysis (QDA) models.
 - a. Use the "Readmitted" variable as the dependent variable.
 - b. **NOTE Be careful with the categorical predictors. Since LDA and QDA use pooled or within group covariance matrices, some of the dummies might give problems due to linear dependencies or taking on constant values across the groups. You may have to add predictors one at a time to assess which dummy predictor (or predictors create problems so that you don't use them in the final model).**
3. Perform Test validation testing of LDA and QDA using the Test sample.
4. Summarize your results – for both Training and Test. Which of the models yielded the best predictions overall, and which models yielded the best predictions of "Yes"? Do you like the model? Did it predict as well as you expected it to predict?

Total points: 5.