STA321 Week #7 Assignment

Multiple Logistic Regression Modeling

This week's assignment focuses on multiple logistic regression modeling using the same data set you used in the previous week. To be more specific, your data set has to meet the following requirements:

- The response variable must be binary. As I mentioned in the previous assignment, you can make a binary response variable from a continuous response variable dichotomization.

- At least two continuous predictor variables

- At least two categorical predictor variables

Components of the analysis report

The report should contain the same components as I included in the case study in this week's class note. Please keep in mind that the interpretation of results is VERY important.

- Description of your data set and variables

- Research questions

- Data management and variable inspection

    - variable creation based on existing variables

    - variable transformation

    - variable discretization

    - handling sparse categorical variables

- model building process

    - candidate models

    - manual variable selection

    - automatic variable selection

    - final model identification

    - summarize the inferential statistics in the final model.

- Conclusion and discussion

Remarks:

1. This week's assignment focuses only on the association analysis.

2. Convert the regression coefficients in the final to odds ratio and then provide practical interpretation.

3. The global goodness-of-fit measures (deviance, AIC, etc) in all candidate models should be used only for model selection.