CSC 400, Exam #1, At Home, Spring 2024   Name _____

Instructions: This exam is in two parts: Part I is to be completed partly at home using the materials posted in the course for the at-home portion and you will answer questions about that work during the in-class portion of the exam; Part II is to be completed entirely in class. You may not use cell phones, and you may only access internet resources you are specifically directed to use.

At home, prepare for questions in Part I using R. Complete the calculations noted below. You will be asked for additional analysis and interpretation of this data in the in-class portion of the test. Print out the results of your analysis and code, and bring the pages with you to the exam. You will submit all this work along with the in-class exam.

Use the data on wine to complete the following tasks after importing the data **400exam1data.xlsx** in the file into R.
1. Create a correlation table of the variables for both Pearson and Spearman correlations. Make a correlation plot and a pairplot of the data set.

2. Create a classification model for the data set (you do not need to separate the data into test and training sets for this) using the following algorithms: Decision tree, KNN and LDA (Linear Discriminant Analysis).

3. For each model identified above, create a confusion matrix, and create appropriate model or diagnostic graphs.

Import the dataset **PimaIndiansDiabetes** from the mlbench package.
4. Create an SVM model of the classification task. (You do not need to separate into testing and training sets for this task.)

5. Create a simple neural network model of the classification task.

6. Create a model for the classification task using XGBoost.

7. Create a confusion matrix for each of the models above. Create appropriate diagnostic graphs.

Import the dataset **IncomeESL** from the arules package.
8. Perform association rule mining using the Apriori algorithm on the dataset after first removing the ethnic classification column. Include in your analysis the top 5 rules identified by lift, and by support. You'll be asked about these during the in-class portion of the exam. Be sure to include in your analysis appropriate visualizations.

Neural networks:
9. For this exam, be prepared to discuss:
    a. General applications of neural networks
    b. Convolutional Neural Networks, Long Short-term Memory Networks, Graph Neural Networks