Assignment No. 2

EECS 690

Introduction to Machine Learning

Due: 11:59 PM, Friday, September 18, 2020

Submit deliverables in a single zip file to BlackBoard

Name of the zip file: FirstnameLastname\_Assignment2

Name of the Assignment folder within the zip file: FirstnameLastname\_Assignment2

Deliverables:

1. Copy of Rubric2.docx with your name and ID filled out (do not submit a PDF)
2. Python source code for CompareMLModels
3. Screen print showing the successful execution of CompareMLModels
4. Answers to the following questions:
   1. Based on accuracy which model is the best one?
   2. For each of the 6 other models, explain why you think it does not perform as well as the best one.

Assignment:

* For this assignment, we are going to compare how well different ML classifiers classify the iris dataset.
* Write a Python program called CompareMLModels that does the following:
  + Uses 2-fold cross-validation to produce a test set of 150 samples of the iris data set with the following ML models:
    - Naïve Baysian (NBClassifier)
    - Linear regression (LinearRegression)
    - Polynomial of degree 2 regression (LinearRegression)
    - Polynomial of degree 3 regression (LinearRegression)
    - kNN (KNeighborsClassifier)
    - LDA (LinearDiscriminantAnalysis)
    - QDA (QuadraticDiscriminantAnalysis)
    - Remember 2-fold cross-validation involves:
      * Dividing the data set into 2 folds
      * Training the model with fold 1
      * Testing the model with fold 2
      * Training the model with fold 2
      * Testing the model with fold 1
      * Concatenating the test results from the 2 folds to get a test set of 150 samples.
  + For each of the 7 models the program should display (with a label before each model’s display indicating which model the results are for):
    - Confusion matrix
    - Accuracy metric
  + If the values in your confusion matrices do not add up to 150, then you did something wrong.

Remember:

* Your Programming Assignments are individual-effort.
* You can brainstorm with other students and help them work through problems in their programs, but everyone should have their own unique assignment programs.