Microsoft: DAT203x Data Science and Machine Learning Essentials

QUESTION 2 (1/1 point)

When examining the results of cross validation for a machine learning model you notice that the following conditions are true:

- The values of the metrics are similar across the folds.
- The standard deviation of the metrics is small compared to the mean values.
- The mean values of the metrics are in an acceptable range.

Given these conditions you can conclude that which of the following statements is true?

- The model is sensitive to the datasets used to train and test it.
- The model exhibits a poor fit to the data.
- The model performance requires improvement.
- The performance of the model should generalize well. ✔

EXPLANATION

A model which generalizes well should show consistent performance metrics across the folds in cross validation. The consistency indicates that the model generalizes well since the performance is insensitive to the test data set in each of the folds. One measure of the consistency in the performance metrics over the folds is the standard deviation being significantly smaller than the mean for each metric. Finally, the mean performance metrics should exhibit acceptable values.

You have used 1 of 2 submissions

Module 4 Review
Homework due Oct 30, 2015 at 00:00 UTC