



- ▶ Before You Start
- ▶ Module 1: Introduction and Data Science Theory
- ▶ Module 2: Working with Data
- ▶ Module 3: Visualization, and Building and Evaluating Models
- ▼ **Module 4: Regression, Classification, and Unsupervised Learning**
- Chapter 16: Regression Modeling
- Lab 4A: Working with Regression Models
- Chapter 17: Classification Modeling
- Lab 4B: Working with Classification Models
- Chapter 18: Unsupervised Learning Models
- Lab 4C: Working with Unsupervised Learning Models

QUESTION 3 (1/1 point)

While exploring a dataset for training a two-class (binary) classification problem, you notice the following properties of the data set:

- Certain features exhibit noticeable separation in the values or categories based on the categories of the label.
- Certain features exhibit little separation in the values or categories based on the categories of the label.

Which two of the following feature selection actions should you take before training the model?

☐ All features should be retained in the data set.

☒ The features exhibiting separation should be retained in the dataset. ✓

☒ The features with poor separation should be pruned from the dataset. ✓

☐ Features should be randomly selected to avoid biased in training the model.




Note: Make sure you select all of the correct options—there may be more than one!

EXPLANATION

Features for Two-class classification should exhibit separation in the values or the categories for the two label categories. The separation in the features allows the model to separate the label categories. These features should be retained in the dataset. Labels which exhibit poor separation in the values or categories for the two categories of the label are unlikely to aid in classification. These features should be removed from the data set as they can only add noise or result in poor generalization of the model.

You have used 1 of 2 submissions

Module 4 Review

Homework due Oct 30,
2015 at 00:00 UTC 

© All Rights Reserved



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY
OPENedX

