

**Microsoft: DAT203x Data Science and Machine Learning Essentials**

- ▶ Before You Start
- ▶ Module 1: Introduction and Data Science Theory
- ▶ Module 2: Working with Data
- ▼ **Module 3: Visualization, and Building and Evaluating Models**

Chapter 13: Data Exploration and Visualization

Lab 3A: Exploring and Visualizing Data

Chapter 14: Building Models in Azure ML

Lab 3B: Building Models in Azure ML

Chapter 15: Model Evaluation, Comparison, and Selection

Lab 3C: Evaluating Models in Azure ML

Module 3 Review

Homework due Oct 30, 2015 at 00:00 UTC

- ▶ Module 4: Regression,

QUESTION 4 (1/1 point)

You are training a supervised machine learning model.

Which three kinds of features should be pruned from the dataset?

☒ Features that are collinear or codependent on other features in the dataset. ✓

☐ Features that decrease model error during training and testing.

☒ Features that increase model error during training and testing. ✓

☒ Features that have little impact on model performance during training and testing. ✓

☐ Features that are categorical with numeric values.



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

EXPLANATION

You should prune features that are collinear or codependent on other features in the dataset to prevent bias and instability when training the model.

You should prune features that increase model error during training and testing to improve model performance.

You should prune features that have little impact on model performance during training and testing, to reduce model complexity and improve generalization.

You have used 1 of 2 submissions

Classification,
and
Unsupervised
Learning

© 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

