

# Important Terms for Big Data Analytics and Machine Learning

Last updated Sept 5, 2019

# Regression

- Difference between multiple and multivariate linear regression
- What is one way to measure “goodness of fit” in linear regression?
- How do you know your model has improved your knowledge of the problem?

# Classification Basics

- Supervised vs unsupervised learning
- What is the difference between precision and recall?
- What is specificity?
- What does ROC stand for?
  - Where did it originate?
  - How do you calculate it?
- What does AUC stand for?
  - How do you calculate it? Is there more than one way to do this?
- What is a confusion matrix?