Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE 100%

Data: Skewed Distributions

TOTAL POINTS 2

1. Which of the following statements are true? 1 / 1 point

- ○ There is only one reason your data might be skewed: bad data collection.
- ○ Algorithms make no assumptions about the distribution of the labels.
- ◉ Your data might be skewed along both the label distribution and along the feature space distribution.
- ○ Standardization and normalization can never help with skewed distributions in your data.

✓ Correct

Correct! You need to understand the multiple ways in which your data can be skewed. This can be along both the y-axis and the x-axis.

2. Which of the following statements are true? 1 / 1 point

- ◉ Usually when we talk about skewed distributions in machine learning, we are concerned with the feature space.
- ○ You should always make an informed decision on which learning algorithm to use based on your knowledge of the distribution and never rely on empirically testing different algorithms despite the distribution.
- ○ The distribution of your data does not really matter when you have a small dataset.
- ○ The same data collected from two different sources is guaranteed to have the same distribution.

✓ Correct

Correct! That is because a distributional shift in the feature space from your training data to your testing data puts you into a different learning regime.