Back

Module 2 Graded Quiz - KNN
Graded Assignment • 24 min
Due Oct 6, 11:59 PM PDT

Your grade: 100%

Your latest: **100%** • Your highest: **100%**
To pass you need at least 62%. We keep your highest score.

Next item →

1. Which one of the following statements is true regarding K Nearest Neighbors? 1 / 1 point

- ○ For high dimensional data, the best distance measure to use for KNN is the Euclidean distance.
- ○ The Manhattan distance between two data points is the square root of the sum of the squares of the differences between the individual feature values of the data points.
- ◉ K Nearest Neighbors (KNN) assumes that points which are close together are similar.
- ○ The distance between two data points is independent of the scale of their features.

Correct
Correct. The distance between two given points is the similarity measure in the KNN model, where close points are thought to be similar.

2. Which one of the following statements is most accurate? 1 / 1 point

- ◉ K nearest neighbors (KNN) needs to remember the entire training dataset in order to classify a new data sample.
- ○ Linear regression needs to remember the entire training dataset in order to make a prediction for a new data sample.
- ○ KNN determines which points are closest to a given data point, so it doesn't take long to actually perform prediction.
- ○ KNN only needs to remember the hyperplane coefficients to classify a new data sample.

Correct
Correct. KNN needs to remember all of the points. It needs to remember the entire training set, so it's going to be very memory intensive.

3. Which one of the following statements is most accurate about K Nearest Neighbors (KNN)? 1 / 1 point

- ○ KNN is a classification model.
- ◉ KNN can be used for both classification and regression.
- ○ KNN is an unsupervised learning method.
- ○ KNN is a regression model.

Correct
Correct. KNN is known as a classification model, but can also be used for regression. All you have to do is replace KNeighborsClassifier with KNeighborsRegressor.

4. (True/False) K Nearest Neighbors with large *k* tend to be the best classifiers. 1 / 1 point

- ○ True
- ◉ False

Correct
Correct! K Nearest Neighbors with high values of k might likely not generalize well with new data. A best practice is to use the elbow method to find a model with low k and high decrease in error.

5. When building a KNN classifier for a variable with 2 classes, it is advantageous to set the neighbor count k to an odd number. 1 / 1 point

- ◉ True
- ○ False

Correct
Correct! An odd neighbor count works as a tie breaker. It ensures there cannot be a tie in the number of n nearest neighbors for two given classes. You can find more information on the k nearest neighbor lesson.

6. The Euclidean distance between two points will always be shorter than or equal to the Manhattan distance. 1 / 1 point

- ◉ True
- ○ False

Correct
Correct! From trigonometry, you should realize that Euclidian distance is shorter than the Manhattan distance. You can review this on the K Nearest Neighbors lesson.

7. The main purpose of scaling features before fitting a k nearest neighbor model is to: 1 / 1 point

- ○ Ensure decision boundaries have roughly the same size for all classes
- ○ Help find the appropriate value of k
- ○ Break ties in case there is the same number of neighbors of different classes next to a given observation
- ◉ Ensure that features have similar influence on the distance calculation

Correct
Correct! You can find more information in the K Nearest Neighbor lesson.

8. These are all pros of the k nearest neighbor algorithm EXCEPT: 1 / 1 point

- ○ It is simple to implement as it does not require parameter estimation
- ◉ It is sensitive to the curse of dimensionality
- ○ It adapt wells to new training data
- ○ It is easy to interpret

Correct
Correct! You can find more information in the K Nearest Neighbor lesson.