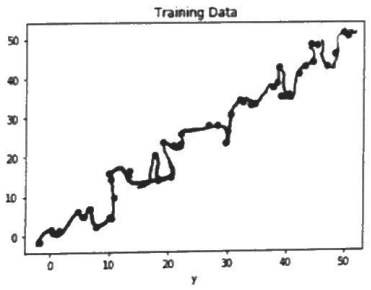
CS691 Quiz 3 Name Solution.

1. Give an example where K-NN and K-Means would give the same result. Be as specific as you can. (4 points)

K-NN with K=1 would have a boundary of A.

K means with K=2 and inital points of b and c would give a boundary of A assuming all + and - points are equidistant.

2. Draw a curve that is overfit to the data below. Explain why it is overfit. (2 points)



It perfectly fits the curve meaning it won't generalize well to future data.

3. Give an example of noise at the feature level. Explain how that example demonstrates feature noise. (2 points)

If I have a camera and it glitches while I am taking a picture, the pixels will be different than the truth. This is feature noise because we use pixel values as features.

4. Hyperparameters are chosen by analyzing the test data. True or False. Circle one and explain. (2 points)

Validation data is used.

Don't touch the test data.

CS491 Quiz 3 Name Solution

1. We use the validation data to determine the max depth of a decision tree. True or False. Circle one and explain. (2 points)

The max depth is a hyper parameter. We use validation to set this.

2. What are two things K-NN and K-Means have in common? What is something that makes them different? (3 points).

Same

1. Both have K as a hyperparameter.
2. Both treat all features equally.

Diff

1. KNN Supervised Kmeans unsupervised.

3. Answer the following questions given the training data below. Each sample is listed as $s = (x_1, x_2, y)$, where x_1 and x_2 are the features and y is the label. (5 points)

$s_1 = (2, 2, -1)$
$s_2 = (0, 0, 1)$
$s_3 = (3, 5, -1)$
$s_4 = (1, 1, 1)$
$s_5 = (1, 0, 1)$
$s_6 = (3, 2, -1)$
$s_7 = (2, 4, -1)$
$s_8 = (0, 1, 1)$

(a) How would a 1-NN classify the following sample? $s_t = (3, 3)$ -1
(b) How would a 1-NN classify the following sample? $s_t = (-1, -1)$ 1
(c) How would a 3-NN classify the following sample? $s_t = (1, 2)$ 1
(d) How would a 5-NN classify the following sample? $s_t = (3, 1)$ 1
(e) How would a 5-NN classify the following sample? $s_t = (0.5, 0.5)$ 1