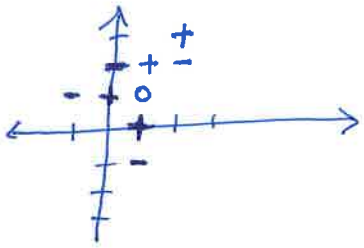
CS 422 Quiz 3 Name Solution

1. Given the training data below, how would 3-NN classify the sample (1,1)? What about 7-NN? (4 points)

Sample	x_1	x_2	label
s_1	-1	1	-1
s_2	0	1	1
s_3	0	2	-1
s_4	1	-1	-1
s_5	1	0	1
s_6	1	2	1
s_7	2	2	-1
s_8	2	3	1

Easiest way to solve is to plot.



3-NN would give +

7-NN would give −

2. When you find noise in data which of the following options would you consider in K-NN? Explain. (3 points)

(a) Increase the value of K.
(b) Decrease the value of K.
(c) Noise has no effect on my choice of K.
(d) None of the above.

Increasing the value of K reduces the effect of noise, effectively drowning it out.

3. True/False: The computational complexity of K-NN for classifying new samples grows linearly with the number of samples in the training dataset in the worst-case scenario. Explain. (3 points)

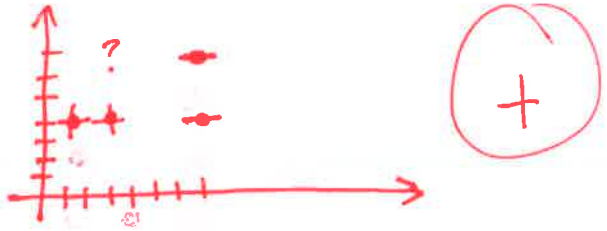
Yes. You have to check every sample against your test sample. ~~Each time you check a training sample~~

So each time you test you check all N training samples.

CS 622 Quiz 3 Name Solution

1. Given the training data below, how would K-NN classify the following sample: (3,7)? Assume K=3. (2 points)

Sample	x_1	x_2	label
s_1	7	7	-1
s_2	7	4	-1
s_3	3	4	1
s_4	1	4	1



2. True/False: K-NN can be used to solve regression problems (predict real values). Explain. (3 points)

Instead of summing labels & returning the sign, we can average the real-valued labels.

3. When you find noise in data which of the following options would you consider in K-NN? Explain. (3 points)

(a) Increase the value of K.
(b) Decrease the value of K.
(c) Noise has no effect on my choice of K.
(d) None of the above.

Same as 422 #2.

4. True/False: K-NN is immediately adapts as we collect new training data. Explain. (2 points)

New training data directly affects predictions.
We calculate distances between test points & all training points.