CS 622 Quiz 3 Name Solution

1. True/(False.) There is no possibility of a tie in K-NN when K is odd. Briefly explain. (2 points)

When $k=1$ if two points are equidistant from the test point and they have different labels, there will be a tie.

2. Given the training data in table 2, what value of K would you choose for K-NN? Justify your choice. (3 points).

x_1	x_2	label
0	0	-1
0	1	1
1	0	1
1	1	1

Table 2: Training data.

$K=1$ Because otherwise (e.g. $k\geq 3$) every test sample would be predicted as positive.

3. Name **two** ways in which K-NN and K-Means are similar and **three** ways in which they differ. (5 points)

Similar
+ Both have hyperparameter K.
+ Both use distance as a measure of similarity.

Different
- K-NN is supervised & K-means is unsupervised.
- No training involved in K-NN / no model.
- K-NN has high computational expense at test time. Calculate N distances. K-means only needs K distances.

CS 422 Quiz 3 Name Solution

1. Name **two** ways in which K-NN and K-Means are similar and **two** ways in which they differ. (4 points)

See 622 #3

2. Given the training data in table 1, for what value of K would K-NN classify the following test point as 1? $t = (-1, 0)$. For what value of K would K-NN classify the test point as -1? (2 points)

$K \geq 3$

x_1	x_2	label
0	0	-1
0	1	1
1	0	1
1	1	1

Table 1: Training data.

$K = 1$

3. What value of K would be overfitting and what value would be underfitting for K-NN and K-Means respectively? Briefly explain. (4 points)

K-NN	K=1 (each point predicted by 1 neighbor)	K=N	always predict same class
K-Means	K=N	K=1	one cluster of all data
	over	under	

each cluster contains 1 point.