Quiz 5

1.
Which of the following statements is true for k-NN classifiers

○	The classification accuracy is better with larger values of k
○	The decision boundary is smoother with smaller values of k
◉	k-NN does not require an explicit training step
○	The decision boundary is linear

2.
What is the best approach to solve this question:

> **What is the average weekly salary of all female employees under forty years of age?**

○	Supervised learning
○	Unsupervised clustering
◉	Data query

3.
A company has build a kNN classifier that gets 100% accuracy on training data. When they deployed this model on client side it has been found that the model is not at all accurate.

Which of the followings could be the reason:

○	None of these
◉	It is probably a overfitted model
○	It is probably a underfitted model

4.

Considering the following training set of m = 4 training examples:

x	y
1	0.5
2	1
4	2
0	0

Consider the linear regression model $h_\theta(x) = \theta_0 + \theta_1 x$.

What are the values of θ_0 and θ_1 that you would expect to obtain upon running gradient descent on this model? (Linear regression will be able to fit this data perfectly.)

○	$\theta_0 = 1, \theta_1 = 0.5$
○	$\theta_0 = 0.5, \theta_1 = 0$
◉	$\theta_0 = 0, \theta_1 = 0.5$
○	$\theta_0 = 0.5, \theta_1 = 1$

5.

For which of following tasks might K-means clustering be a suitable algorithm?

○	Given historical weather records, predict if tomorrow's weather will be sunny or rainy
◉	Given a set of news articles from many different websites, find out what topics are the main topics covered
○	Given sales data from a large number of products in a supermarket, estimate future sales for each of these products.