Name:_____ _____ Reg #:_____
Section:___

Data Science Quiz # 2: **Marks: 10**
2017-04-05

Q1: Multiclass Classification: (1+2+1 Marks)

(A) We have to Draw the plot for the training data, where each class should be represented by a different symbol. (keep $\mathbf{x_1}$ on x-axis and $\mathbf{x_2}$ on y-axis).

(B) First we will label one of the class as positive and all rest data as negative and we will train a logistic regression classifier (h(x)).
Then take 2nd class as positive and all rest data as negative and we will train a logistic regression classifier (h(x)).
Then take 3rd class as positive and all rest data as negative and we will train a logistic regression classifier (h(x)).

So to conclude we will train separate model h(x) for each class. So we will have three h(x).

(C) On new input x (new patient), how we will predict if the patient has “flu”, “cold” or is “not-ill”. We will take x as input and see output for each trained model (h(x)), the new patient x will belong to the class with highest h(x) value.

Q2.

Solution: (d)

Q3. For which of the following tasks might K-means clustering be a suitable algorithm? Select all that
apply. **[1 mark]**

Solution: A) and B)

Q4.

Solution:

$$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$
$$\text{temp1} := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$
$$\theta_0 := \text{temp0}$$
$$\theta_1 := \text{temp1}$$

Name:_____ _____ Reg #:_____
Section:___

Q5.

Solution: D

Q6.

Solution: Simple model because complex model can lead to overfitting problem.