**Machine Learning: Dr. Shahid Mahmood Awan**

Quiz #1

Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ ID:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Question 1**

A computer program is said to learn from experience E with respect to some task T and some performance measure P if its performance on T, as measured by P, improves with experience E. Suppose we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather. What would be a reasonable choice for P?

The process of the algorithm examining a large amount of historical weather data.

The probability of it correctly predicting a future date's weather.

None of these.

The weather prediction task.

**Question 2**

Suppose you are working on weather prediction, and use a learning algorithm to predict tomorrow's temperature (in degrees Centigrade/Fahrenheit). Would you treat this as a classification or a regression problem?

Regression

Classification

**Question 3**

Suppose you are working on stock market prediction. You would like to predict whether or not a certain company will win a patent infringement lawsuit (by training on data of companies that had to defend against similar lawsuits). Would you treat this as a classification or a regression problem?

Classification

Regression

**Question 4**

Some of the problems below are best addressed using a supervised learning algorithm, and the others with an unsupervised learning algorithm. Which of the following would you apply supervised learning to? (Select all that apply.) In each case, assume some appropriate dataset is available for your algorithm to learn from.

In farming, given data on crop yields over the last 50 years, learn to predict next year's crop yields.

Given a large dataset of medical records from patients suffering from heart disease, try to learn whether there might be different clusters of such patients for which we might tailor separate treatments.

Given data on how 1000 medical patients respond to an experimental drug (such as effectiveness of the treatment, side effects, etc.), discover whether there are different categories or "types" of patients in terms of how they respond to the drug, and if so what these categories are.

Examine a web page, and classify whether the content on the web page should be considered "child friendly" (e.g., non-pornographic, etc.) or "adult."

**Question 5**

Which of these is a reasonable definition of machine learning?

Machine learning is the field of allowing robots to act intelligently.

Machine learning is the science of programming computers.

Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.

Machine learning learns from labeled data.

**Machine Learning: Dr. Shahid Mahmood Awan**

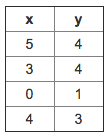
Quiz #2

## 1. Question 1

Consider the problem of predicting how well a student does in her second year of college/university, given how well she did in her first year.

Specifically, let x be equal to the number of "A" grades (including A-. A and A+ grades) that a student receives in their first year of college (freshmen year). We would like to predict the value of y, which we define as the number of "A" grades they get in their second year (sophomore year).

Here each row is one training example. Recall that in linear regression, our hypothesis is *hθ*(*x*)=*θ*0+*θ*1*x*, and we use *m* to denote the number of training examples.



For the training set given above (note that this training set may also be referenced in other questions in this quiz), what is the value of *m*? In the box below, please enter your answer (which should be a number between 0 and 10).



## 2. Question 2

For this question, assume that we are using the training set from Q1. Recall our definition of the cost function was 

What is *J*(0,1)? In the box below, please enter your answer (Simplify fractions to decimals when entering answer, and '.' as the decimal delimiter e.g., 1.5).



## 3. Question 3

Suppose we set ***θ*0=−1,*θ*1=2** in the linear regression hypothesis from Q1. What is ***hθ*(6)?**



## 4. Question 4

Let *f* be some function so that *f*(*θ*0,*θ*1) outputs a number. For this problem, *f* is some arbitrary/unknown smooth function (not necessarily the cost function of linear regression, so *f* may have local optima). Suppose we use gradient descent to try to minimize *f*(*θ*0,*θ*1) as a function of *θ*0 and *θ*1. Which of the following statements are true? (Check all that apply.)

If the learning rate is too small, then gradient descent may take a very long time to converge.

 Even if the learning rate *α* is very large, every iteration of gradient descent will decrease the value of *f*(*θ*0,*θ*1).

 If *θ*0 and *θ*1 are initialized so that *θ*0=*θ*1, then by symmetry (because we do simultaneous updates to the two parameters), after one iteration of gradient descent, we will still have *θ*0=*θ*1.

If *θ*0 and *θ*1 are initialized at a local minimum, then one iteration will not change their values.

## 5. Question 5

Suppose that for some linear regression problem (say, predicting housing prices as in the lecture), we have some training set, and for our training set we managed to find some *θ*0, *θ*1 such that *J*(*θ*0,*θ*1)=0.

Which of the statements below must then be true? (Check all that apply.)

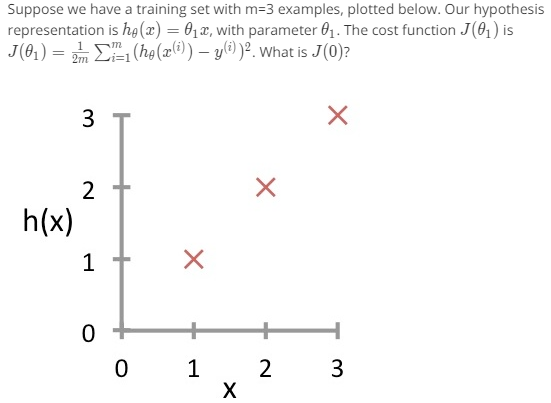
For this to be true, we must have *y*(*i*)=0 for every value of *i*=1,2,…,*m*.

For this to be true, we must have *θ*0=0 and *θ*1=0 so that *hθ*(*x*)=0

Our training set can be fit perfectly by a straight line, i.e., all of our training examples lie perfectly on some straight line.

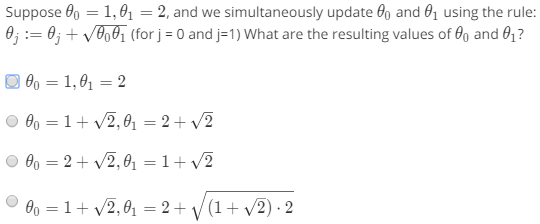
 Gradient descent is likely to get stuck at a local minimum and fail to find the global minimum.

Q1

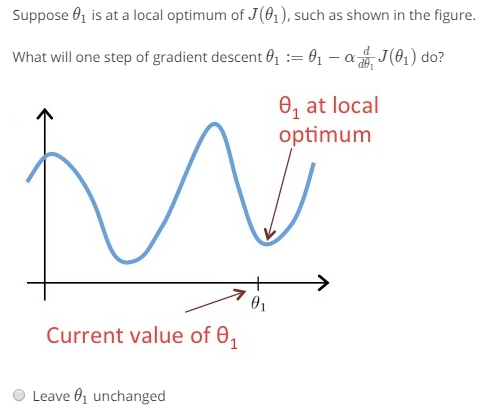


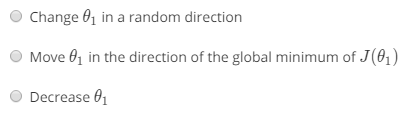
* 0
* 1
* 1/6
* 14/6

Q3.



Q3





Q4

