The Coders' Club

Machine Learning: G1

Week 1: Assignment

Topics:

- Supervised Learning: Classification & Regression
- Linear Regression with One Variable
- Linear Algebra
- Miscellaneous

Instructions:

- Send your answers on thecodersclub2017@gmail.com strictly in PDF format.
- Show necessary calculations and steps for numerical questions.
- Deadline: 19/12/2019

Supervised Learning: Classification & Regression

Q.1. You are given with statistical data of head-to-head matches played between India and Pakistan

The problem is to predict the probability of either of the teams winning the next match to be played.

This is an example of -

- Classification
- Regression
- K-means Clustering

Q.2. Which of the following are regression problems? Check all that apply.

- To predict the age of a person.
- To predict whether the monsoon will be scarce or normal in 2020.
- To predict whether CST-Panvel train will be late or not.
- To predict your CGPA in next semester.

Q.3. Which of the following are true about Supervised Learning? Check all that apply.

- It is used to predict the future outcome based on the historical data.
- The model learns through observation and find structures/patterns in the data.
- Models are given clear instructions right from the beginning as in what needs to be learnt and how it needs to be learnt.
- The model is trained by trial and error method.

Q.4. In which problem, the predicted result is mapped into discrete categories?

- Regression
- K-means Clustering
- Classification

Q.5. 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 regression problem?

- Regression
- Classification

Linear Regression with One Variable

Q.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. We would like to predict the value of y, which we define as the number of “A” grades they get in their second year.

Refer to the following training set of a small sample of different students’ performances. Each row represents a training set example. Recall that in linear regression, our hypothesis is $h_\theta(x)= \theta_0 + \theta_1x$, and we use m to denote the number of training examples.

x	y
3	4
2	1
4	3
0	1

For the training set given above, what is the value of m?

Q.2. Many substances that can burn (such as gasoline and alcohol) have a chemical structure based on carbon atoms; for this reason they are called hydrocarbons. A chemist wants to understand how the number of carbon atoms in a molecule affects how much energy is released when that molecule combusts (meaning that it is burned). The chemist obtains the dataset below. In the column on the right, “kJ/mol” is the unit measuring the amount of energy released.

Name of molecule	No. of hydrocarbons in molecule (x)	Heat released when burned (kJ/mol) (y)
methane	1	-890
ethene	2	-1411
ethane	2	-1560
propane	3	-2220
cyclopropane	3	-2091
butane	4	-2878
pentane	5	-3537
benzene	6	-3268
cyclohexane	6	-3920
hexane	6	-4163
octane	8	-5471
naphthalene	10	-5157

You would like to use linear regression $h_\theta(x) = \theta_0 + \theta_1 x$, to estimate the amount of energy (y) released as a function of the no. of carbon atoms (x). Which of the following do you think will be the values you obtain for θ_0 and θ_1?

- $\theta_0 = -1780.0$, $\theta_1 = 530.9$
- $\theta_0 = -569.6$, $\theta_1 = 530.9$
- $\theta_0 = -1780.0$, $\theta_1 = -530.9$
- $\theta_0 = -569.6$, $\theta_1 = -530.9$

Q.3. Consider 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 algorithm on this model? (Note that Linear Regression will be able to fit this data perfectly).

- $\theta_0 = 1$, $\theta_1 = 0.5$
- $\theta_0 = 1$, $\theta_1 = 1$
- $\theta_0 = 0.5$, $\theta_1 = 0$
- $\theta_0 = 0.5$, $\theta_1 = 0.5$
- $\theta_0 = 0$, $\theta_1 = 0.5$

Q.4. Suppose we set $\theta_0 = -1$, $\theta_1 = -2$ in the linear regression hypothesis. What is $h_\theta(6)$?

Q.5. Suppose that for some linear regression problem (say, predicting house prices), we have some training set, and for our training set we managed to find out some θ_0 and θ_1 such that $J(\theta_0 , \theta_1) = 0$.
Which of the following options are true? Check all that apply.

- Our training set can be fit perfectly in a straight line i.e. all our training examples lie perfectly on some straight line.
- For this to be true, we must have $\theta_0 = 0$ and $\theta_1 = 0$ so that $h_\theta(x)=0$
- For this to be true, we must have $y^{(i)} = 0$ for every value of i = 1, 2, 3, … m.
- Gradient Descent is likely to get stuck at a local minimum and fail to find the global minimum.

Q.6. Let f be a function so that $f(\theta_0 , \theta_1)$ outputs a number. For this problem, f is some arbitrary smooth function (not necessarily the cost function of linear regression, so f may have local optima). Suppose we use gradient descent to try to minimise $f(\theta_0 , \theta_1)$ 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.
- If θ_0 and θ_1 are initialized so that $\theta_0 = \theta_1$, then by some symmetry (because we do simultaneous updates to the two parameters), after one iteration of gradient descent, we will still have $\theta_0 = \theta_1$.
- If θ_0 and θ_1 are initialized at a local minimum, then one iteration will not change their values.
- Even if the learning rate α is very large, every iteration of gradient descent will decrease the value of $f(\theta_0 , \theta_1)$.

Linear Algebra

Q.1 Which of the following statements are true? Check all that apply.

$\begin{bmatrix} 1 & 2 \\ 4 & 0 \\ 0 & 1 \end{bmatrix}$ is a 3×2 matrix.

$\begin{bmatrix} 0 & 1 & 4 & 2 \\ 3 & 4 & 0 & 9 \end{bmatrix}$ is a 4×2 matrix.

$\begin{bmatrix} 0 & 4 & 2 \\ 3 & 4 & 9 \\ 5 & -1 & 0 \end{bmatrix}$ is a 3×3 matrix.

$\begin{bmatrix} 1 & 2 \end{bmatrix}$ is a 1×2 matrix.

Q.2. Let A be a matrix as shown below. A_{32} is one of the elements of this matrix.

$$A = \begin{bmatrix} 85 & 76 & 66 & 5 \\ 94 & 75 & 18 & 28 \\ 68 & 40 & 71 & 5 \end{bmatrix}$$

What is the value of A_{32}?

Q.3. What is

$$\begin{bmatrix} 8 & 6 & 9 \\ 10 & 1 & 10 \end{bmatrix} + \begin{bmatrix} 3 & 10 & 2 \\ 6 & 1 & -1 \end{bmatrix}?$$

Q.4. What is

$$2 \times \begin{bmatrix} 4 & 5 \\ 1 & 7 \end{bmatrix}?$$

Q.5. What is

$$\begin{bmatrix} 4 \\ 6 \\ 7 \end{bmatrix} / 2 - 3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}?$$

Q.6. Consider the product of these two matrices.

$$\begin{bmatrix} 1 & 2 & 1 & 5 \\ 0 & 3 & 0 & 4 \\ -1 & -2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

Find the value of the resultant matrix. Also find its dimensions.

Q.7. What is

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 5 \\ 3 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}?$$

Q.8. In the equation

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ a & b \\ c & d \end{bmatrix}$$

Find the values of a, b, c and d.

Q.9. What is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}?$$

Q.10. What is

$$\begin{bmatrix} 0 & 3 \\ 1 & 4 \end{bmatrix}^T?$$

Q.11. Let two matrices be

$$A = \begin{bmatrix} 4 & 3 \\ 6 & 9 \end{bmatrix}, \qquad B = \begin{bmatrix} -2 & 9 \\ -5 & 2 \end{bmatrix}$$

What is A + B ?

Q.12.

$$\textbf{Let } x = \begin{bmatrix} 2 \\ 7 \\ 4 \\ 1 \end{bmatrix}$$

What is ½ * x ?

Q.13. Let u be a 3D vector, where specifically

$$u = \begin{bmatrix} 5 \\ 1 \\ 9 \end{bmatrix}$$

Find u^T.

Q.14. Let u and v be two 3D vectors where,

$$u = \begin{bmatrix} -3 \\ 4 \\ 3 \end{bmatrix} \qquad v = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$$

Find u^Tv.

Q.15. Let A and B be two 3x3 matrices. Which of the following must necessarily hold true? Check all that apply.

- If A is a 3x3 identity matrix, then A * B = B * A
- A + B = B + A
- A * B * A = B * A * B
- A * B = B * A

Miscellaneous

Q.1. Give any two examples each of Supervised and Unsupervised Machine Learning you have observed.

Q.2. What is **“Batch Gradient Descent”** ?

Q.3. Write the basic steps of Gradient Descent algorithm.

Q.4. Write down the equations/formulae for the following:

(i) Hypothesis for Linear Regression
(ii) Cost function
(iii) Gradient Descent
(iv) Gradient Descent with Linear Regression