Sheet 1

1 K Nearest Neighbor (KNN)

The results of the exam were recorded along with some data about the students. The results can be found in the table below.

ID	No. Programming Languages used	No. delivered assignments	GPA	Passed exam
1	3	5	3.1	Yes
2	3	5	2.0	No
3	2	7	3.5	Yes
4	4	9	2.5	Yes
5	3	11	3.9	No
6	2	3	2.9	No
7	3	3	1.9	No
8	4	9	3.2	Yes

1. Knowing that Passed exam is the class label, predict the class of the following students using K-Nearest Neighbor for K = 3.

ID	No. Programming Languages used	No. delivered assignments	GPA	Passed exam
9	2	2	3.0	?
10	3	3	4.0	?
11	4	6	2.0	?
12	2	5	3.5	?

2. Show the effect of feature scaling by transforming each feature(column) to be in the range [0,1]. The min value of the feature in the training data maps to 0 and the max value of same feature maps to 1. Predict the labels again after scaling.

2 Linear Regression

1. Find the least square regression line for the following set of data

$$\{(-1\ ,\ 0),(0\ ,\ 2),(1\ ,\ 4),(2\ ,\ 5)\}$$

Then plot the given points and the regression line.

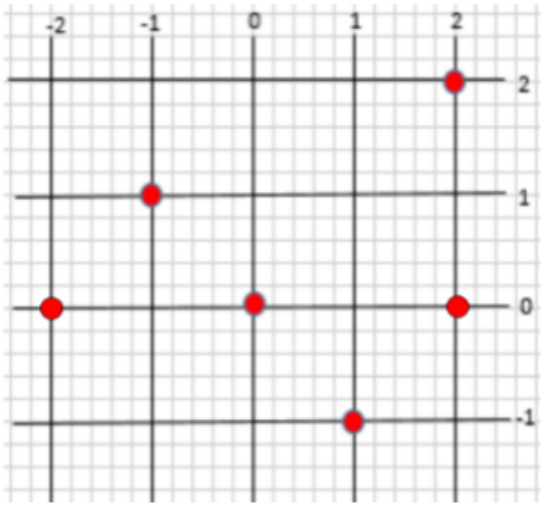
2. The values of x and their corresponding values of y are shown in the table below

x	0	1	2	3	4
y	2	3	5	4	6

(a) Find the least square regression line $y = a\,x + b$.

(b) Estimate the value of y when $x = 10$.

3.



A. Apply linear regression analytic form, use matrix inverse, to **find** the parameters of the best-fit line through the 6 points{(x,y)}= {(2,2), (0,0), (-1,1), (1,-1), (-2,0),(2,0)} shown in the figure.
B. **Draw** the best-fit line on the answer sheet.
C. **Find** the sum of the squared loss.
D. **Discuss** how sensitive is the linear regression to the noise. **Illustrate** your answer by finding the best model if we consider the point at (2,2) as an outlier.
E. **Estimate** the y for x= - 0.5, x = 0.5 and for x = 1.5.

3 Logistic Regression

1. State True or False

   (a) A perceptron is guaranteed to perfectly learn a given linearly separable function within a finite number of training steps.

   (b) A single perceptron can compute the XOR function.

   (c) Each of the inputs to the perceptron is multiplied by a number to give it a weight. These weights allow the strength of the different connections to be changed so that the perceptron can learn.

2. Design one layer perceptron to represent the following logic functions:

   (a) AND

   (b) OR

   (c) NOT

3. A perceptron has input weights w1 = 3.1 and w2 = - 1.9 and a threshold value T = 0.4. What output does it give for the input x1 = 1.2, x2 = 2.3?

4. Suppose the logistic regression classifier model is given as

$$p(y = 1|x) = g(w_0 + w_1x_1 + w_2x_2)$$

where, g is sigmoid function.

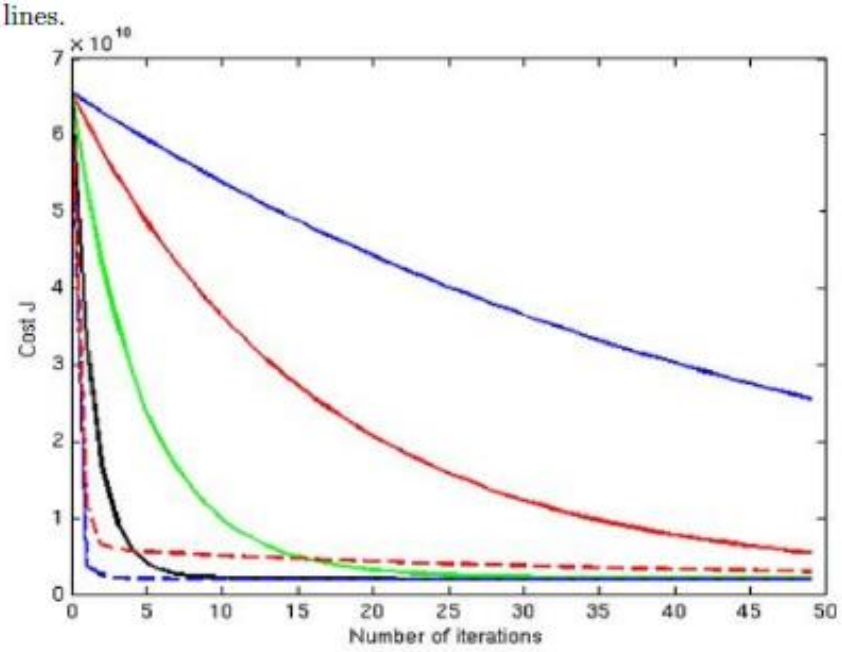
Plot the decision boundary in each of the following weights assignment:

(a) $w_0 = 6, w_1 = 0, w_2 = -1$

(b) $w_0 = 8, w_1 = 2, w_2 = 0$

(c) $w_0 = -4, w_1 = 1, w_2 = 1$

5. Suppose the following figure represents the change of the cost function during the training iterations for 3 different learning rates represented by the 3 blue, red and green undashed lines.



Arrange the 3 learning rates on ascending order explaining your intuition.

6. Consider the following 2-dimensional dataset. We have 8 points and 2 classes {+, -}

Feature 1	Feature 2	Class
-1	0	+
-0.5	0.5	+
2	2	-
0.5	1	-
1	0	+
1	-1	+
0	3	-
3	1	-

Apply 3 epochs to fit this data using logistic regression model.
The initial weights will be $w_0 = -2, w_1 = -2, w_2 = 1$
Report the training accuracy before training then after each epoch.