Question-1

1.Machine Learning is the field of study that gives computers the

ability to learn without being explicitly programmed.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.

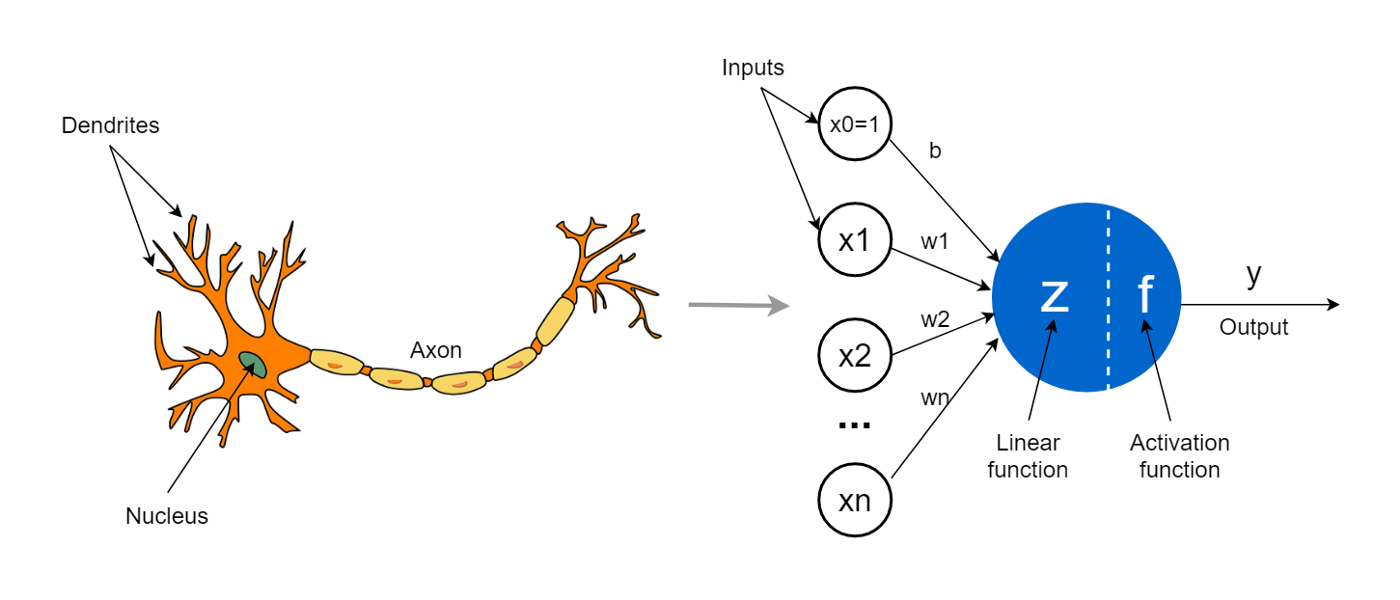
2.Supervised Learning is type of machine learning in which machines are trained using well "labelled" training data, and on basis od that data, machines predict output.This input/output pair is usually referred to as "labeled data".As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately, which occurs as part of the cross validation process.

3.Unsupervised learning is an algorithm that learns from the data which is neither classified nor labeled or raw data to analyze and cluster them.These algorithms discover hidden patterns or data groupings without the need for human intervention.It's ability to discover similarities and differences in information make it the ideal solution for exploratory data analysis, cross-selling strategies, customer segmentation, image recognition and extracting generative features.

4.Reinforcement learning is the training of machine learning models to make a sequence of decisions.To get the machine to do what the programmer wants, the algorithm gets either rewards or penalities for the actions it performs.It's goal is to maximize the total reward.The environment is going to give the agent a state S(t) and then the agent is going to take an action A(t) and the environment is going to give back a reward to the agent which is R(t) and then next state S(t+1) and this keeps going on loop unitl the environment gives back a terminal state which then ends the episode.

5.A series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Artificial neural networks (ANNs) are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer.Each node, or artifical neuron, connects toa nether and has an associated weight and threshold. If the output opf any individual node is above the specified threshold value,that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.

6.An artificial neuron is a digital construct that seeks to stimulate the behavior of a biological neuron in the brain. Artificial neurons are typically used to make up an artificial neural network these technologies are modeled after human brain activity.A biological neuron recieves its input signals from other neurons through dentrites.Likewise, a perceptron recieves its data from other perceptrons through input neurons that take numbers. The connection points betweem dendrites and biological neurons are called synapses. Lisewise, the connections betweem inputs and perceptrons are called weights.They measure the importance level of each input.



Question-2

## **Perceptron**

Perceptron was introduced by Frank Rosenblatt in 1957. He proposed a Perceptron learning rule based on the original MCP neuron. A Perceptron is an algorithm for supervised learning of binary classifiers. This algorithm enables neurons to learn and processes elements in the training set one at a time.

There are two types of Perceptrons: Single layer and Multilayer.

Single layer - Single layer perceptrons can learn only linearly separable patterns Multilayer - Multilayer perceptrons or feedforward neural networks with two or more layers have the greater processing power The Perceptron algorithm learns the weights for the input signals in order to draw a linear decision boundary.

This enables you to distinguish between the two linearly separable classes +1 and -1.

## **Perceptron Function**

Perceptron is a function that maps its input “x,” which is multiplied with the learned weight coefficient; an output value ”f(x)”is generated.

f(x) = {lifw,c + b>0., 0otherwise

In the equation given above:

“w” = vector of real-valued weights “b” = bias (an element that adjusts the boundary away from origin without any dependence on the input value) “x” = vector of input x values Συγά, 2 =1

“m” = number of inputs to the Perceptron The output can be represented as “1” or “0.” It can also be represented as “1” or “-1” depending on which activation function is used.

## **Inputs of a Perceptron**

A Perceptron accepts inputs, moderates them with certain weight values, then applies the transformation function to output the final result.

A Boolean output is based on inputs such as salaried, married, age, past credit profile, etc. It has only two values: Yes and No or True and False. The summation function “∑” multiplies all inputs of “x” by weights “w” and then adds them up as follows:

w0+w1x1+w2x2+.......+wnxn

## **Weights and Bias**

Weights are the co-efficients of the equation which you are trying to resolve. Negative weights reduce the value of an output. weights control of the signal ( or the strength of the connection ) betweem two neurons. In other words, a weight decides how much influence the input will have on the output.

Bias term is an adjustable, numerical term added to a perceptron;s weighted sum of inputs and weights that can increase classification model accuracy. Bias units are not influenced by the previous layer but they do have outgoing connections with their own weights. The bias unit gurantees that even when all the inputs are zeros there will still be an activation in the neuron.

## **Summing Functions of Perceptron**

The perceptron is a simple function that takes several inputs, sums them, and passes them through another function to give a single result. Summing just goes through the inputs and sums them. The function that the sum is passes through is called transfer functions.

## **Activation Functions of Perceptron**

The activation function applies a step rule (convert the numerical output into +1 or -1) to check if the output of the weighting function is greater than zero or not.

activation function (X1W1 + X2W2 + X3W3 + ............. + XnWn + bias)

## **Output of Perceptron**

Perceptron with a Binary output:

Inputs = x1… xn

Output: o(x1… xn)

Weights: wi = contribution of input xi to the Perceptron output;

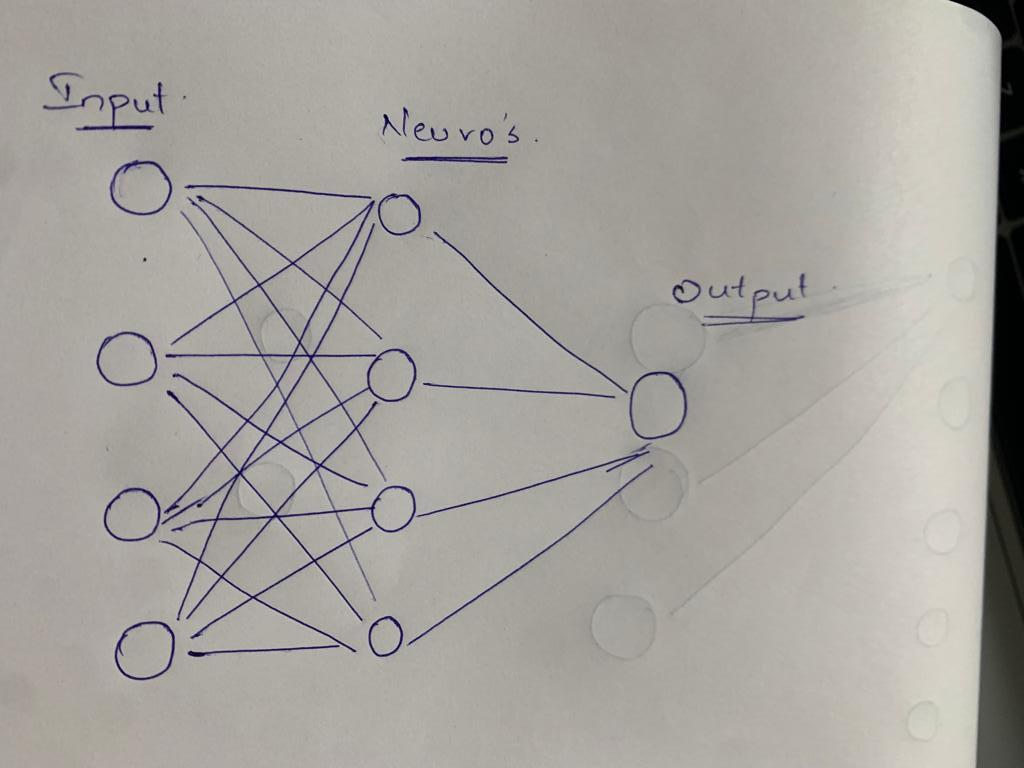
w0 = bias or threshold

An output of +1 specifies that the neuron is triggered. An output of -1 specifies that the neuron did not get triggered.

“sgn” stands for sign function with output +1 or -1.

# **Question-3**

## **Perceptron Architecture**



* A Deep learning model learns to map the inputs to outputs from the given training set of examples. The training process is progressive or iterative to find a weight value which is optimum to solve the given problem. The weight values updated are small but in return which gives a change in performance.
* Expected output It is defined as the determined output for the given set of training examples.
* Actual output The output given by the model for the given set of training data.

The difference in Expected output and actual output gives us the error in the model.

* If we input an image which was not trained by the network in our case a bridge the following output will be : random output from the training sample. This problem is caused because our model didn't train in this praticular case bridge image will throw an error to the console which in turn will cause the deflection in the ouput.

To solve this problem we constantly check and update our weights after every epoch which in return reduces the error in the output.

Question-4

1. Binary - step function

Binary step functions depends on a threshold value which tells whether the neuron is activated or not. The input fed to the activation function is graded to this threshold value.If the input is greater than the threshold value then the neuron is activated else deactivated. Thus, output is not Transferred to the next layer.

Mathematical formula of binary-step function:

If ∑w.x > 0, output is +1, else -1. The neuron gets triggered only when weighted input reaches a certain threshold value.

f(x) = 1, x>=0  
 = 0, x<0

1. Sigmoid function

This function takes any real value as an input and gives binary classification output i.e., O or 1. It is a bounded function and differentiable.

S(x) = {1}/{1+e^{-x}}

1. Tanh function

If the input is a larger positive value , the closer the output will be to 1, and if the input is a small value close to negative then the output will be -1.

The Tanh activation function is calculated as follows: (e^x – e^-x) / (e^x + e^-x)

1. Rectified Linear Activation function

ReLU function is a piece wise linear function in which if the input is a positive number then it directly puts 1 to the output if not 0.

f(x) = max(0,x)

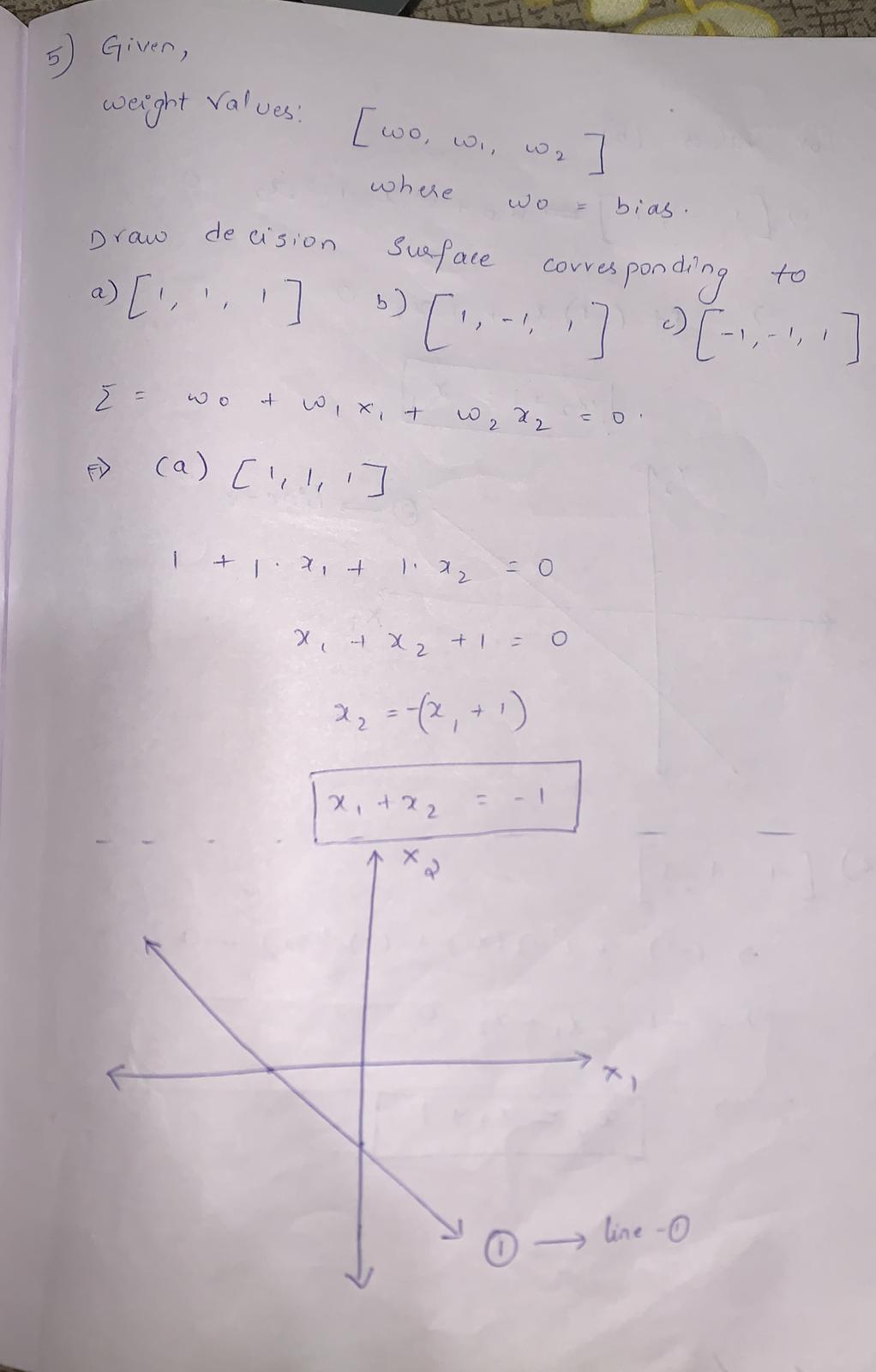
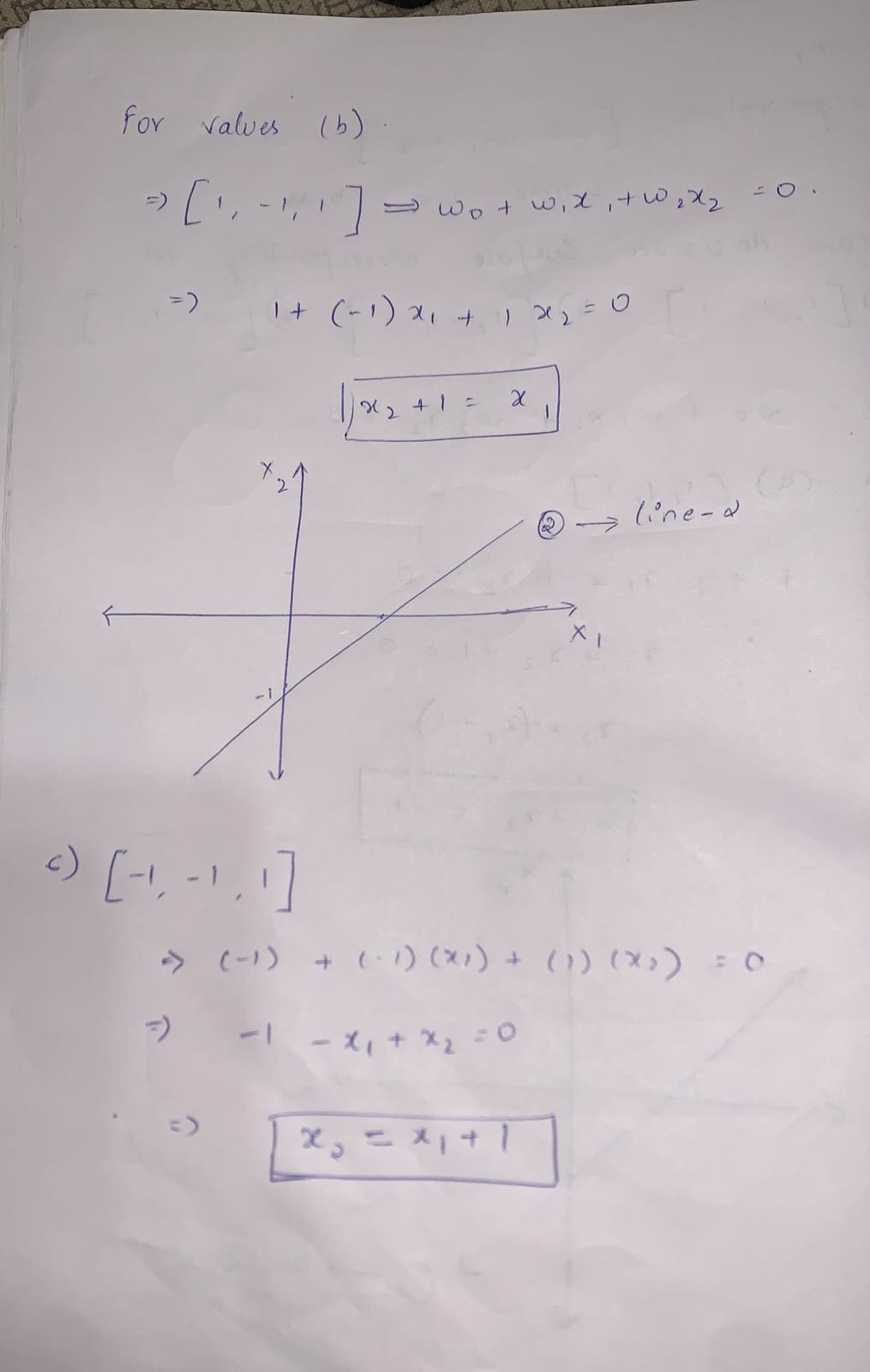
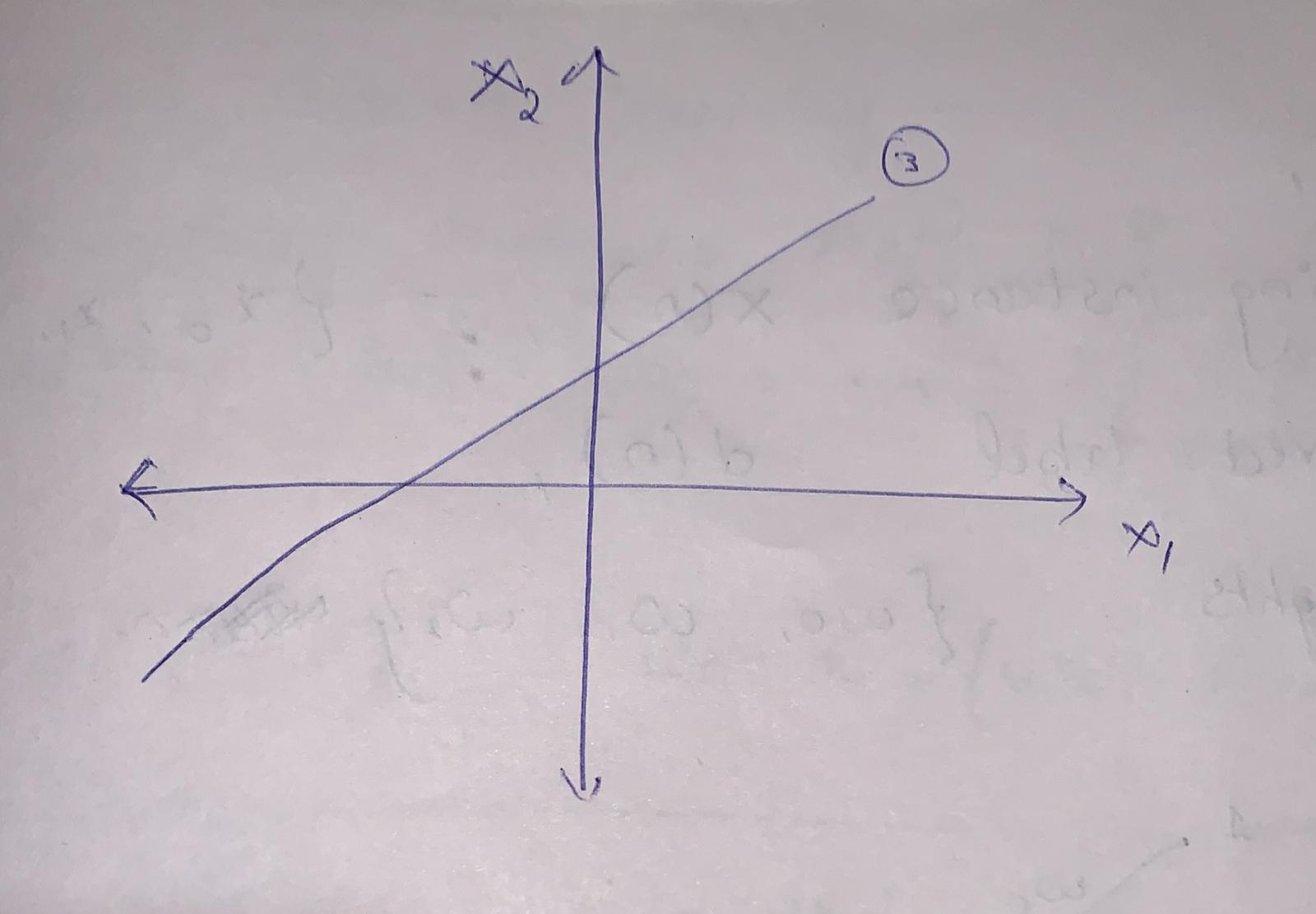
# **Question -5**[**¶**](http://localhost:8888/notebooks/Homework.ipynb#Question--5)

* Training the neuron We can train the neuron with the data by using the large margin classifier or support vector machine ( SVM )

It will always try to choose the decision boundary to be in between the separation with exactly same margin on both sides. The weight values of neuron are actually used to tell the neuron about how much a previous layer neuron weighs on the input for that neuron

* Roles of weight values

-In Deep learning, data is broken into small input pieces. These input pieces are mixed in “right proportion” to get desired output. It is similar to chef mixing the ingredient in “right proportion” to prepare a food. “Weight” captures the “proportion” in which input pieces are mixed together. -Weight is a parameter inside the neural network which converts the data to the hidden layers. -When an input enters the network is get multiplied by the weight value then it is either observed or passed onto the next layers.



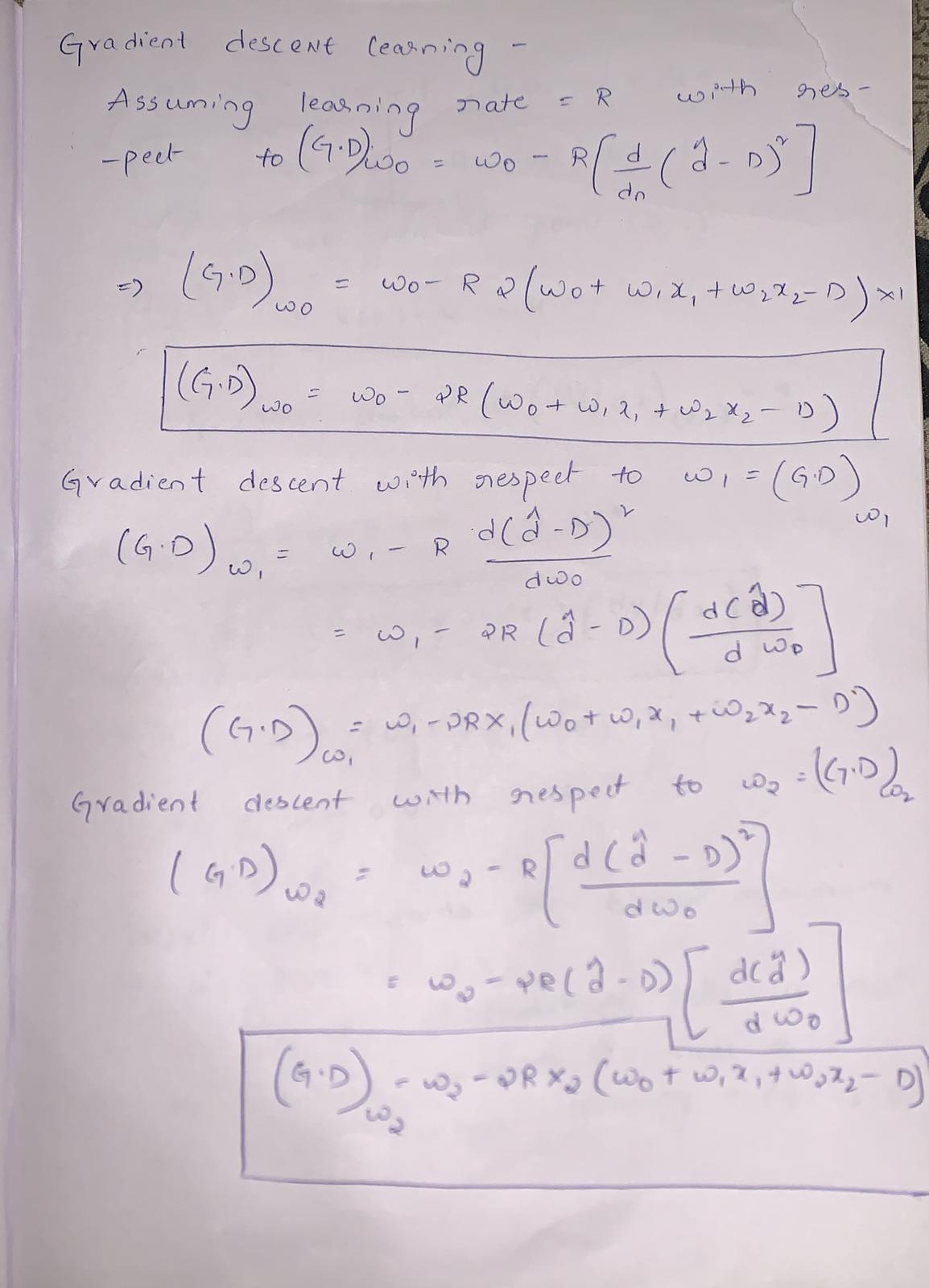
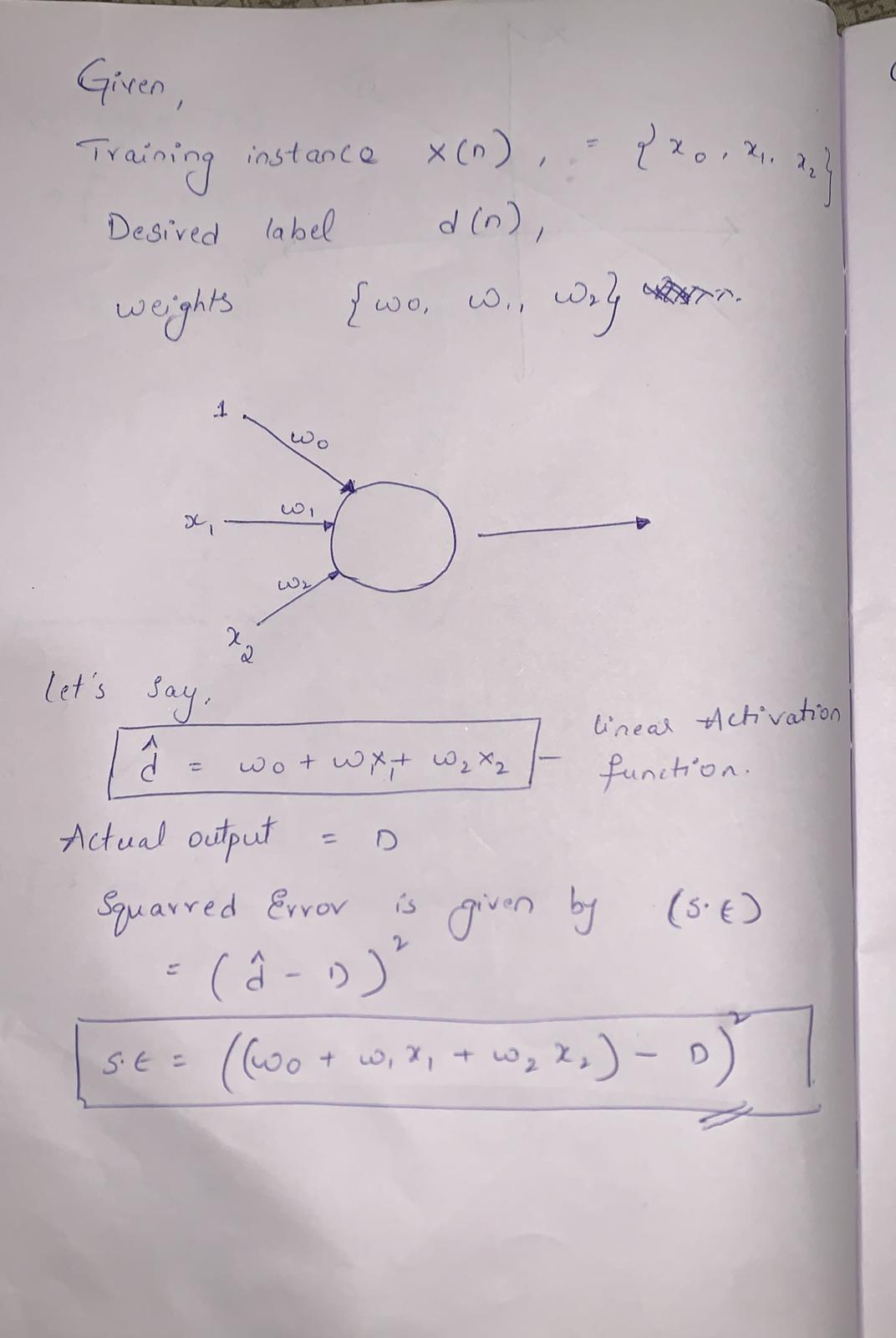
Weight-values

The weight values serves as support to balance the movement of the decision line. All this operation varies in three parameters; w1, w2 and w0. Depending on how you tune the decision boundary line, the angle theta may also increase or decrease.weights are used so that we can scale individual inputs.If input 𝑥3 for example isn't contributing enough to the right classification the perceptron will assign a small value to diminish it's output signal.

* Best decision

The line - (c) crossing at [-1, -1, 1] is the best line to seperate the instances as it perfectly classifies the given data set into two distintive sub patterns.

Question-6



Question-7

