Definition: A perceptron is a basic model that helps computers make simple decisions by learning from data

Example: Suppose we want to decide whether to go outside

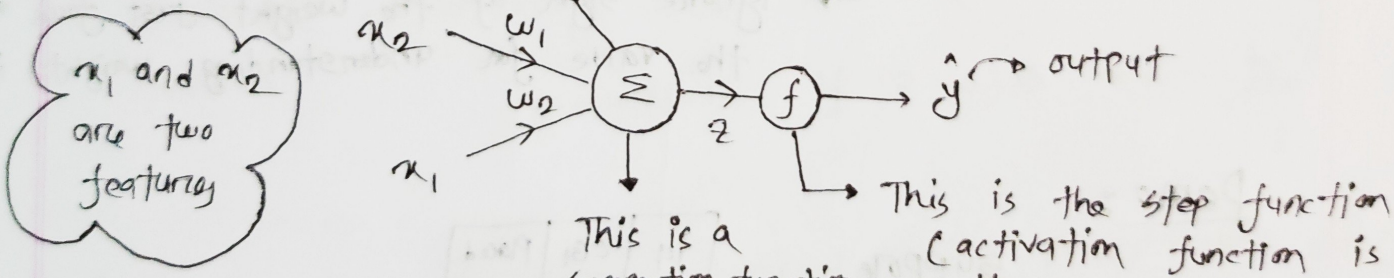
input (i): Is it sunny? (1 for yes, 0 for no)

input (ii): Do you have free time? (1 or 0)

The perceptron combines this info and decides "Yes, go out!" or "No, stay in."

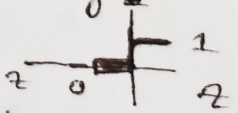
* It is a supervised ML algorithm and mathematical model / mathematical function.

Design:



step function

$$\text{output} = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{otherwise} \end{cases}$$



* It's imagined that perceptron works like Neuron (Human's nerve cell)

Perceptrons are inspired by neurons

* Activation function could be: relu, tanh, sigmoid, softmax, Leaky Relu

Perceptron has the two steps:

i) Data trainning for finding the value of weights and bias.

ii) Prediction (after getting the value of weights and bias we can make prediction)

* Keep in mind: weights are nothing but the importance of any particular feature. It tells us how important / impact of the feature to make the prediction.

w ↑ High impact
w ↓ Low impact
> on prediction

↳ ignore sign of the weight just focus on the value for understanding weight's impact.

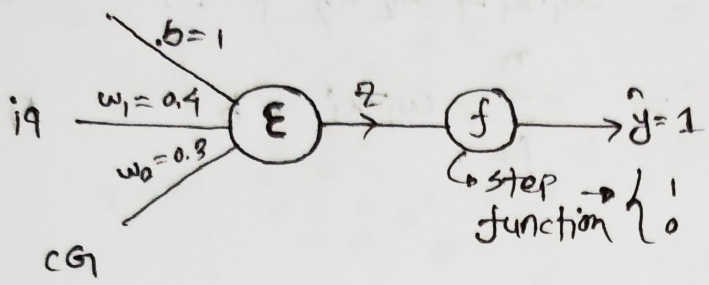
Demo:

Suppose

iq	CGI	Placed
78	3.5	1
60	3.6	0

A student has iq = 70 and CG = 3.40. Find the placement status.

Suppose we got weights $w_1 = 0.4$ and $w_2 = 0.3$ and bias $b = 1$ after trainning the model by given data then



Here,

$z = 7.0 \times 0.4 + 3.4 \times 0.3 + 1$

$= (+)$ ve

we get the student is placed for the seat.

Geometrical Intuition:

* For a single featured dataset the perceptron's decision boundary is just a point on a number

$z = wx + b$ if $z = 0 \Rightarrow x = -b/w$

* For two featured dataset the peceptron's decision boundary is 2D line.

$z = w_1 x_1 + w_2 x_2 + b$

* For three featured dataset is a the decision boundary is a plane

$z = w_1 x_1 + w_2 x_2 + w_3 x_3 + b$

* But for more than three features the decision boundary represents hyperplane.

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b$$

$$= \sum_{i=1}^{n} w_ix_i + b$$

$n \rightarrow$ no. of features / dimensions

1 D

$wx + b = z$

↳ $wx + b \geq 0$

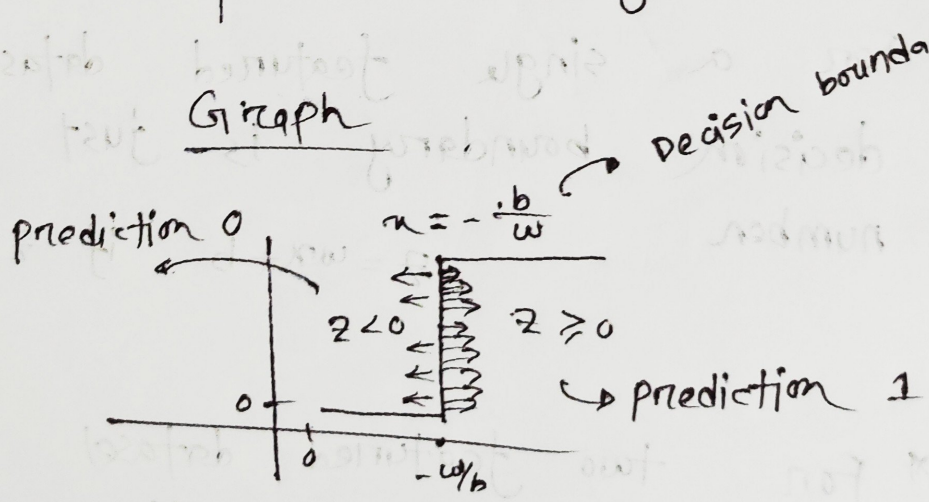
$wx + b < 0$

since z enters in activation function and we have for

$z \geq 0 \rightarrow \hat{y} = 1$

$z < 0 \rightarrow \hat{y} = 0$

Graph



At a glance:

perceptron create

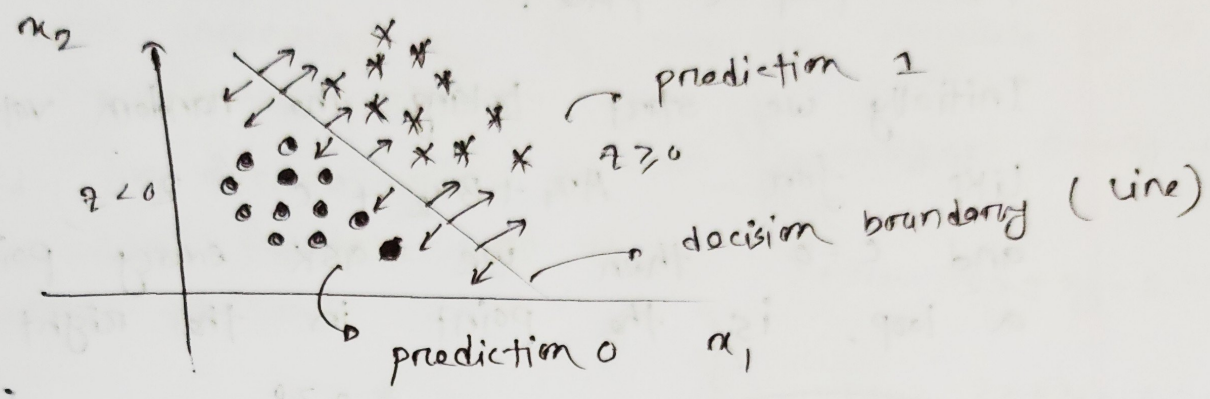
- point for 1D
- Line for 2D
- Plane for 3D
- Hyperplane for higher dimension (4D, 5D ....)

Then activation function provides a value

** perceptron classifies the data between two regions.

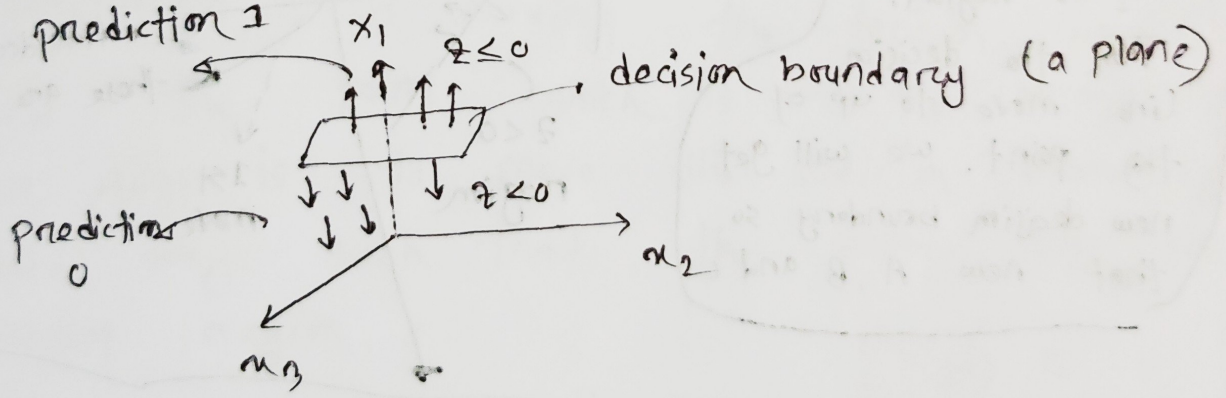
2D

$$w_1 x_1 + w_2 x_2 + b \geq 0$$
$$w_1 x_1 + w_2 x_2 + b < 0$$



3D

$$w_1 x_1 + w_2 x_2 + w_3 x_3 + b \geq 0$$
$$w_1 x_1 + w_2 x_2 + w_3 x_3 + b < 0$$



So perceptron is nothing but a binary classifier which classify the data between two classes

** Limitation: we show that perceptron classifies the data between two classes by a line / plane / hyperplane. So it can works only for linear / sort of linear data but failed in non-linear