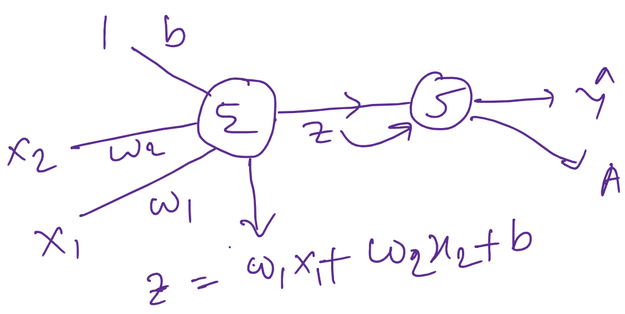
Perceptron

→ Algorithm → Supervised ML → Design → Deep learning build



Activation → Step Function

-1, 1
0, 1

Activation Function	Output Range	Common Use
ReLU	0 to ∞	Hidden layers
Sigmoid	0 to 1	Binary classification
Tanh	-1 to 1	Hidden layers
Softmax	0 to 1 (sum=1)	Multi-class output

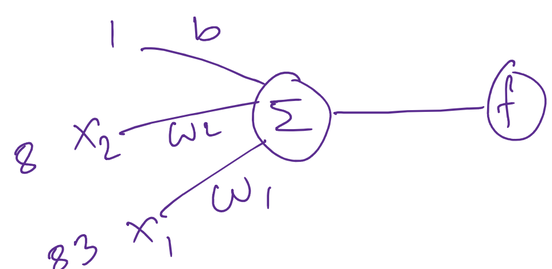
$z \geq 0 \rightarrow 1$

$z < 0 \rightarrow 0$

Student → models

IQ	Hours	Result
83	8	1
20	12	1
70	1	0
65	5	1

1) Training
2) Predict



Training Purpose → w_1, w_2, b

$w_1 = 1.5$

$w_2 = -2.6$

$w_2 = 2.6$

$b = 3$

New prediction

Eq Hour Result

60 3

$x_1 = 60$

$x_2 = 3$

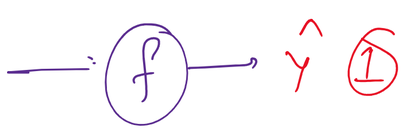
$z = w_1 x_1 + w_2 x_2 + b$

$= 1.5 \times 60 + 2.6 \times 3 + 3$

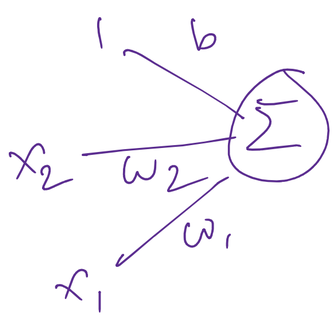
$= (+ve)$

$z \geq 0 \rightarrow 1$

$z < 0 \rightarrow 0$



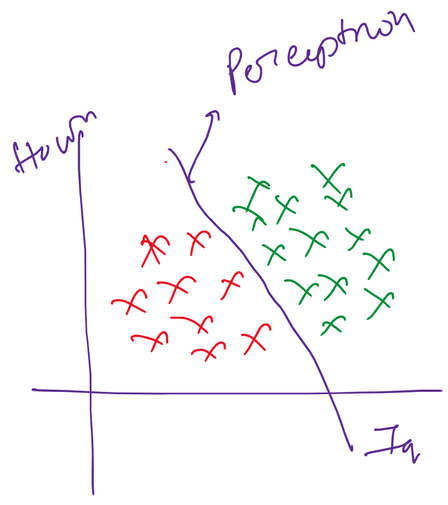
Prediction Perceptron → (w, b)

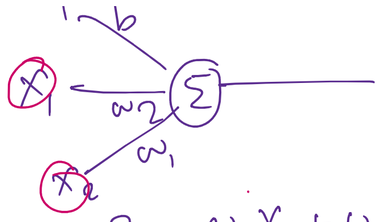


w = weights = Feature Importance

Students

Iq	Hours	Result
83	8	1
20	12	1
70	1	0
65	5	1





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

$$= \boxed{Ax + By + C}$$

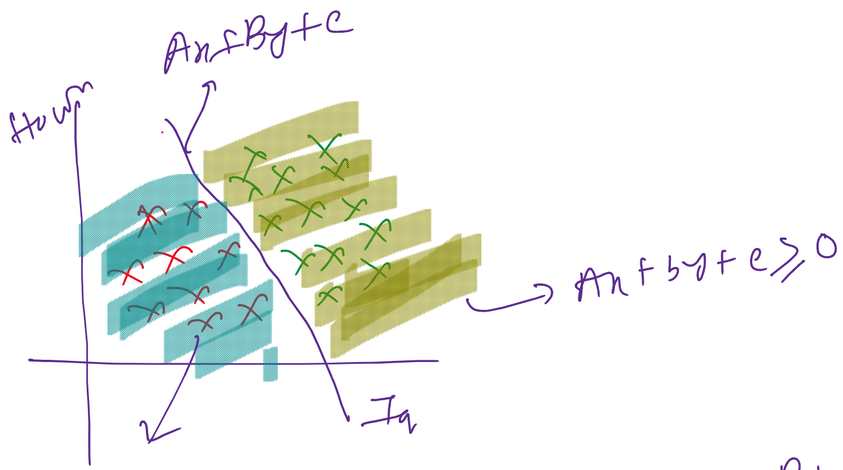
$w_1 = A$, $w_2 = B$, $b = C$

$x_1 = x$, $x_2 = y$

↓

Line

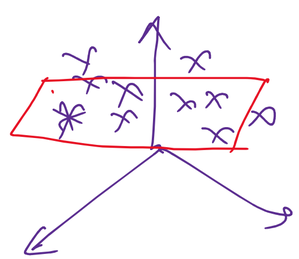
Perceptron → Line



Perceptron → Binary classification

Feature

Iq | Hours | Attendance



2D → Line

3D → plan

4D, 5D ... nD → Hyper plan

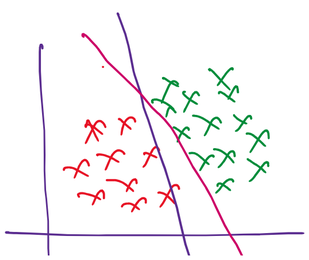
Limitation → Linear data ✓ ✓

non Linear data ✗ ○

Problem in Perceptron tricks

Loss Function

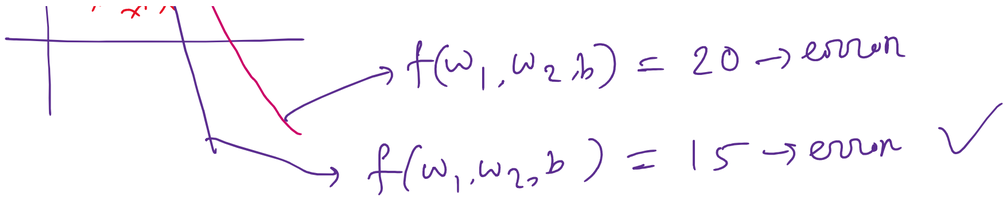
A way → ML Model Performance → How Good / How bad



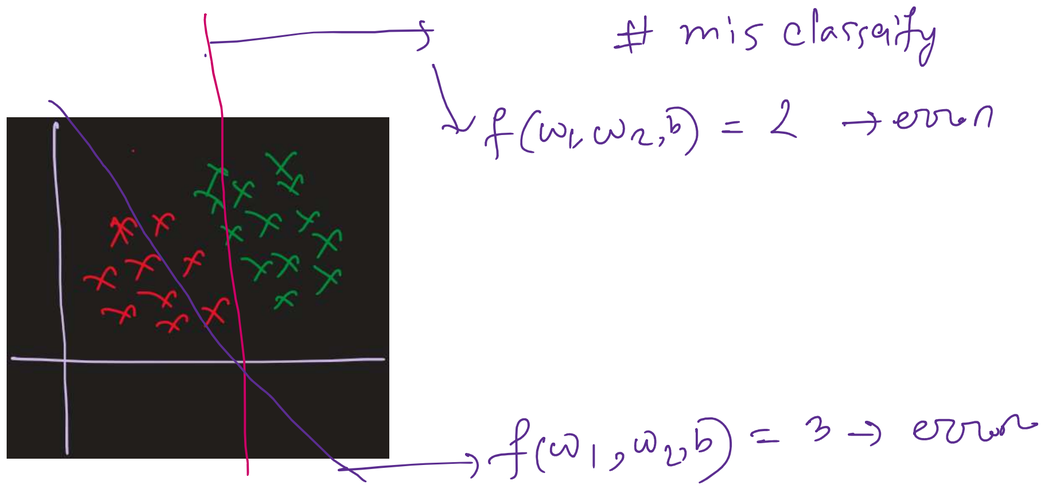
Loss → minimum

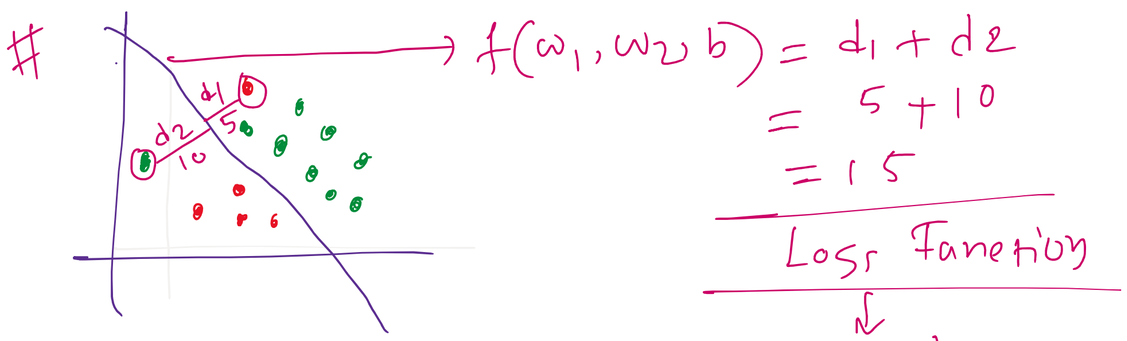
$f(w_1, w_2, b) = \text{number}$

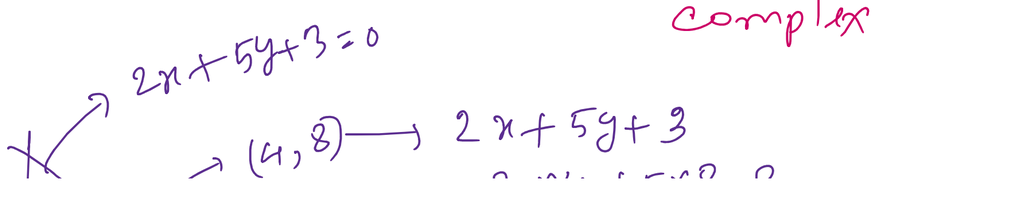
[illegible] → error

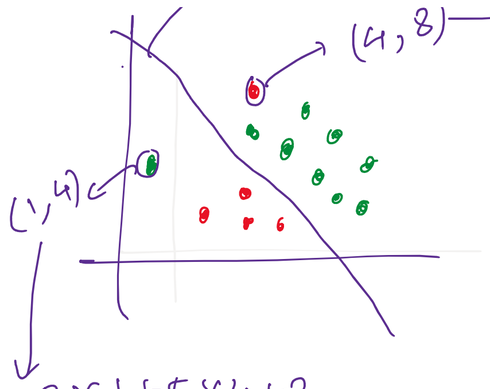


Algorithm Name	Loss Function Name	Loss Function Formula	Used For
Perceptron	Perceptron Loss	$\max(0, -y(wx + b))$	Binary Classification
Linear Regression	Mean Squared Error (MSE)	$(y - \hat{y})^2$	Regression
Linear Regression	Mean Absolute Error (MAE)	$\lvert y - \hat{y} \rvert$	Regression
Logistic Regression	Log Loss / Binary Cross Entropy	$-[y \log(\hat{y}) + (1-y) \log(1-\hat{y})]$	Binary Classification
Neural Network	Cross Entropy Loss	$-\Sigma\, y \log(\hat{y})$	Classification
Support Vector Machine (SVM)	Hinge Loss	$\max(0, 1 - y(wx + b))$	Classification
Neural Network	Categorical Cross Entropy	$-\Sigma\, y_i \log(\hat{y}_i)$	Multi-class Classification
Neural Network	Binary Cross Entropy	$-[y \log(\hat{y}) + (1-y) \log(1-\hat{y})]$	Binary Classification
Decision Tree	Gini Impurity	$1 - \Sigma\, p^2$	Classification
Decision Tree	Entropy Loss	$-\Sigma\, p \log(p)$	Classification









$2x+5y+3$

$= 2\times4 + 5\times8 + 3$

$= 8 + 40 + 3$

$= 51$

$2\times1 + 5\times4 + 3$

$= 2 + 20 + 3$

$= 25$

Total error $= 51 + 25$

$= 76$

↑

Perceptron Loss function

Loss Function

$$L(w_1, w_2, b) = \frac{1}{n}\sum_{i=1}^{n} L(y_i f(x_i)) + \underbrace{\alpha R(w_1, w_2)}_{\text{Ignore}} \quad \leftarrow \text{Regularization}$$

$$L(y_i f(x_i)) = \max(0, -y_i f(x_i))$$

$\hookrightarrow f(x_i) = w_1 x_1 + w_2 x_2 + b$

$$L = \frac{1}{n}\sum_{i=1}^{n} \max(0, -y_i f(x_i)) \rightarrow L(w, b)$$

n = number of rows

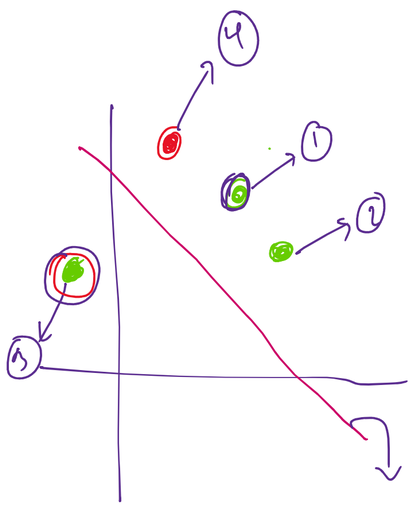
y_i = Actual output

$$L = \underset{w_1\, w_2\, b}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n} \max(0, -y_i f(x_i)) \rightarrow \text{minimum}$$

Gradient descent

$z \geq 0 \rightarrow 1$

$z < 0 \rightarrow 0$



	x_1 IQ	x_2 Hours	Y Result	Predict	
1	83	8	1	1	✓
2	20	12	1	1	~
3	70	1	-1	1	x
4	65	5	1	-1	x

Student

$z \geq 0$ ↓ $w_1x_1 + w_2x_2$

✓ 0 ✓ 0

$$L = \frac{1}{4}\left[\max(0, -y_1 f(x_1)) + \max(0, -y_2 f(x_2)) + \max(0, -y_3 f(x_3)) + \max(0, -y_4 f(x_4))\right] = \frac{0}{4} = 0$$

(+ve) (+ve)

Loss function

now = 1

$\max(0, -y_1 f(x_1)) \rightarrow y_1$ (+ve)

$\max(0, -(+ve)(+ve))$

$\max(0, -ve)$ → 0

$f(x_1)$

$w_1x_{11} + w_2x_{12} + b \geq 0$ (+ve)

now = 3

$\max(0, -y_3 f(x_3)) \rightarrow y_3$ (-ve)

$\max(0, -(-ve)(+ve))$

$f(x_3)$

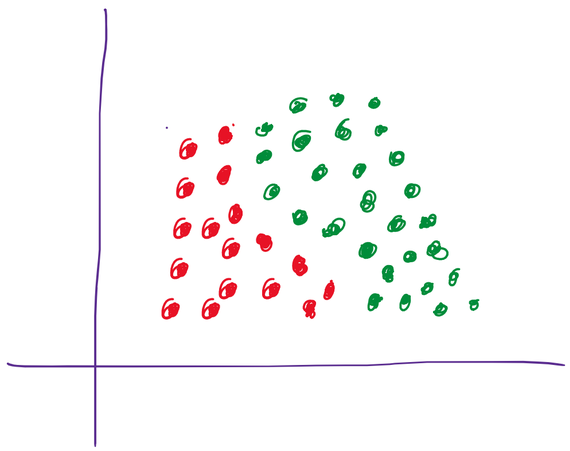
$w_1x_{31} + w_2x_{32} + b \geq 0$ (+ve)

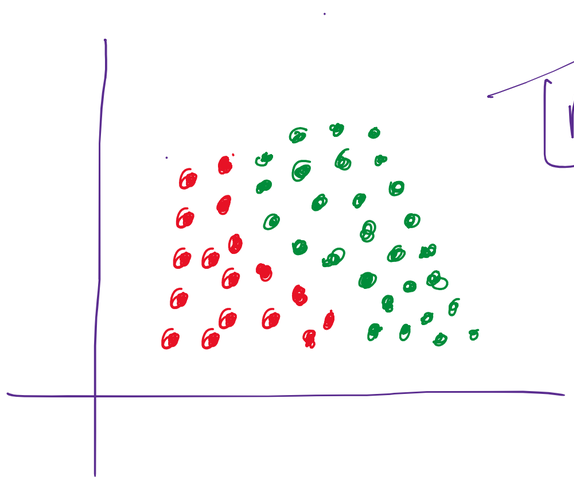
max(0, -ve) (+ve) (-ve) (+ve)

max(0, +ve)

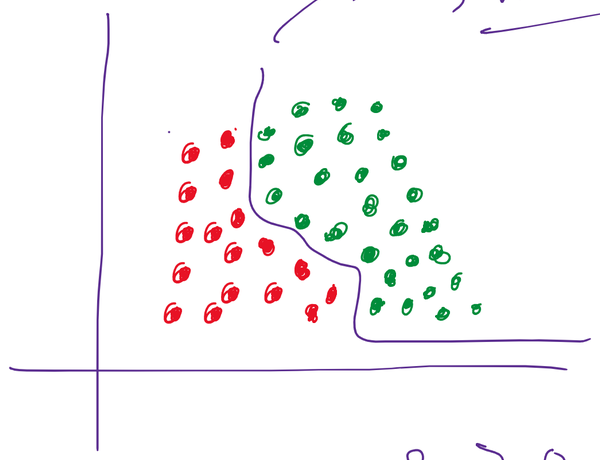
+ve →

Multi Layer Perceptron





MLP

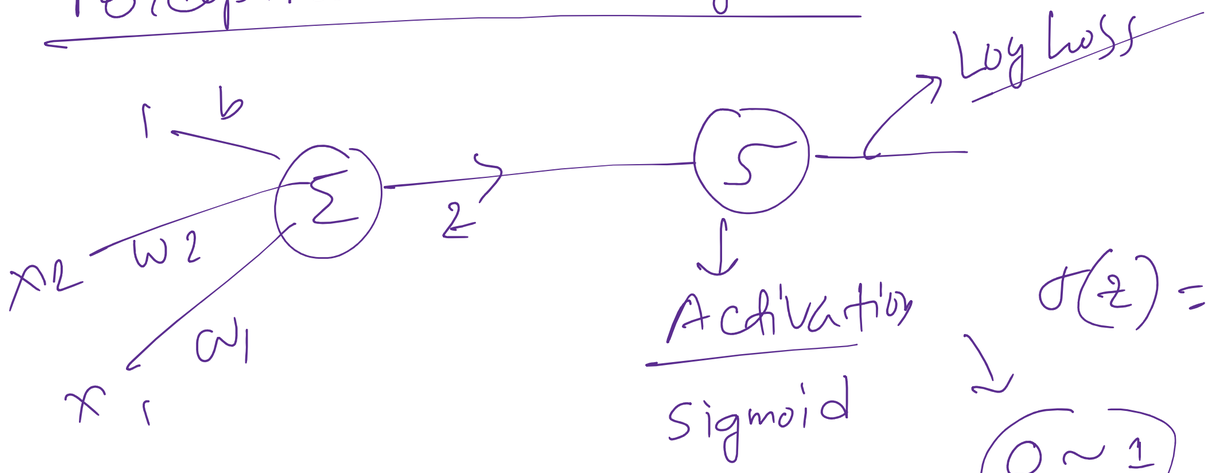


$z \geq 0 \rightarrow 1$

$z < 0 \rightarrow 0$

Perceptron with sigmoid

Perceptron with sigmoid

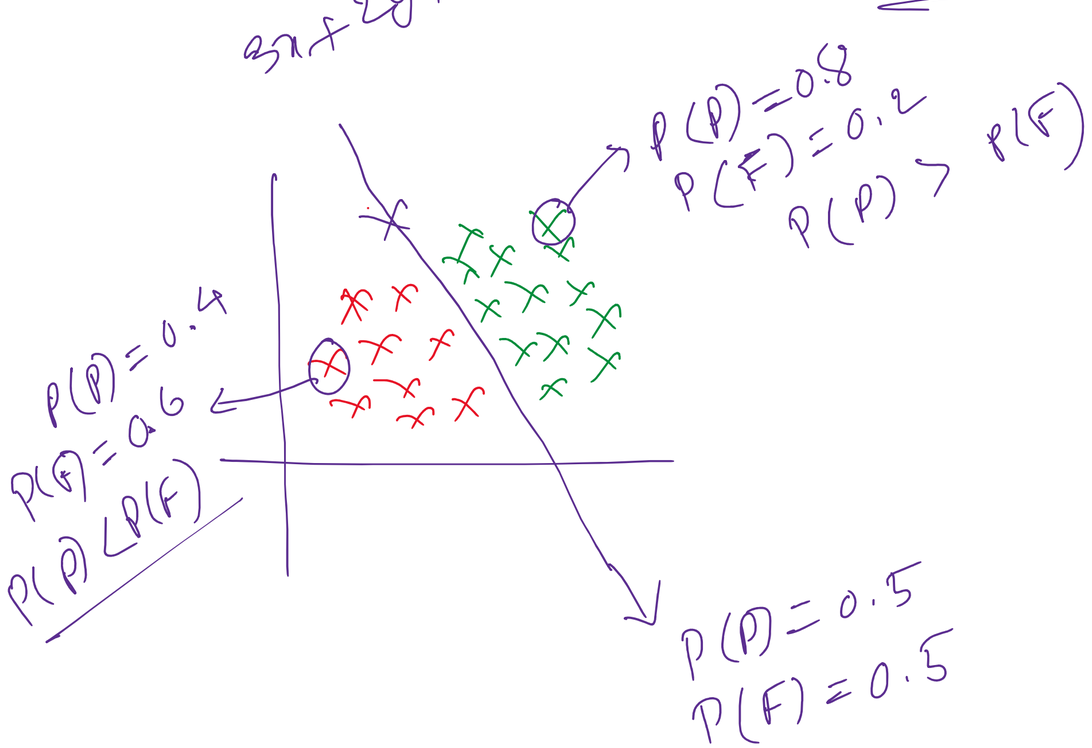


$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$0 \sim 1$

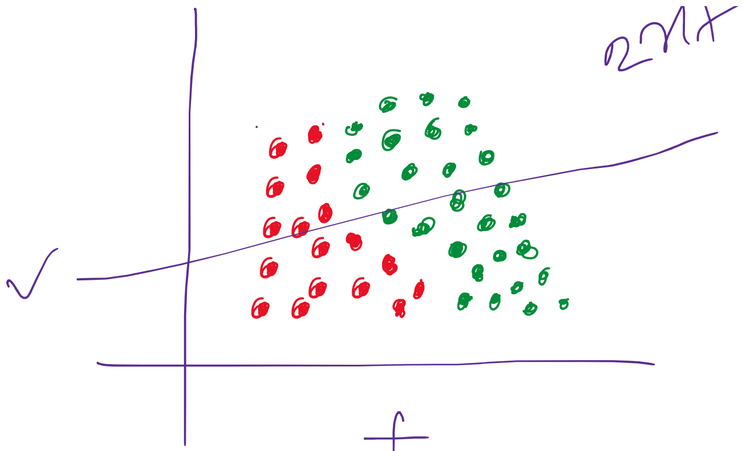
Probability

$ax+2y+5=0$

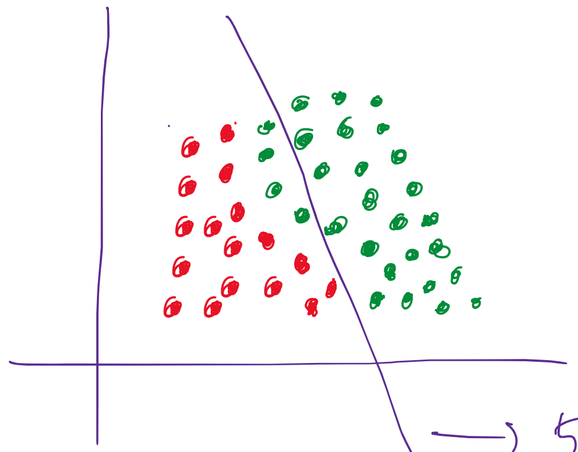


1

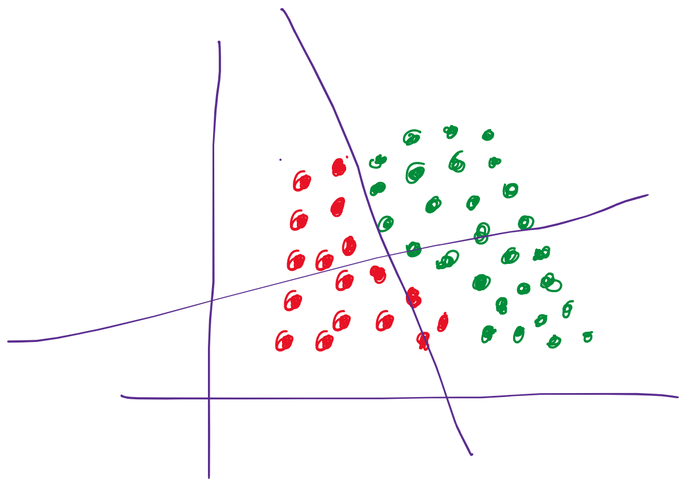
$2x+4y+2=0$



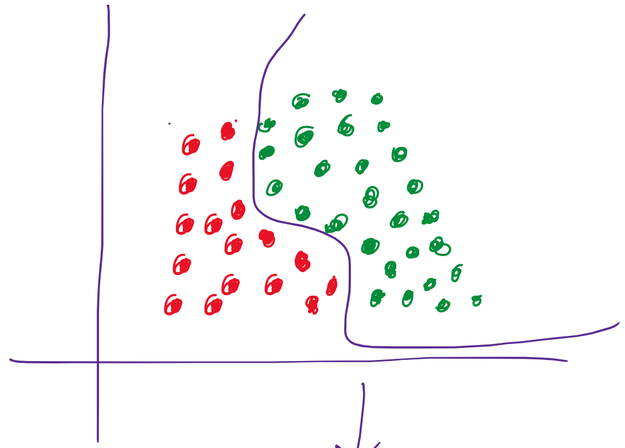
+



→ $5x+3y+2=0$

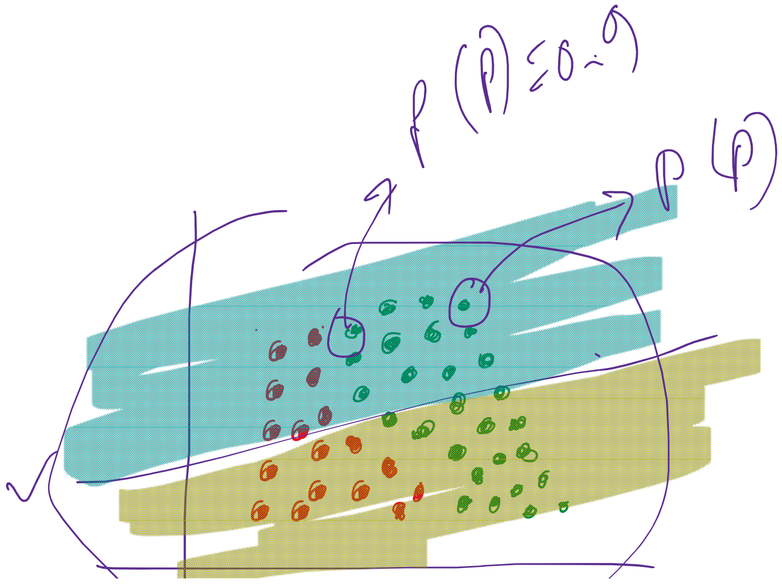


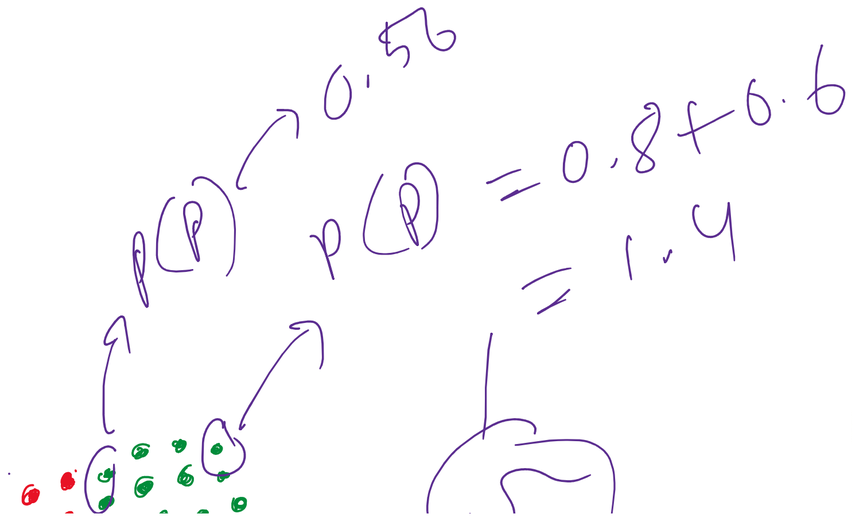
=

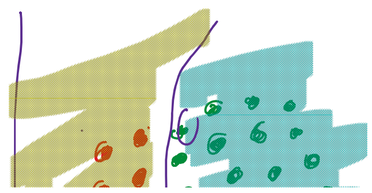


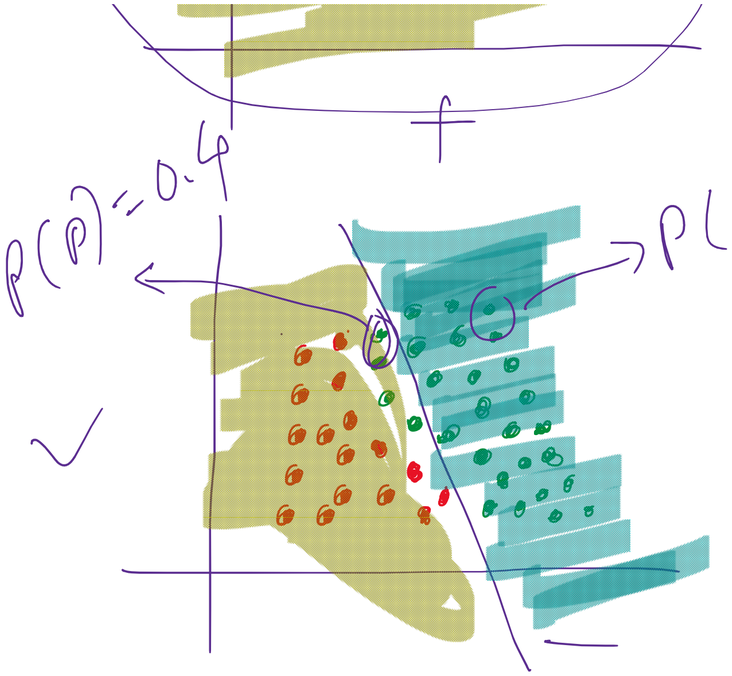
Smoothing

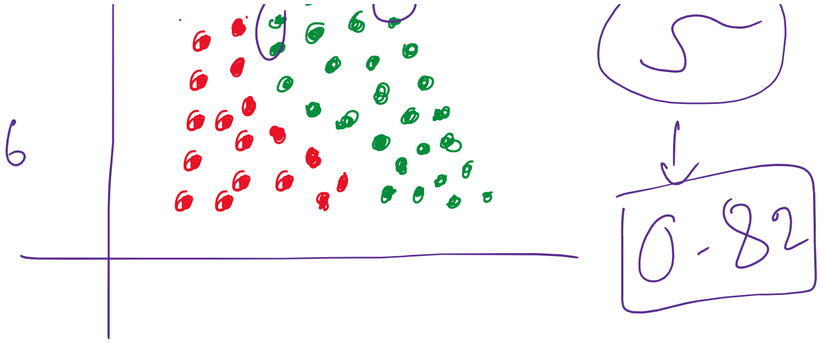
MLP



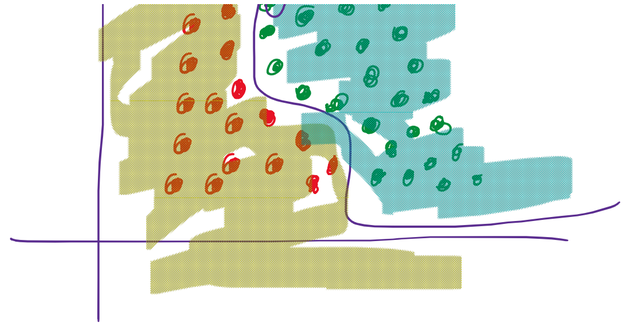




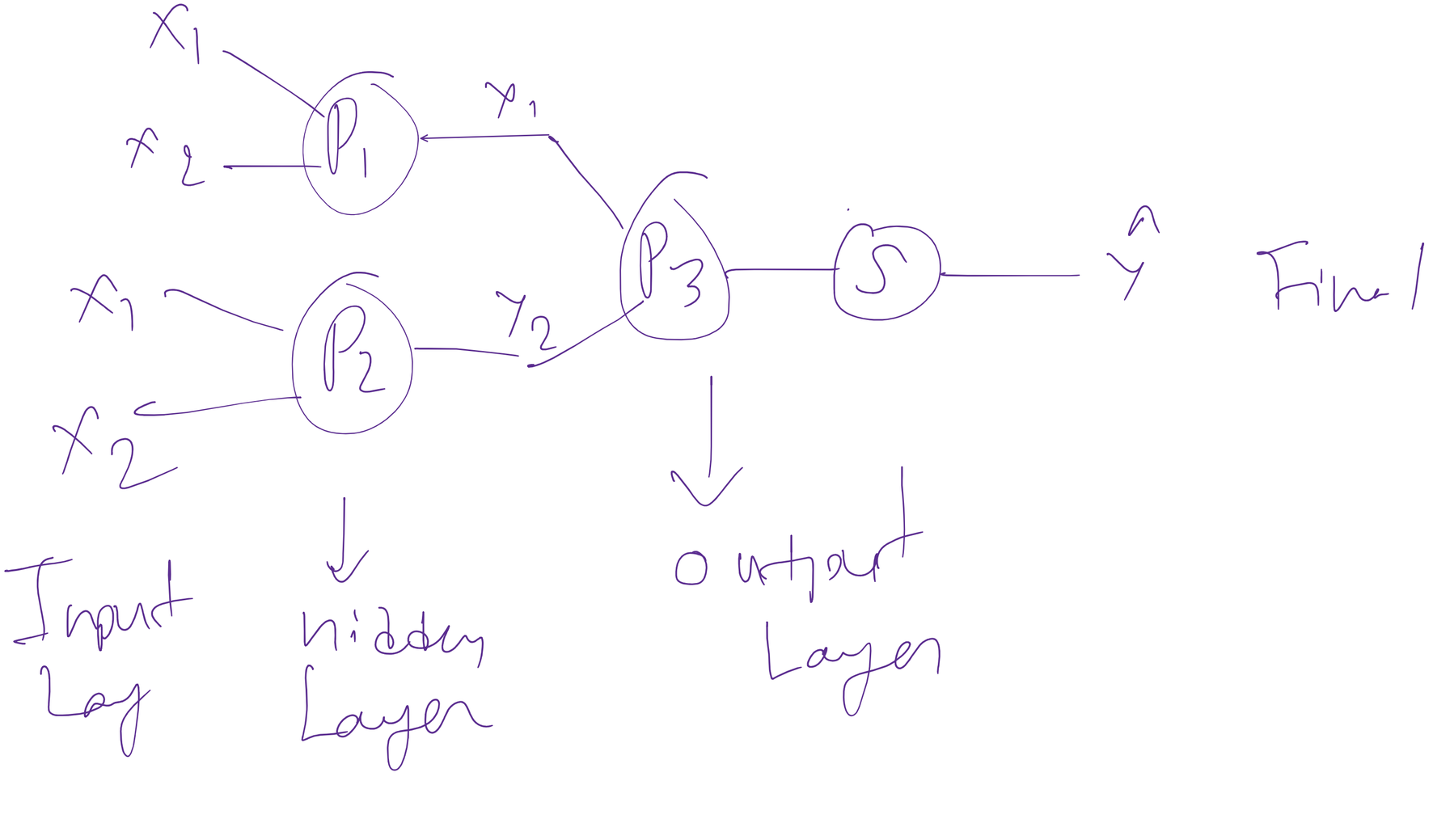

P(P)=0.4
P(P)=0.6



S
0.82



MLP


x_1
x_2
P_1
y_1
x_1
x_2
P_2
y_2
P_3
S
$\hat{y}$
Final
Input Lay
hidden Layer
output Layer

1)