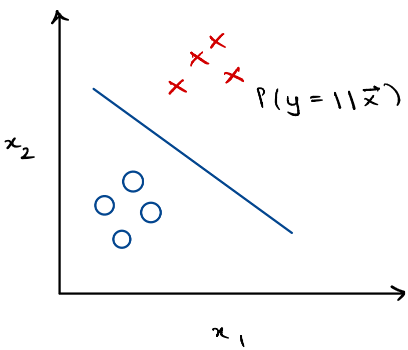
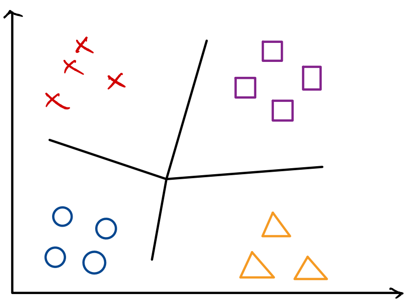
A type of classification algorithm in which y can have more than two possible categories is called a multiclass categories.

eg. recognizing parts defects in factories, finding whether a patient has multiple diseases or not, etc.



binary classification



multiclass classification

softmax algorithm is used for this!

Softmax classification

- Basically used for probability distribution of all possible outputs that the algorithm can classify into.

- General formula :- $a_j = \dfrac{e^{z_j}}{\sum_{k=1}^{N} e^{z_k}}$ $[P(y = j \mid x)]$

↳ In logistics regression :- only 2 outputs

eg. $a_1 = g(z) = \dfrac{1}{1+e^{-z}} = P(y=1 \mid \vec{x}) = 0.71$

Since total probability is always 1 . $P(y=0 \mid \vec{x})$

$= 1 - P(y=1 \mid \vec{x}) = 1 - 0.71 = 0.29$

⇒ Chance of probability of y being 1 = 71 %

and " " " y being 0 = 29 %

Softmax has more than 1 output

For eg. (for 4 outputs 1–4)

$z_1 = \vec{w}_1 \vec{x} + b_1$

$z_2 = \vec{w}_2 \vec{x} + b_2$

$z_3 = \vec{w}_3 \vec{x} + b_3$

$z_4 = \vec{w}_4 \vec{x} + b_4$

$a_1 = \dfrac{e^{z_1}}{e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}}$

$= P(y = 1 \mid \vec{x})$

0.30

0.20

$a_2 = \dfrac{e^{z_2}}{e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}} = P(y=2 \mid \vec{x})$

0.15

$a_3 = \dfrac{e^{z_3}}{e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}} = P(y=3 \mid \vec{x})$

$1-(0.30+0.20+0.15) = 1-0.65 = 0.35$

$a_4 = 1 - (a_1 + a_2 + a_3) = P(y=4 \mid \vec{x})$

↳ Last one can also be calculated by subtracting from the total.

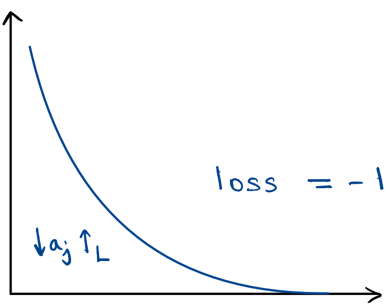
Cost Function for Softmax

$$a_1 = \frac{e^{z_1}}{e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}} = P(y=1 \mid \vec{x})$$

$$\vdots$$

$$a_N = \frac{e^{z_N}}{e^{z_1} + e^{z_2} + e^{z_3} + \dots + e^{z_N}} = P(y=N \mid \vec{x})$$

$$\text{loss}(a_1, \dots, a_n, y) \begin{cases} -\log a_1 & \text{if } y = 1 \\ -\log a_2 & \text{if } y = 2 \\ \vdots & \\ -\log a_N & \text{if } y = N \end{cases}$$



$$\text{loss} = -\log a_j \quad \text{if } y = j$$