

# A Brief Introduction to ML

## 2

credit:[https://speech.ee.ntu.edu.tw/~hylee/ml/ml2021-course-data/classification\\_v2.pdf](https://speech.ee.ntu.edu.tw/~hylee/ml/ml2021-course-data/classification_v2.pdf)

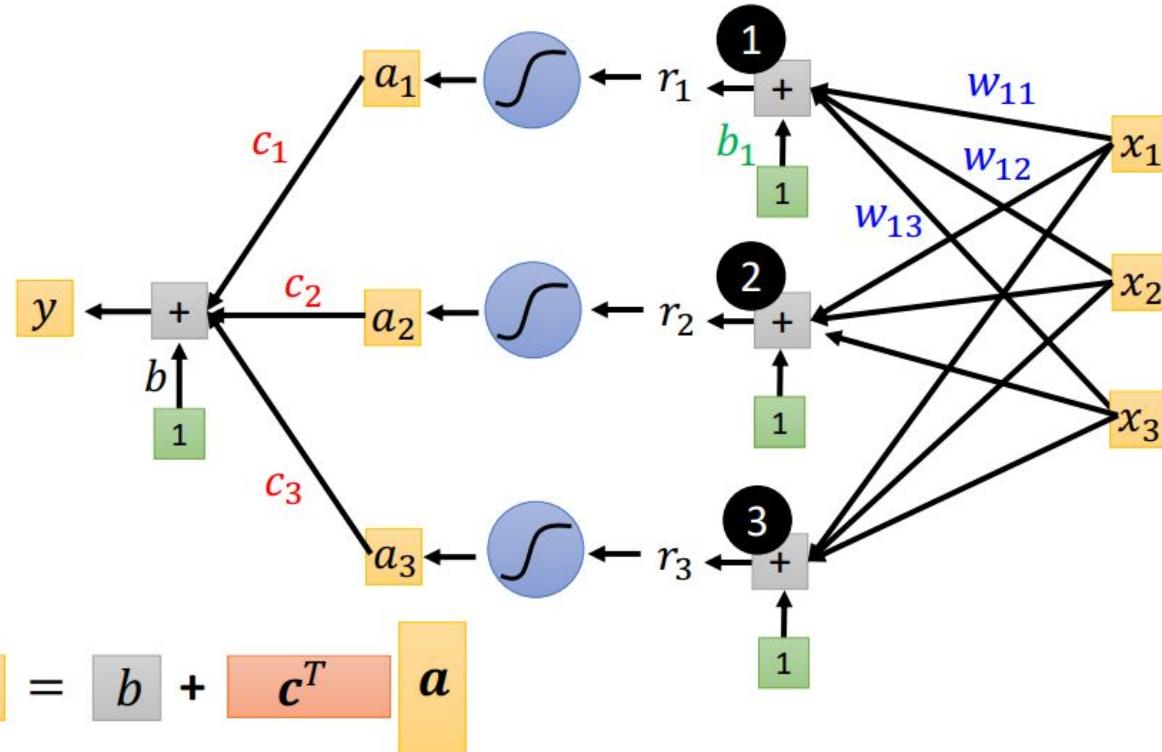
# Regression

How to do classification?

How to indicate the output class?

$$y = b + \sum_i c_i \text{sigmoid} \left( b_i + \sum_j w_{ij} x_j \right)$$

$i: 1, 2, 3$   
 $j: 1, 2, 3$

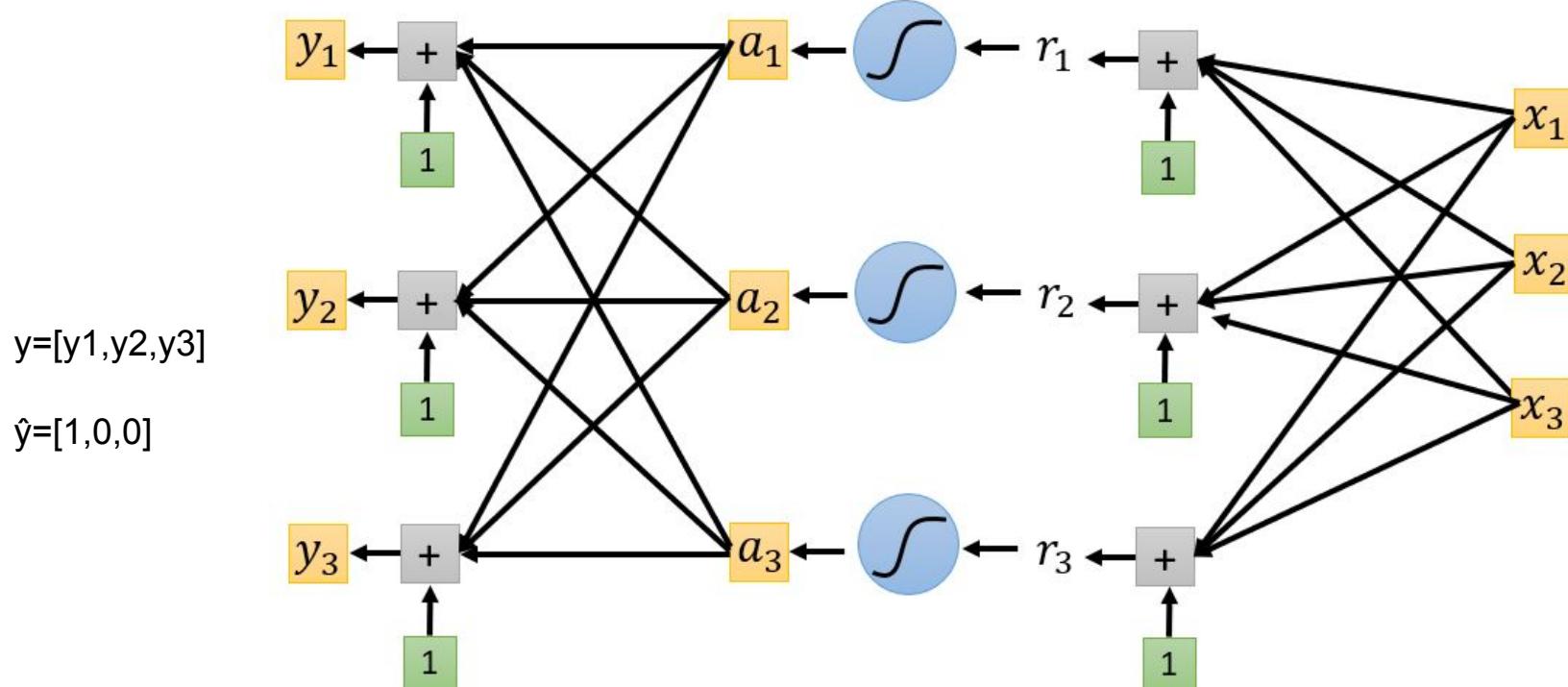


# One-hot Vector

$$\hat{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Class 1      Class 2      Class 3

# Classification



## Regression

label

$$\hat{y} \longleftrightarrow y = b + c^T \sigma(b + Wx)$$

feature

## Classification

$$y = b' + W' \sigma(b + Wx)$$

feature

label  $\hat{y} \longleftrightarrow y' = \text{softmax}(y)$

0 or 1

Make all values  
between 0 and 1

Can have  
any value

# softmax

$$y = [0.88, 0.12, 0]$$
$$\hat{y} = [1, 0, 0]$$

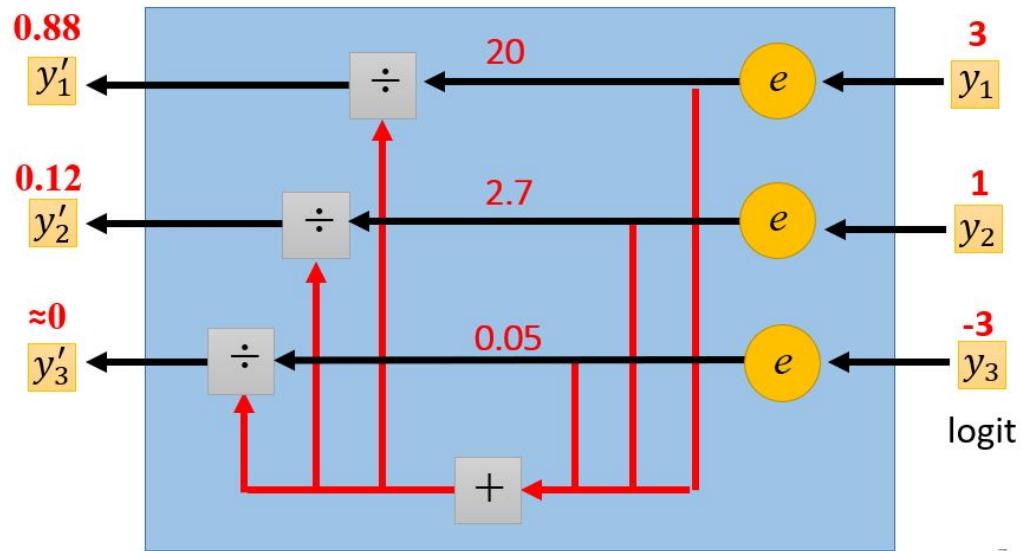
Soft-max

$$y'_i = \frac{\exp(y_i)}{\sum_j \exp(y_i)}$$

- $1 > y'_i > 0$
- $\sum_i y'_i = 1$

Softmax

How about **binary classification?** ☺



# Loss Functions

Mean Square Error (MSE)

$$e = \sum_i (\hat{y}_i - y'_i)^2$$

Cross-entropy

$$e = - \sum_i \hat{y}_i \ln y'_i$$

$\ln(1)=0$ ,  $\ln(0.1)=-2.30258\dots$ ,  $\ln(0.5)=-0.69314\dots$

