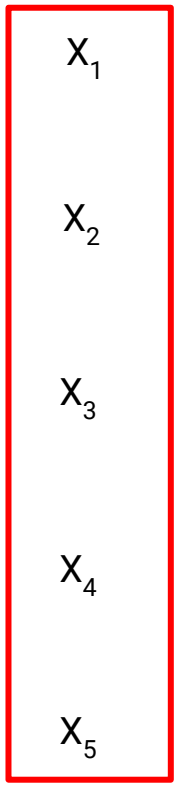
Softmax Activation Function

Multiclass Classification Problem

Multiclass Classification Problem

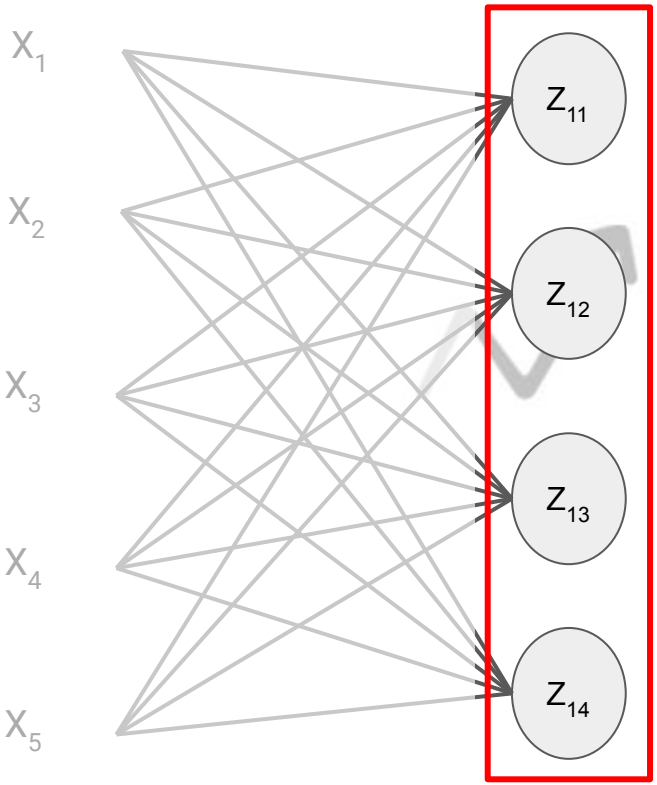
Feature X1	Feature X2	Feature X3	Feature X4	Feature X5	Target
1	22	569	35	0	Class 1
1	7	351	75	1	Class 2
1	45	451	542	1	Class 2
1	5	572	8	0	Class 1
0	22	565	44	1	Class 3
0	24	243	546	1	Class 3
1	78	953	42	0	Class 2

Multiclass Classification Problem



Input Layer

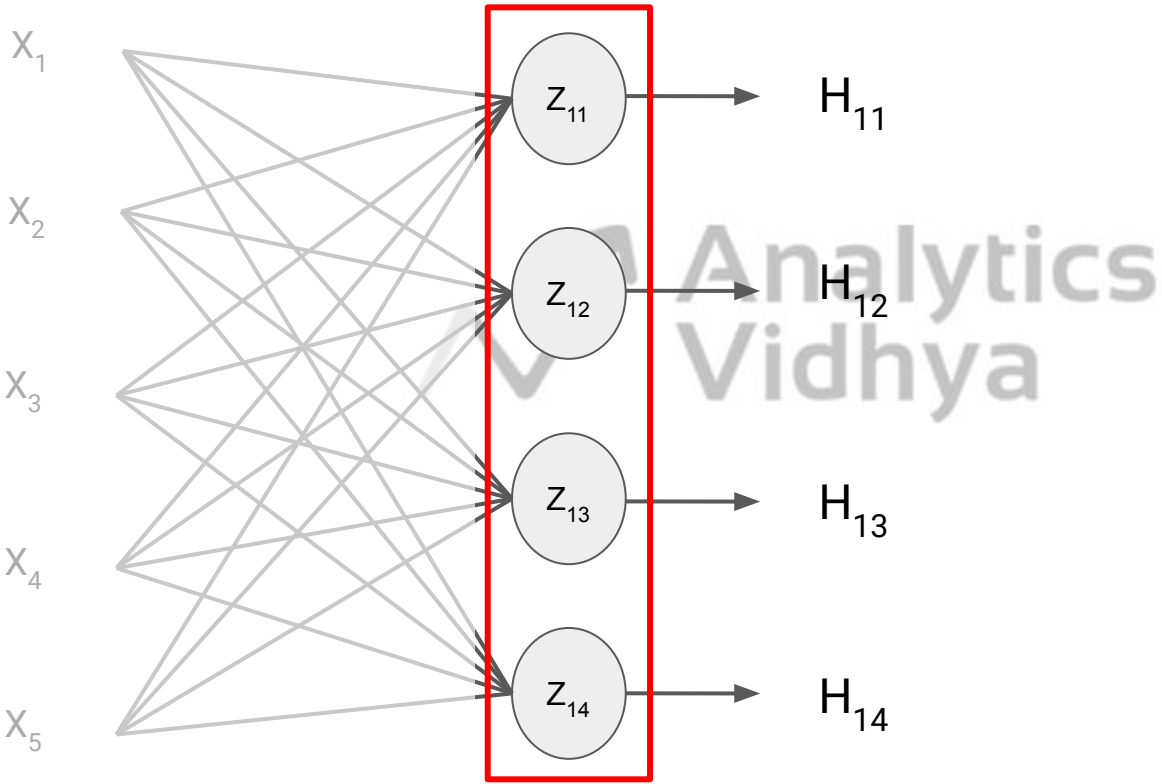
Multiclass Classification Problem



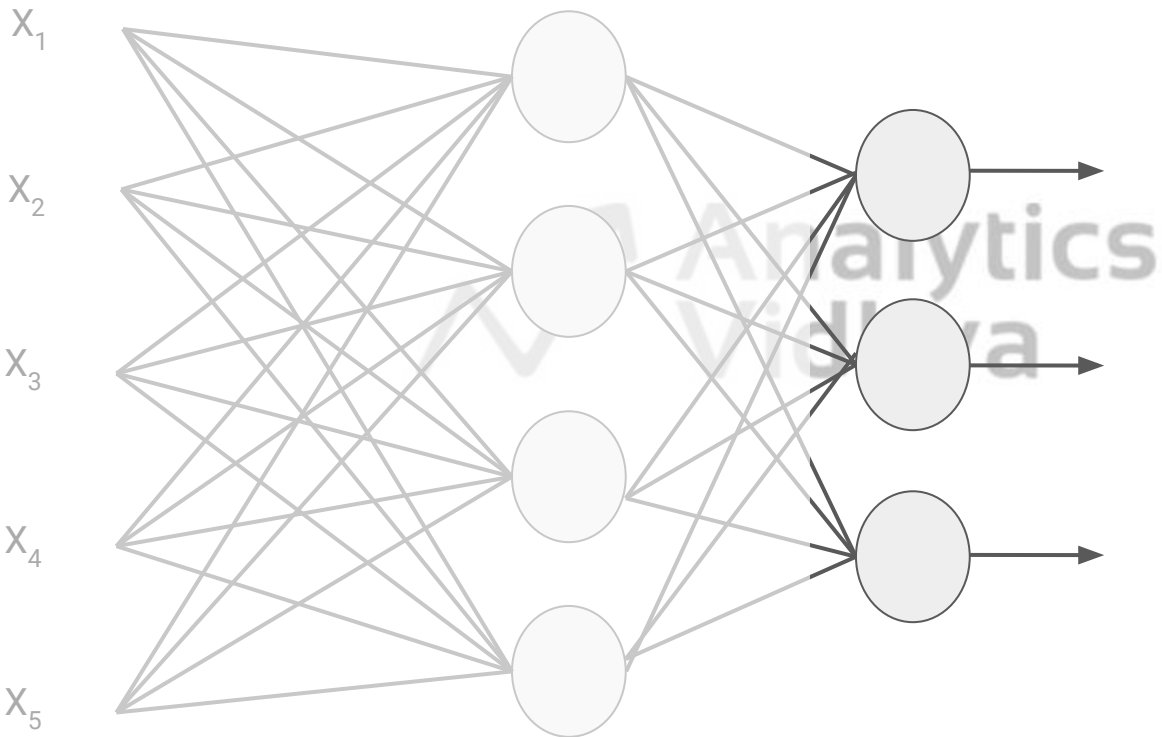
Input Layer

Hidden Layer

Multiclass Classification Problem

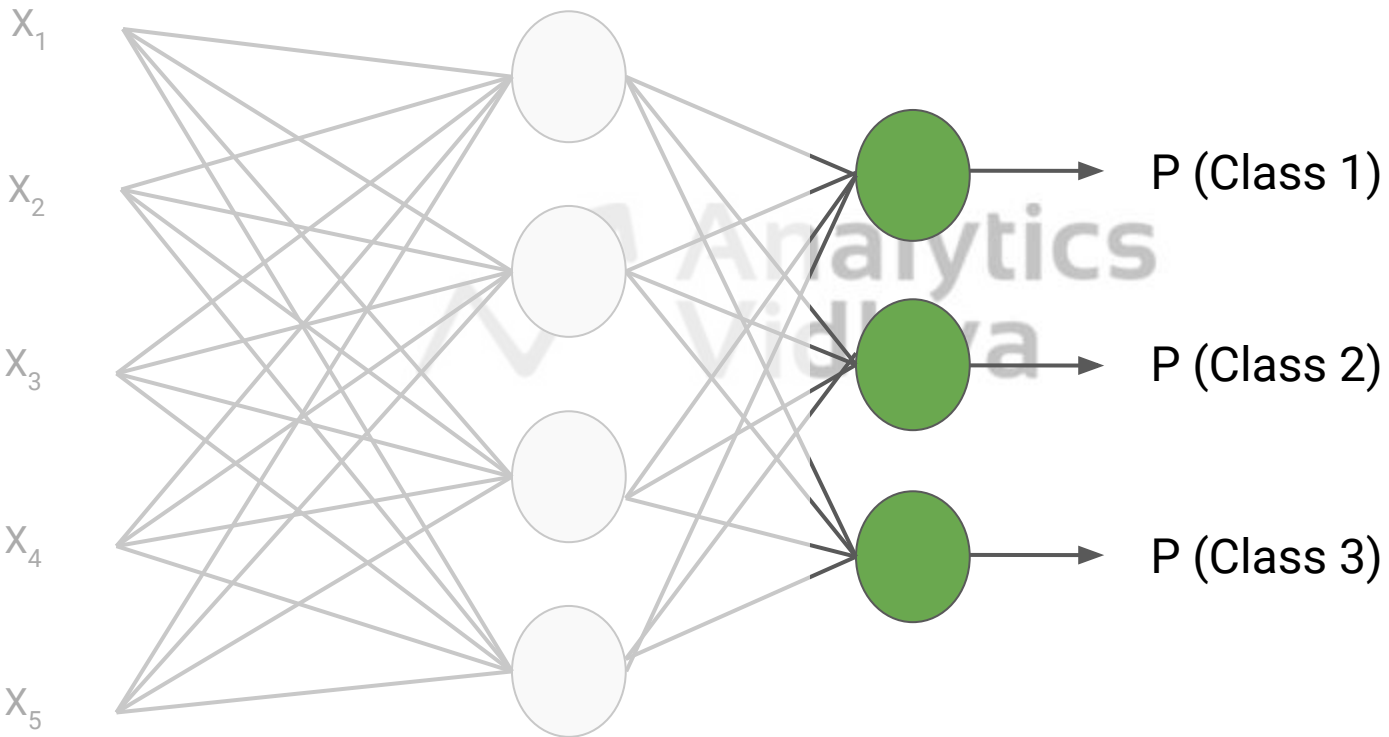


Multiclass Classification Problem

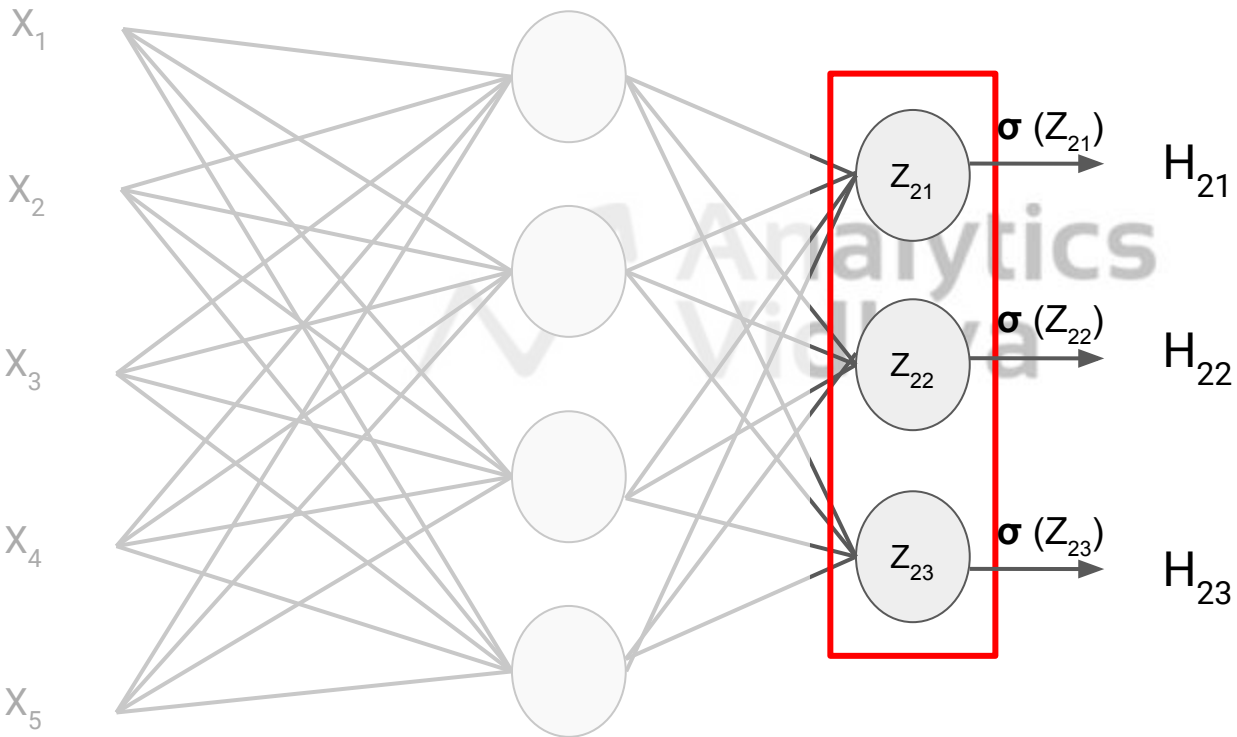


Input Layer Hidden Layer Output Layer

Multiclass Classification Problem

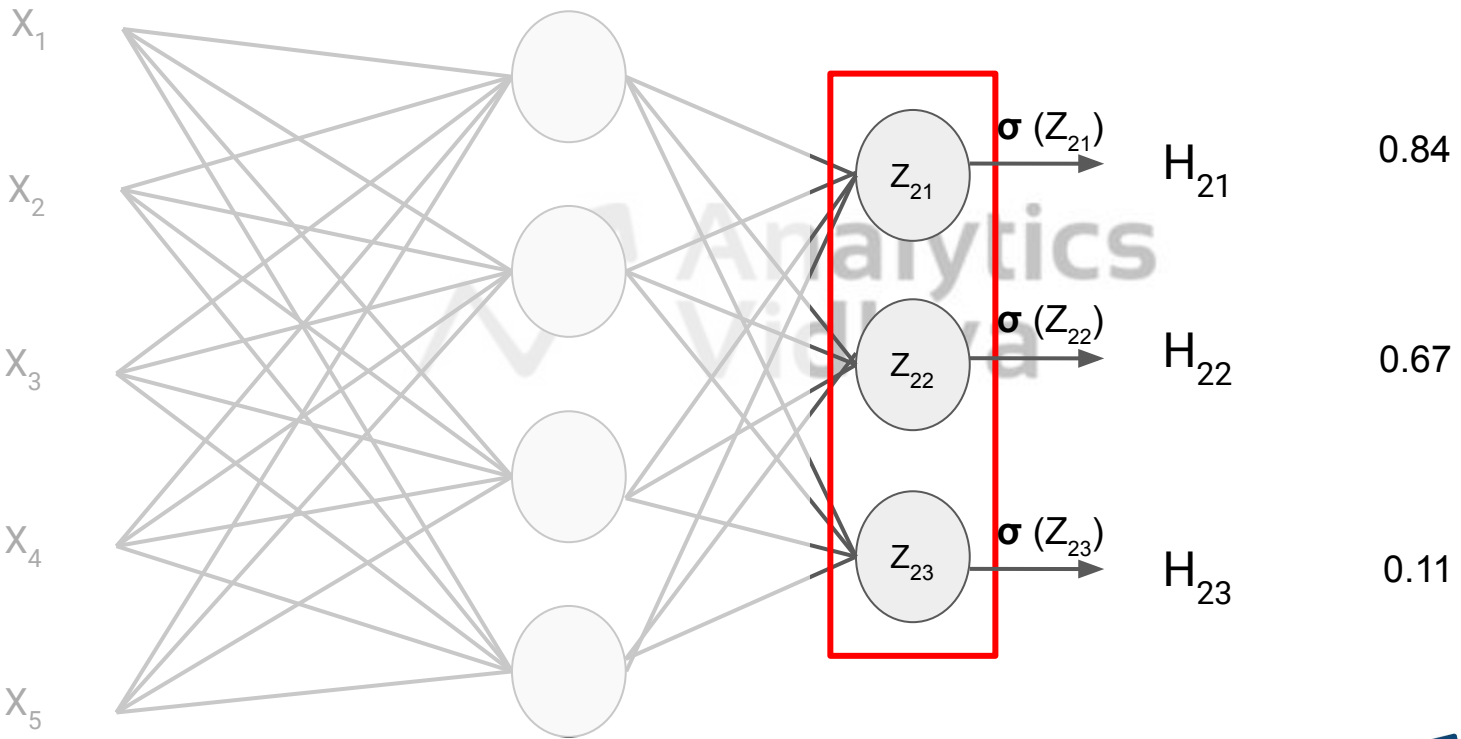


Multiclass Classification Problem: Sigmoid

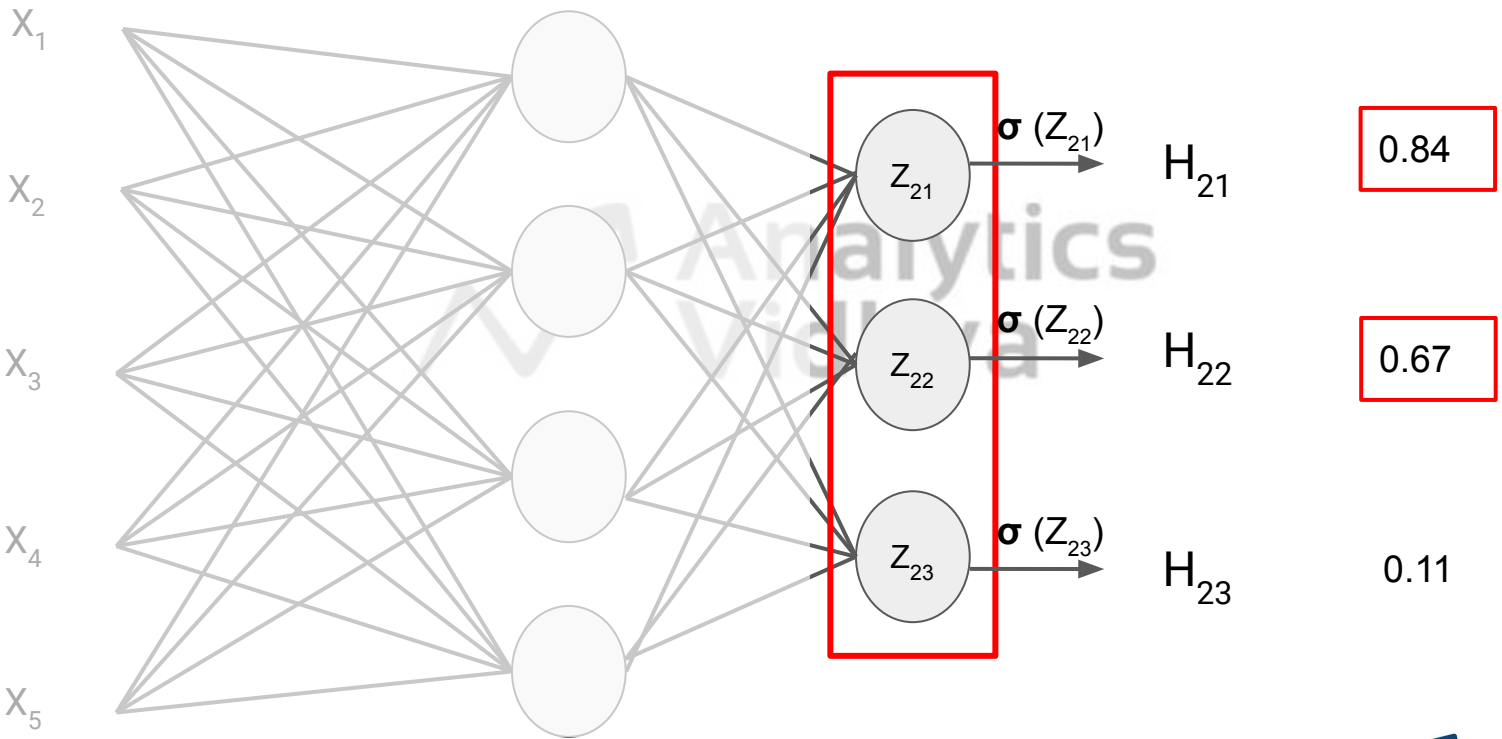


Input Layer Hidden Layer Output Layer

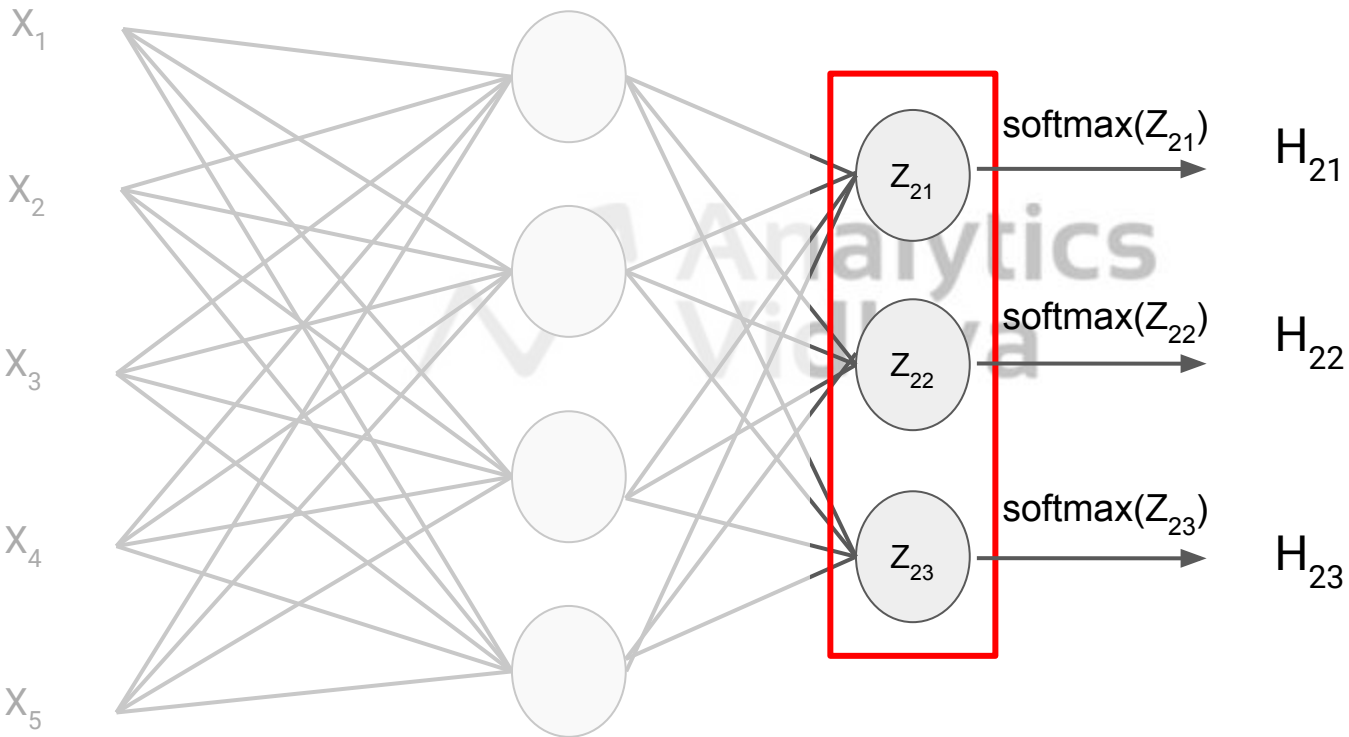
Multiclass Classification Problem: Sigmoid



Multiclass Classification Problem: Sigmoid



Multiclass Classification Problem: Softmax



Input Layer Hidden Layer Output Layer

Softmax Activation Function

Softmax Activation Function

- Returns probability for each class

Softmax Activation Function

- Returns probability for each class
- Sigmoid Activation Function: $$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$

Softmax Activation Function

- Returns probability for each class
- Sigmoid Activation Function:

$$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$

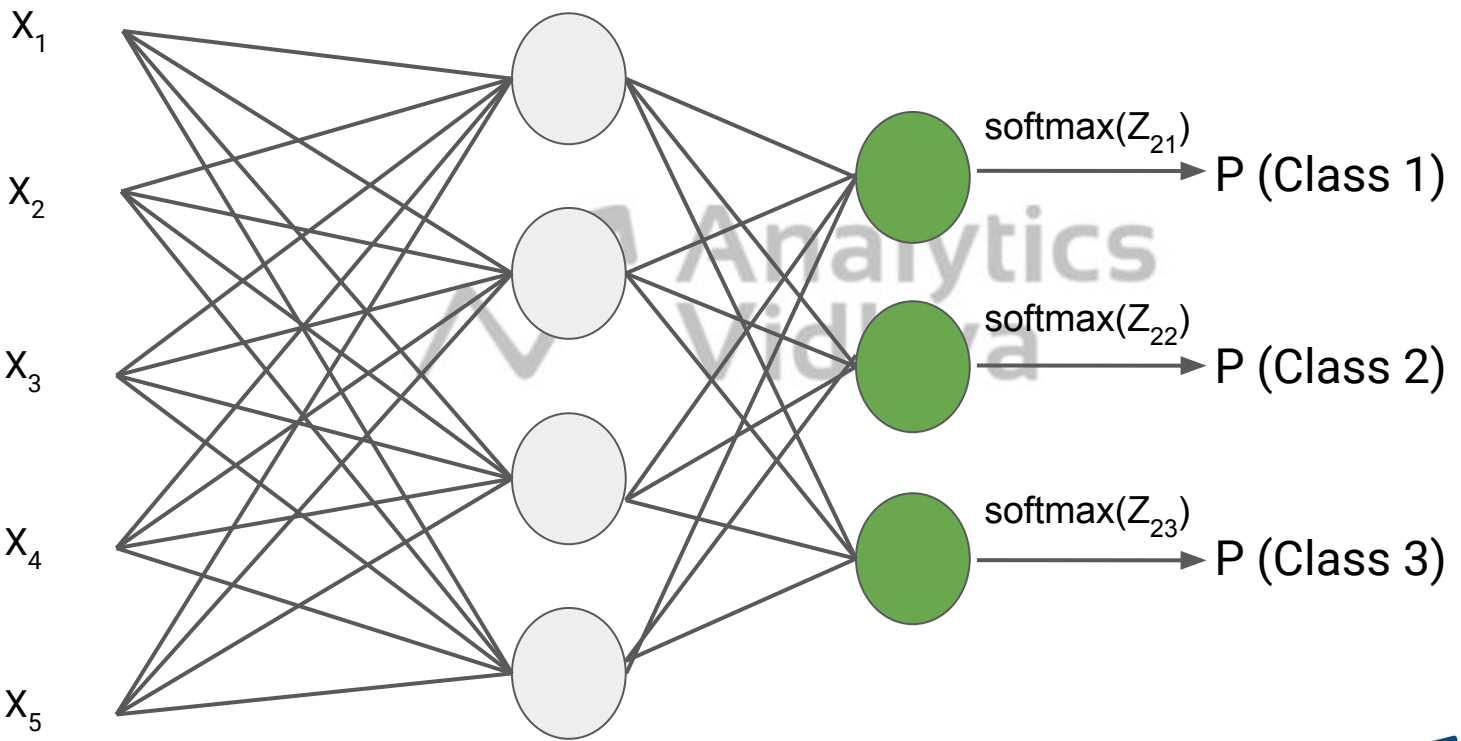
Softmax Activation Function

- Returns probability for each class
- Sigmoid Activation Function: $$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$
- Non-linear activation function

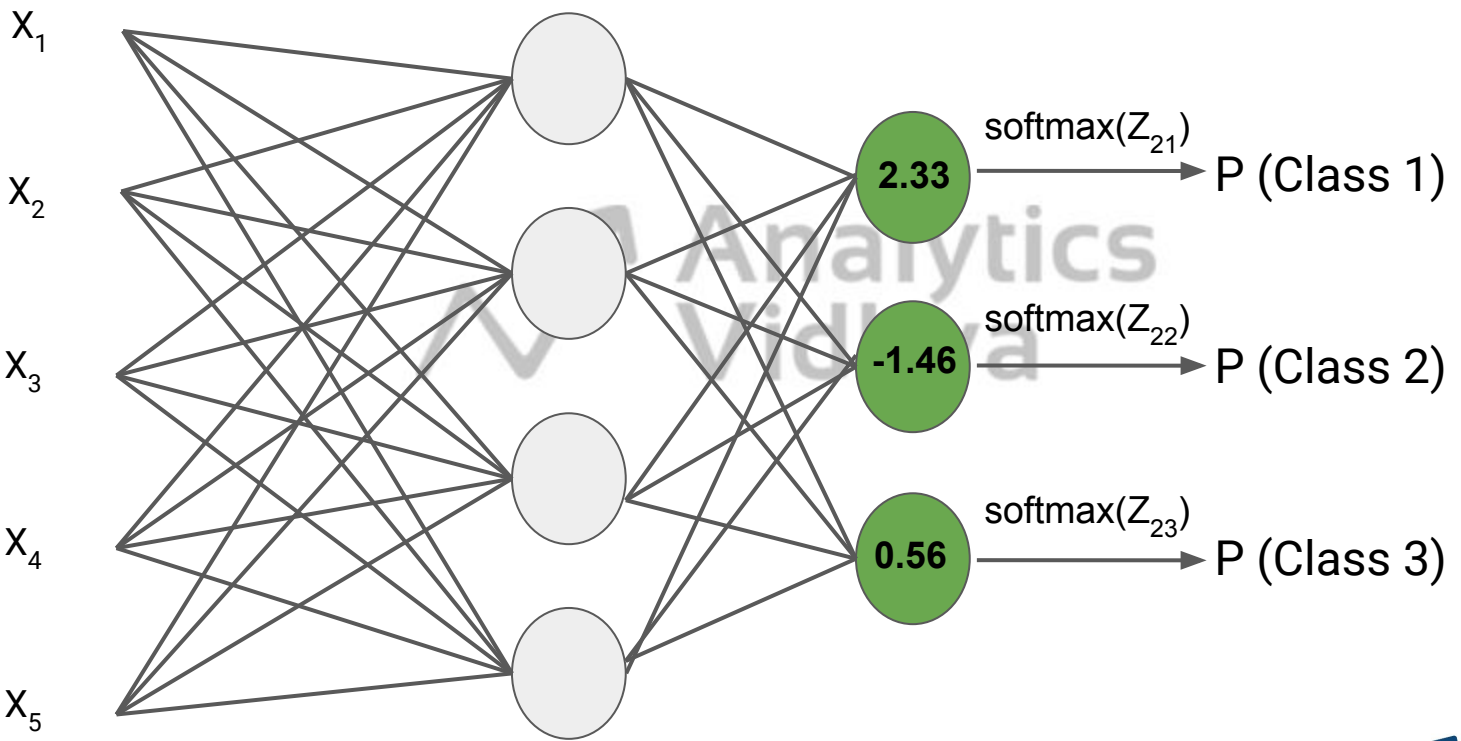
Softmax Activation Function

- Returns probability for each class
- Sigmoid Activation Function: $$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$
- Non-linear activation function
- Probability of all classes

Softmax Activation Function



Softmax Activation Function



Softmax Activation Function

Z_{21} → P (Class 1) = $\frac{\exp(Z_{21})}{\exp(Z_{21}) + \exp(Z_{22}) + \exp(Z_{23})}$

$$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)}$$

Z_{22} → P (Class 2) = $\frac{\exp(Z_{22})}{\exp(Z_{21}) + \exp(Z_{22}) + \exp(Z_{23})}$

Z_{23} → P (Class 3) = $\frac{\exp(Z_{23})}{\exp(Z_{21}) + \exp(Z_{22}) + \exp(Z_{23})}$

Softmax Activation Function

Example :

2.33 → $P(\text{Class 1}) = \frac{\exp(2.33)}{\exp(2.33) + \exp(-1.46) + \exp(0.56)} = 0.83827314$

-1.46 → $P(\text{Class 2}) = \frac{\exp(-1.46)}{\exp(2.33) + \exp(-1.46) + \exp(0.56)} = 0.01894129$

0.56 → $P(\text{Class 3}) = \frac{\exp(0.56)}{\exp(2.33) + \exp(-1.46) + \exp(0.56)} = 0.14278557$

Softmax Activation Function

Example :

2.33 → $P(\text{Class 1}) = \frac{\exp(2.33)}{\exp(2.33) + \exp(-1.46) + \exp(0.56)}$ = 0.83827314

-1.46 → $P(\text{Class 2}) = \frac{\exp(-1.46)}{\exp(2.33) + \exp(-1.46) + \exp(0.56)}$ = 0.01894129

0.56 → $P(\text{Class 3}) = \frac{\exp(0.56)}{\exp(2.33) + \exp(-1.46) + \exp(0.56)}$ = 0.14278557

Thank You