**Deep Learning (ANN):**

* No. of nodes in Input layer are equal to no. of features in all cases (Binary & Multiclass)
* In other cases, the no. of category (in labels) decide no. of nodes in output layer.
* No. of nodes in hidden layer = ( 2/3 \* nodes in input layer )+ node in output layer
* There is no method to decide no. of hidden layer

More layers leads to overfitting

Less layers leads to underfitting

* o/p = af (input \* weight +bias)

Activation Function:

1. Relu -

* If x < 0 o/p = 0
* If x > 0 o/p = x
* Used on Hidden Layer

1. Sigmoid –

* If probability < 0.5 o/p = 0
* If probability > 0.5 o/p = 1
* Used in Binary Classification (0/1 , True/ False , Yes/No) in labels
* Used on Output Layer
* Only one node is sufficient in output layer (Binary Classification)

1. SoftMax –

* Used for Multiclass classification in labels
* Used on output layer
* The no. of category (in labels) decides the no. of nodes in output layer.
* In case of Regression no activation function on output layer is required