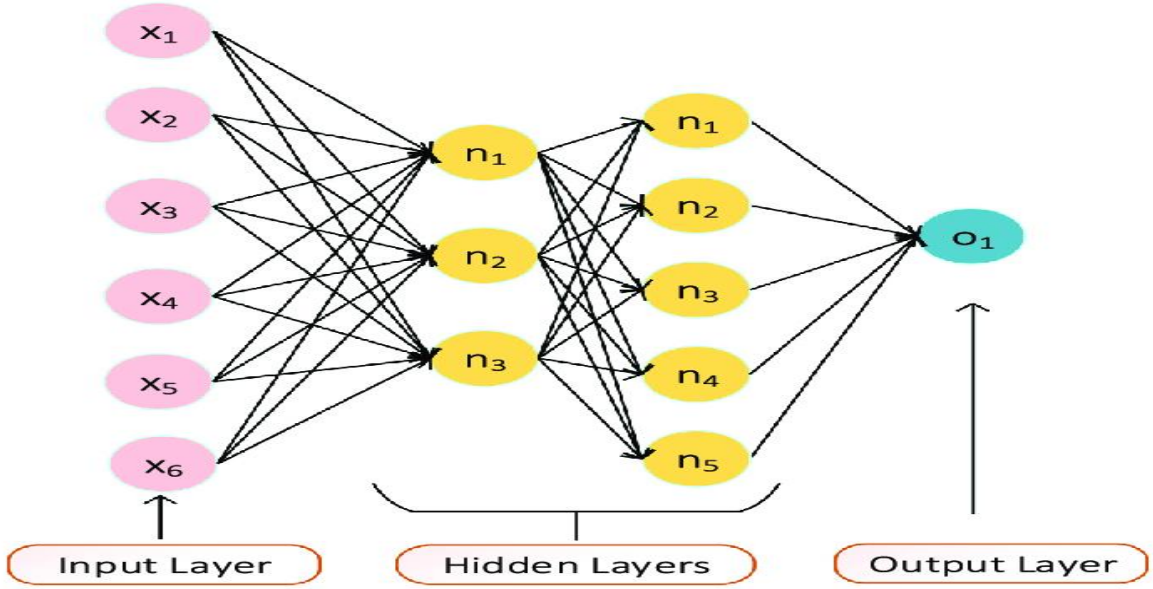
What is ANN ?

Artificial Neural Network (ANN): A computer model that works like a tiny version of the human brain. It has layers of “neurons” that learn patterns from data to make decisions or predictions.



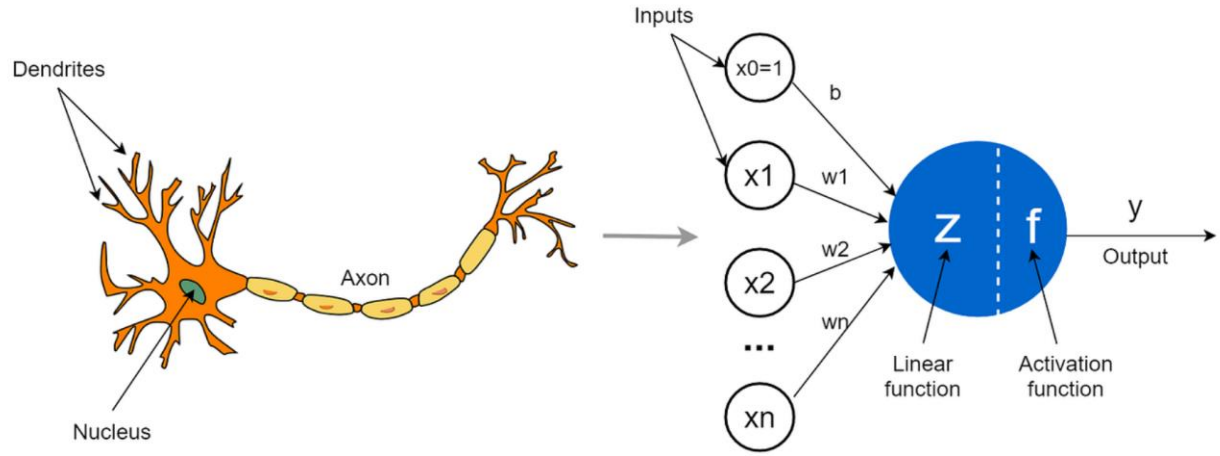
It consist of the following layers :

- Input layer
- Hidden layer
- Output layer

The human brain consist of neurons. In Artificial Neural Network has concept of Perceptron. In the next slide we will study perceptron and its types(single layer perceptron and multiple layer perceptron)

What is Perceptron / (single layer perceptron)?

Perceptron: The simplest type of neural network. It takes some inputs, gives them weights, adds them up, and then decides an output (like yes/no).

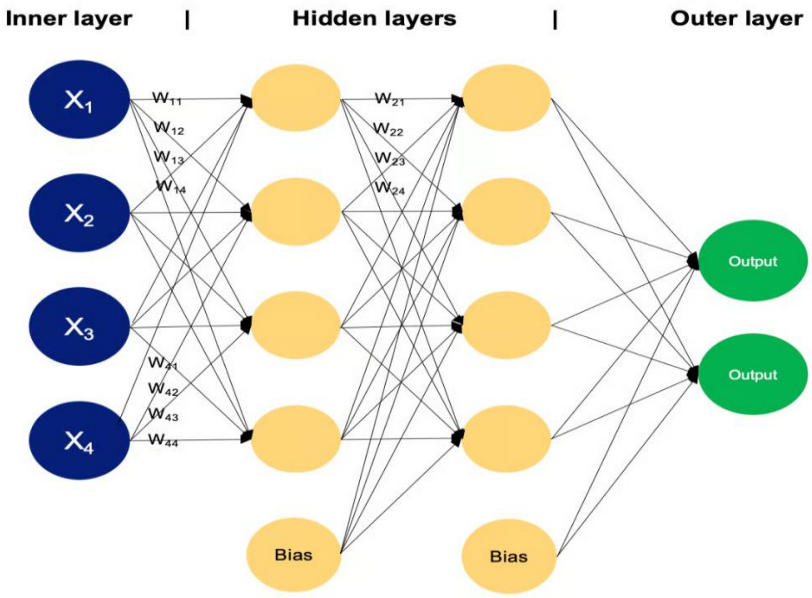


Through this we can solve different problem like regression classification we already studied that in the previous lectures.

It makes that algorithms more perfect.

What is Multi-Layer Perceptron or ANN ?

Multi-Layer Perceptron (MLP): A neural network made of many perceptrons arranged in layers (input, hidden, output). It can learn complex patterns, not just simple ones.



These perceptron takes input data or training data trained the network on that data then we give new unseen data and it gives us result.

What is Forward Propagation ?

Forward Propagation: The process of sending input data through the layers of a neural network to get an output (prediction or decision)

Input are multiply by weights and bias are added, applies activation and output flow to the next layer.

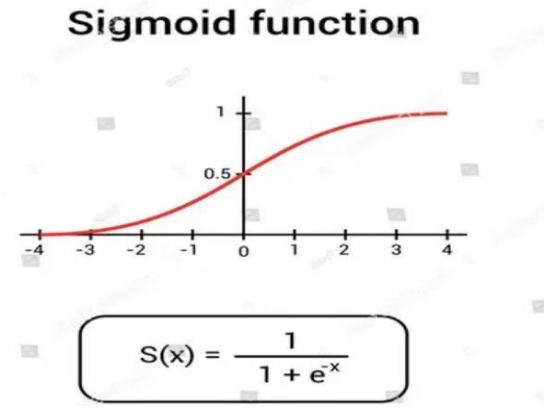
In simple words the prediction process as known as forward propagation.

Activation Function :

Activation Function: A rule that decides whether a neuron should “fire” (activate) or not. It adds non-linearity so the network can learn complex patterns.

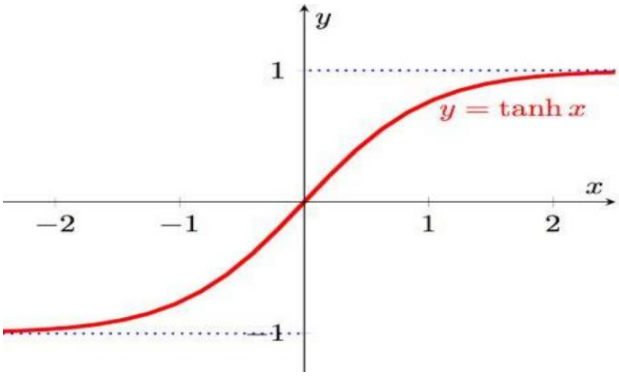
- Sigmoid function :

  Turns the output between range of 0 and 1 make the output probabilistic.
  Graph:



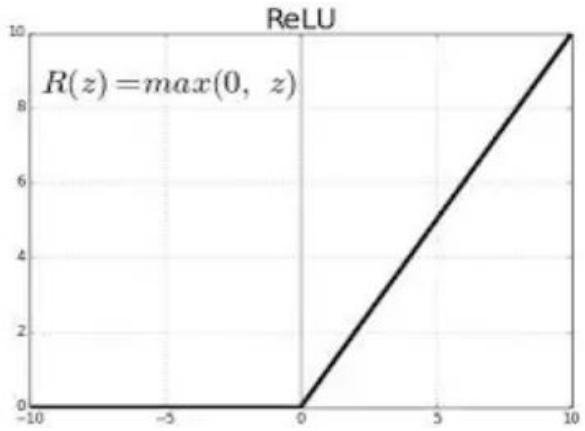
- Tanh function :

  Similar to sigmoid but output is in the range of -1 and 1, helps in centering data.
  Graph:



- ReLU (Retified Linear Unit) :

  If the input is positive, it keeps it, if negative it gives zero, it fast and most widely use function.
  Graph:



- Softmax :

  Convert all number into probabilities that all add up 1, use in classification. The graph is similar to sigmoid or tanh.

Loss function :

It measures prediction error (e.g: MSE) to guide learning.

There are different types of loss functions like.

- Binary_crossentropy.
- Categorical_crossentropy, etc.

What is Backward propagation ?

Backward Propagation: A learning step where the network checks the error in output, then sends it backward through the layers to adjust weights and improve accuracy.

For updating each neuron output depends on the back output which it takes is an input.

Gradient Descent :

Gradient Descent: A method used by neural networks to learn. It slowly adjusts the weights step by step in the direction that reduces the error, like moving downhill to reach the lowest point.

e.g: SGD,adam.

see types and more explanation in linear regression lecture

Learning rate:

Controls step size in updates; too high overshoots, too low slows convergence.

Epoch :

One full pass over the training data; multiple epochs improve learning until convergence.

Overfitting :

Model fits training data too closely and fails to generalize to new data.

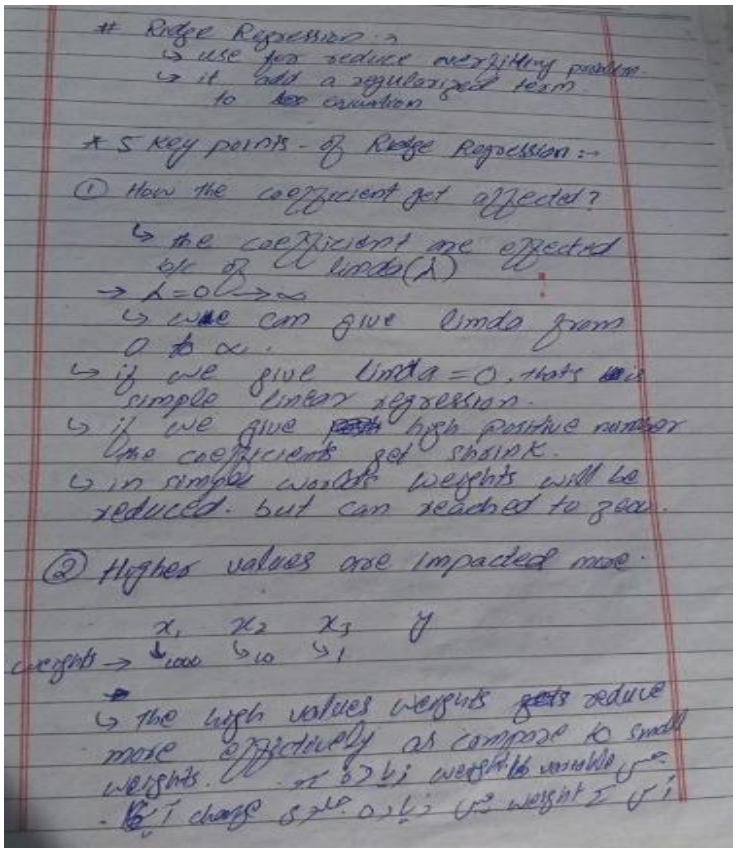
See the linear regression lecture for overfitting there I explain overfitting in details.

Regularization :

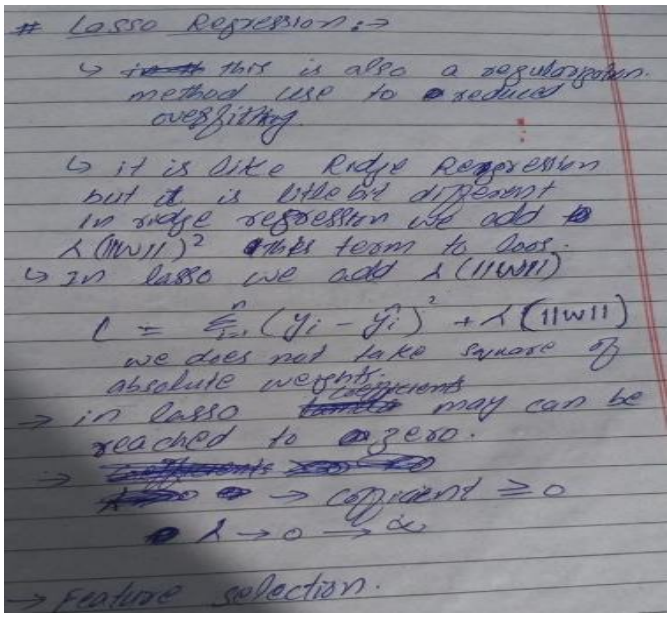
Techniques (L1/L2, dropout, early stopping) that reduce overfitting and improve generalization.

Types of regularization.

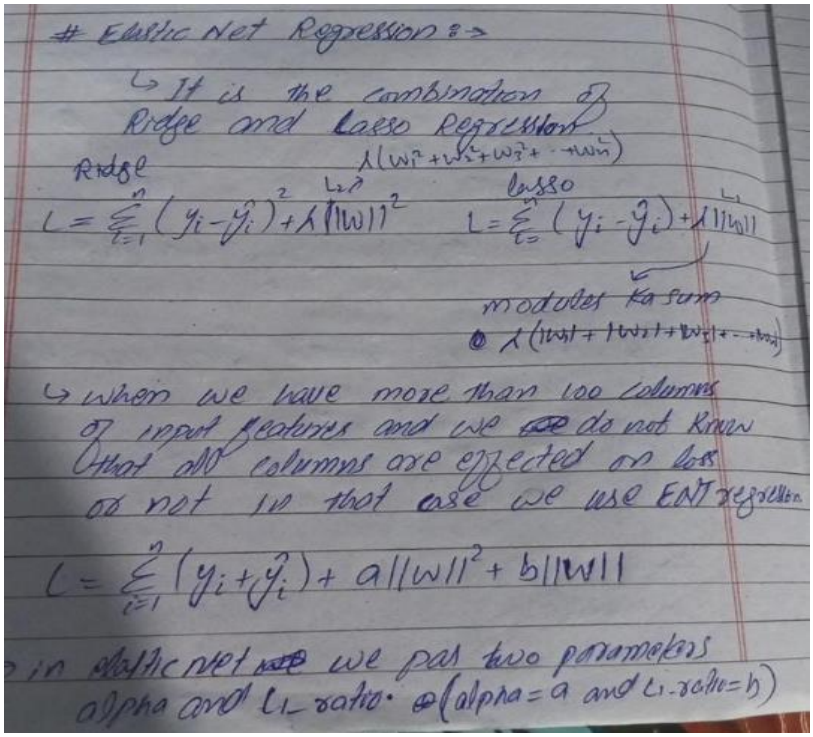
- Ridge regression(L2) :



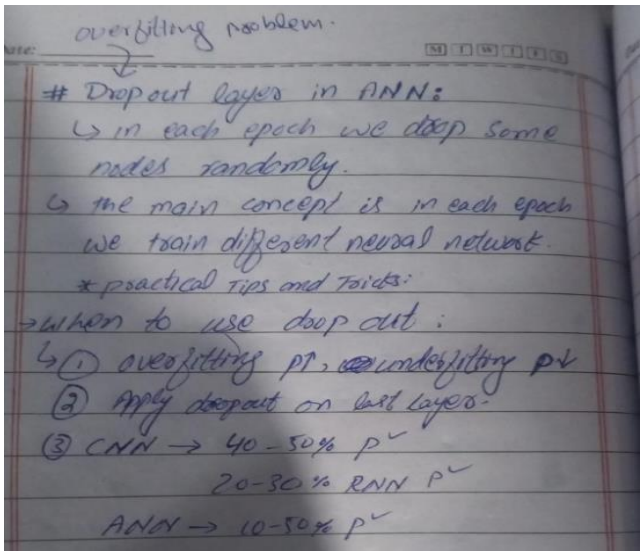
- Lasso regularization(L1) :



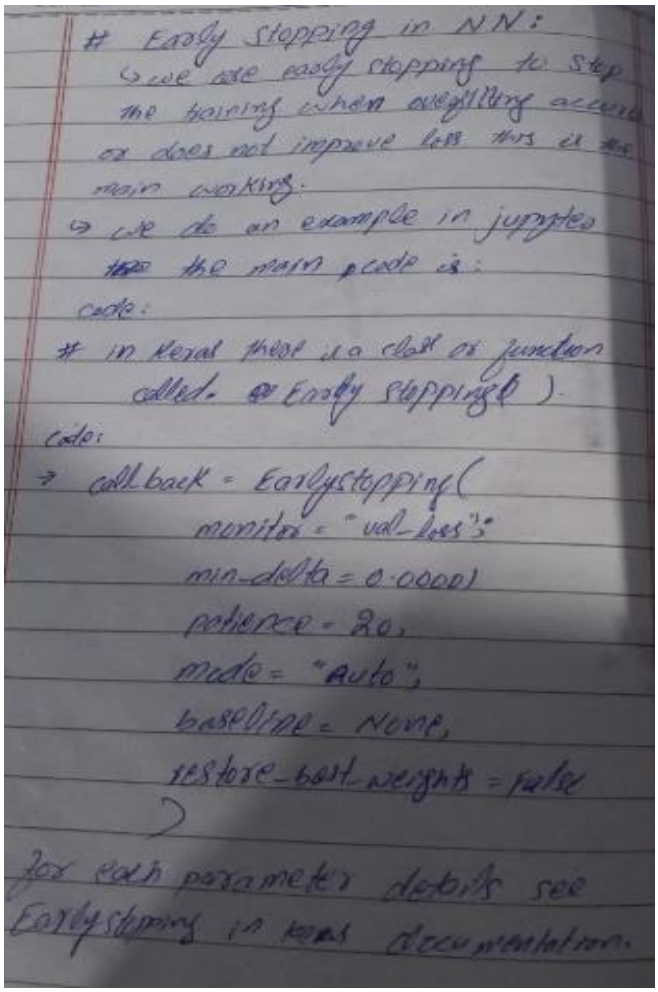
- ElasticNet Regression :



- DropOut Layer :



- Early Stoping :



Prepared by Sayyad khan

End of lecture