## horizontal line

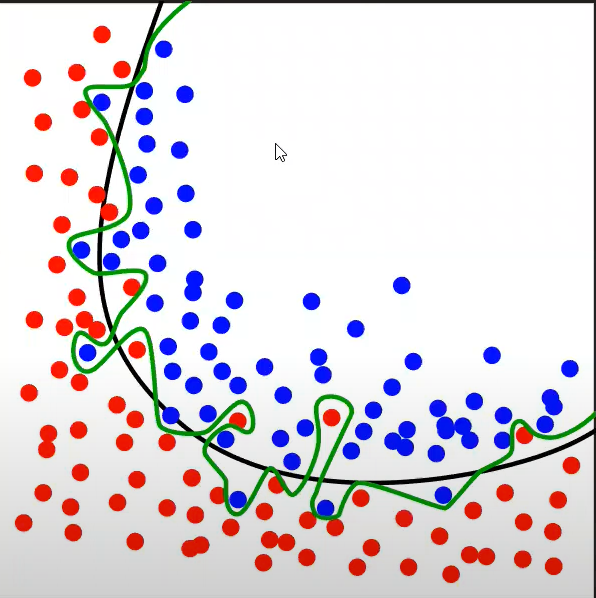
Dropout Layers

25.01.2025

[Dropout in classification](https://colab.research.google.com/drive/1UfGQhNcO1D8GF2huMfQvwIGQxLriOw7H)

[Dropout in regression](https://colab.research.google.com/drive/1YQisUjo7Xa7CNIXzQM_zB_5NZNmxr5Ci)

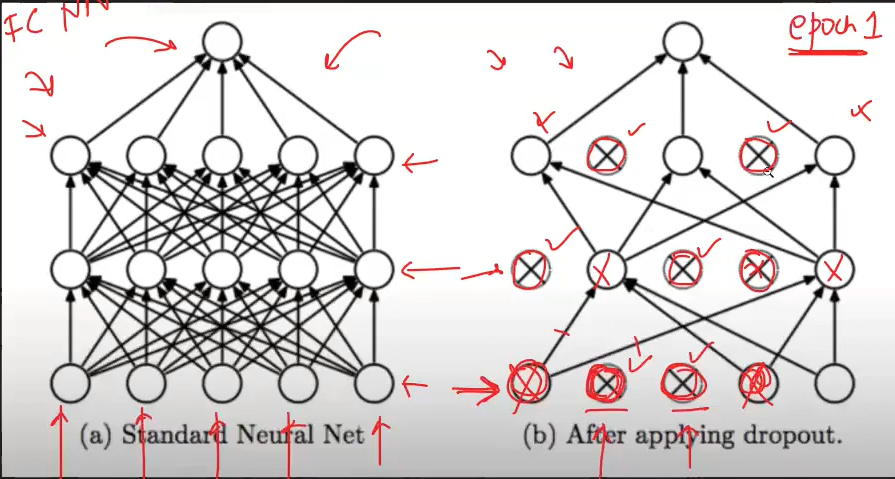
# Problem of Overfitting

Overfitting is a condition when a model learns the minute patterns of the training data , not the concept . In that case it performs well on training data but poorly on test data.

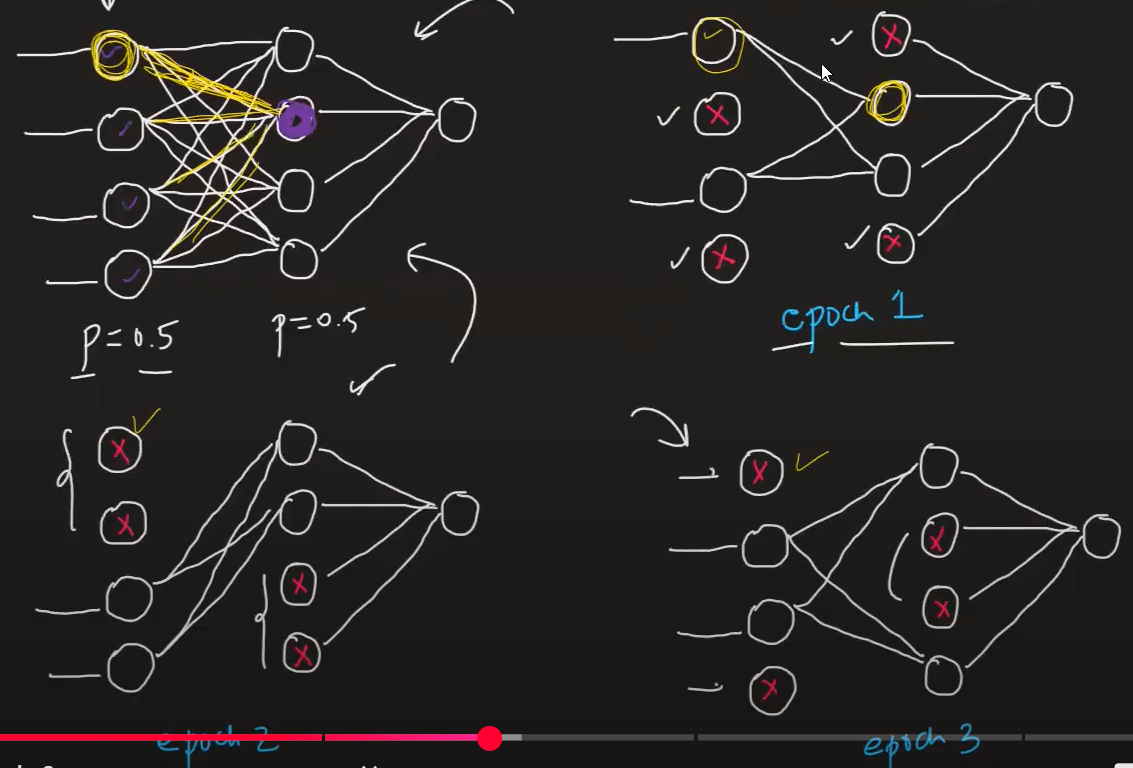
Possible Solutions can be :

1. Add more data
2. Reduce complexity
3. Early Stopping
4. Regularization (L1 and L2)
5. Dropout

Dropout means dropping out a percent of nodes in layers , every time we have a unique set of nodes in the layers this means the model focuses on all the nodes equally not on few ones.



This works because



Every time a unique set has been taken , now it learns from all nodes equally. Outliers are being ignored as weights didn’t repeat they either come or not in the next epoch.

# Dropout same as Random forest analogy

We have X1 , X2 , X3 - - - - - - - - - - - X10 .

We will make 100 DTs and feed them on the basis of row/column sampling .

50% of the data given to the model in 1 epoch , for next epochs data changes randomly

Thus every time model trains on a uniquely different set .

Test data row given to all 50 DTs and prediction that comes by majority be the output.

Same as in Dropouts .

P is the dropout ratio.

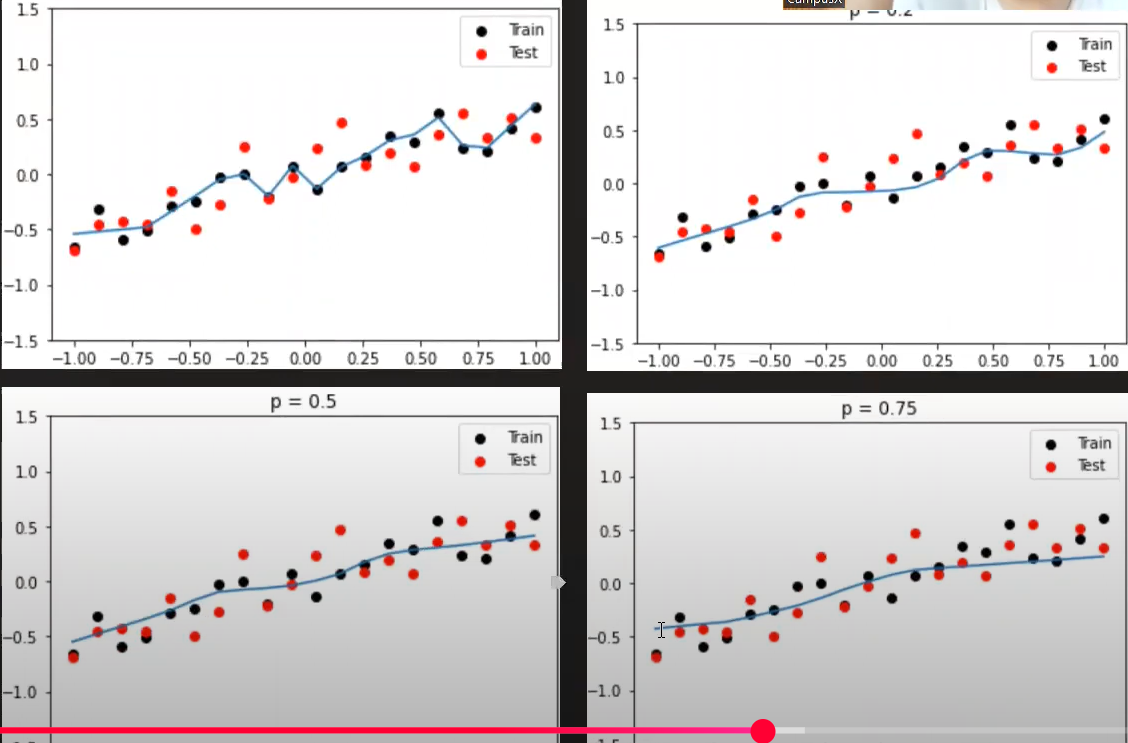
# How does prediction work ?

One weight didn’t feed to the model completely over all epochs .

For e.g. p=0.25 means every weight used only 75 times out of 100. So for testing the weight will be:

w = w(1-p)

# Effect of p



1. Increasing value of p underfitting , decreasing causes overfitting . Best range is (0.2 - 0.5)
2. It is suggested to put dropouts on the last layer . if not good output , use in between as well.
3. For CNN - 04 - 0.5 , RNN - 0.2 - 0.3 , ANN - 0.1 - 0.5 dropout ratio.

# Drawbacks

1. Convergence is delayed.
2. Difficult to calculate loss function / cost functions.