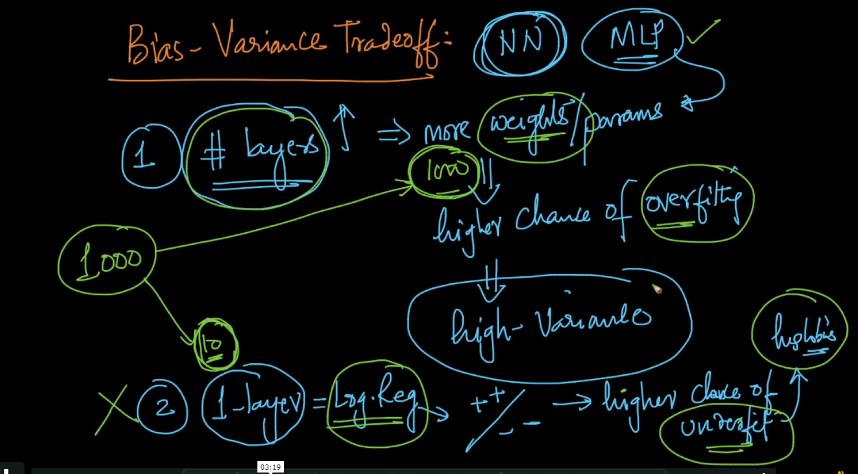
**Bias-Variance tradeoff**

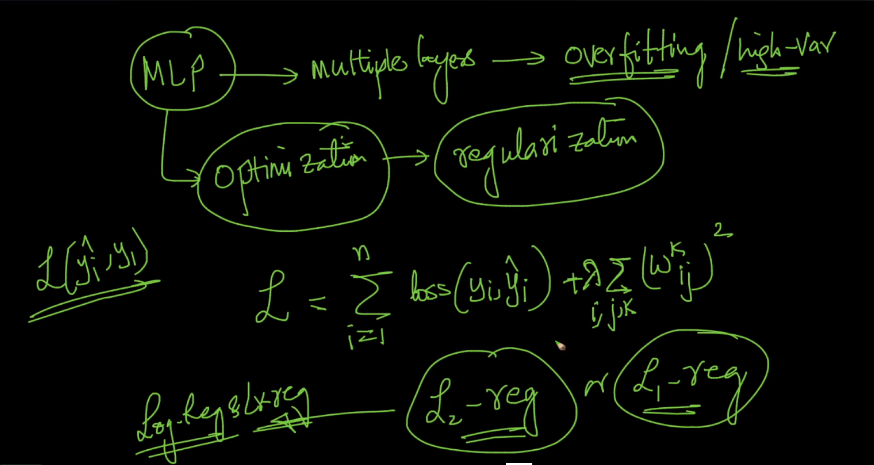
In neural network specially in multi layer perceptron as no. of layers increases then there are more weights therefore by this there are higher chances of overfitting i.e high-variance.

And if there is few layers like 1-layer then it is just like a logistic regression which have single plane to separate points therefore in this case there is higher chance of underfitting i.e high bias.

But in MLP high variance problems occurs mostly.

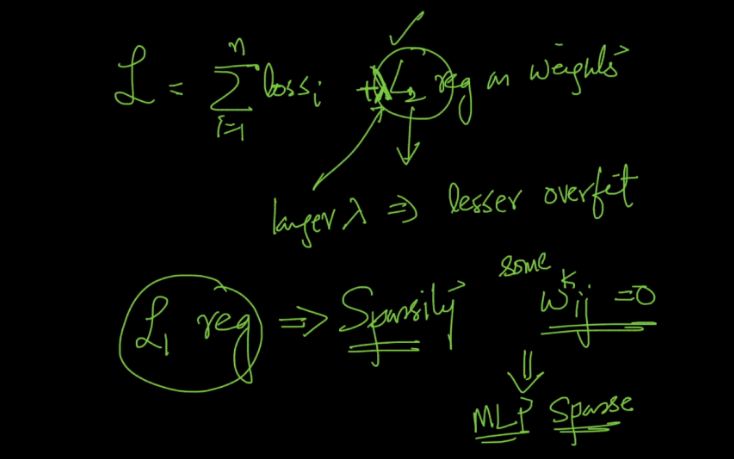


As we know MLP is also an optimization problem therefore we can remove the problem of high-variance by using regularization term either L2 or L1



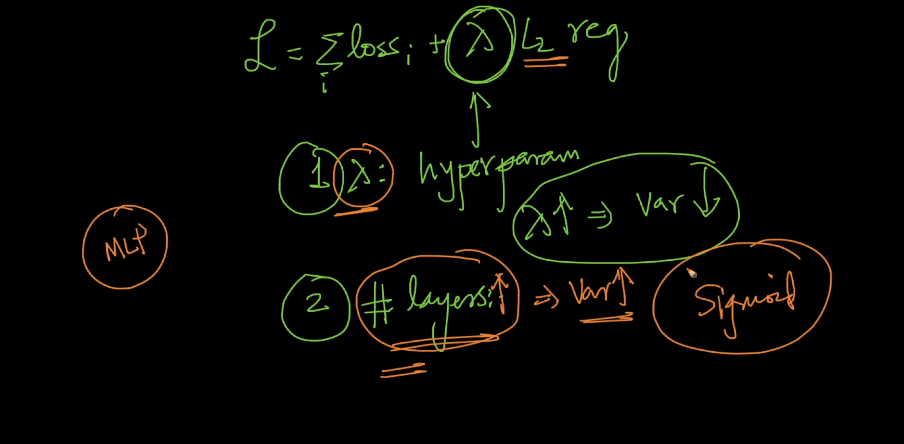
We also use one hyperparameter with regularization which is lamda, so if there is larger the lambda lesser overfit because we focus on regularization more.

If we use L1 reg. then creates sparsity means some weights becomes 0 means some connections b/w neurons gets disconnected because weights on that connection becomes 0



Therefore in MLP there are 2 hyperparameter

1. Lambda : as lambda increases, variance decreases.
2. No. of layers : as number of layers increases, variance also increases therefore we should choose it wisely.



Comments :

