author : Swapnil Narwade
date : Monday-22/11/2021
references : google & some books.

---

1. Configure Capacity with Nodes and Layers

- One can have a question such as how to choose or decide the number of hidden layers or units/nodes in a neural network to build a model.

To answer such a question we have to understand one concept called 'Capacity of the model'.

1.1 Neural Network's Capacity

- The capacity of a network refers to the range or scope of the type of functions that the model can approximate.
- "Informally, a model's capacity is its ability to fit a wide variety of functions."
- A model with less capacity may not be able to sufficiently learn the training dataset, meaning it will underfit.
- Whereas a model with too much capacity may learn or memorize the training dataset, meaning it will overfit or it may get stuck or may get lost in the process.
- Therefore we can say that the resultant model wouldn't be able to generalize.

1.2 How to define a model's capacity ?

- The capacity of the model is defined by configuring the number of nodes and the number of layers.

1.2 Width & Depth

1.2.1 Width :

- The number of nodes in a layer is referred to as the width of the network.
- In theory, a network with enough nodes in a single hidden layer network can learn to approximate any mapping function, although in practice we don't know how many nodes are sufficient or how to train such a model.

1.2.2 Depth :

- The number of layers in a neural network is referred to as the depth of the network.
- Increasing depth, increasing the capacity of the model.
- Training deep models, e.g those with many hidden layers can be computationally more efficient than training a single layer network with a vast number of units.
- By adding more hidden layer and more units within the layers, a deep network can represent a function of increasing complexity.

- Solution to above : Modern Deep Learning methods :
- Traditionally it has been challenging to train neural network models with more than few hidden layers, due to problems such as vanishing gradient descent.
- 'More recently there is an advancement in deep learning modern methodologies which allows us to train deep neural models, with large depths, and can achieve an impressive performance over the challenging problems in a wide range of domains.'