Selecting Hyperparameters

Flexibility of ANN is also a drawback
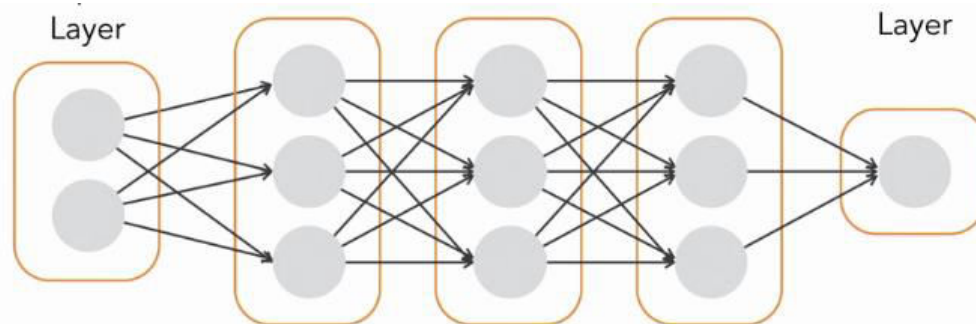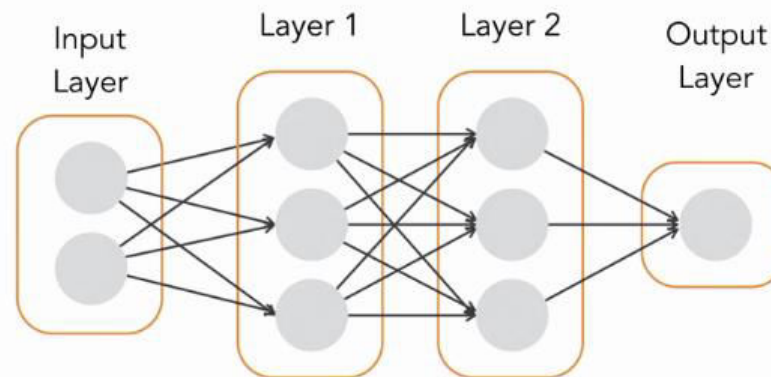
Hyperparameters

- Number of Layers
- Number of neurons per layer
- Type of activation function
- Learning rate
- Weight initialization logic
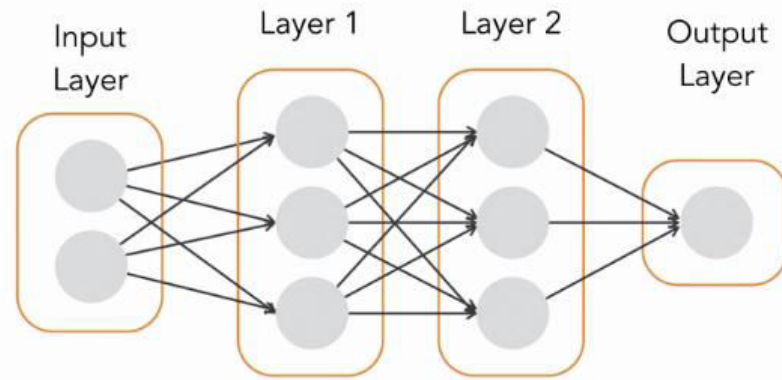- And many more

Selecting Hyperparameters

- For most problems single layer is sufficient
- But going deeper increases parameter efficiency
- Transfer Learning

Number of Hidden Layers

Selecting Hyperparameters

Number of Neurons per Hidden Layers

- \# Neurons in Input and output hidden layers are according to your data or problem
- Earlier Pyramid structures were preferred (large number of neurons in 1st layer and then gradually decreasing to output layer)
- Now same number of neurons in all hidden layer are preferred

Selecting Hyperparameters

Hyperparameters

- Learning rate
  To tune learning rate start with a large value, that makes algorithm diverge, then divide this value by 3 and try again

- Batch size
  Should be < 32 and don't make it too small

- Epochs
  Use early stopping