**A typical neural network architecture is made-up of following parts**

|  |  |
| --- | --- |
| **Hyperparameter** | **Example Value** |
| Input Layer Shape | Same shape as the number of features (e.g., in housing price predictions, if features are Size, # Rooms, # Garage, then Input layer shape = 3) |
| Hidden Layer(s) | Problem specific. Minimum 1, maximum unlimited |
| Neurons per hidden layer | Again, problem specific. Generally, 10-100 |
| Output Layer Shape | Desired prediction shape. Typically, 1 for regression problems, problem specific for classification problems |
| Hidden Activations | ReLU, Sigmoid. ReLU is popular choice |
| Output Activations | None, ReLU, tanh |
| Loss Functions | MSE (Mean Squared Error), MAE (Mean Absolute Error), Categorical cross-entropy, Binary cross-entropy |
| Optimizers | SGD, Adam |