Hyper-parameter Optimisation in Machine Learning Models

Hyper-parameter: is a configuration that is external to the model and whose value cannot be estimated from data.

Hyper-parameters can be divided into 2 categories:

- Optimizer hyperparameters,
- Model Specific hyper-parameters

Hyper-parameters help answer questions like:

1. The depth of the decision tree
2. How many trees are required in random forest
3. How many layers should a neural network have
4. The learning rate for gradient descent method

The goal of hyper-parameter exploration is to search across various hyper-parameter configurations to find a configuration that results in the best performance.

Methods:

1.Manual

2.Grid Search: lay down a grid over the space of possible hyperparameters, and evaluate at each point on the grid; the hyper-parameters from the grid which had the best objective value is then used in production.

3.Random Search: provides a statistical distribution for each hyper-parameter from which values maybe randomly sampled.

4.Bayesian Optimisation : is a SMBO(sequential Model Based) algorithm that allows for one to use results from a previous iteration to improve sampling method of the next experiment. It is based on randomness and probability distributions.