Hyperparameter tuning

A Machine Learning model is defined as a mathematical model with a number of parameters that need to be learned from the data. By training a model with existing data by training a model with existing data, we are able to fit the model parameters.

However, there is another kind of parameters, known as Hyperparameters, that cannot be directly learned from the regular training process. They are usually fixed before the actual training process begins. These parameters express important properties of the model such as its complexity or how fast it should learn.

1.The penalty in Logistic Regression Classifier i.e. L1 or L2 regularization
2.The learning rate for training a neural network.
3.The C and sigma hyperparameters for support vector machines.
4.The k in k-nearest neighbors.

GridSearchCV

In GridSearchCV approach , machine learning model is evaluated for a range of hyperparameter values. This approach is called GridsearchCV, because it searches for best set of hyperparameters from a grid of hyperparameters values.

For example , if we want to set two hyperparameters C and Alpha of Logistic Regression Classifier model, with different set of values. The gridsearch technique will construct many versions of the model with all possible combinations of hyperparameters, and will return the best one.

As in the image, for C = [0.1, 0.2, 0.3, 0.4, 0.5] and Alpha = [0.1, 0.2, 0.3, 0.4].
For a combination *C=0.3 and Alpha=0.2*, performance score comes out to be **0.726(Highest)**, therefore it is selected.

		0.1	0.2	0.3	0.4
	0.5	0.701	0.703	0.697	0.696
	0.4	0.699	0.702	0.698	0.702
	0.3	0.721	0.726	0.713	0.703
	0.2	0.706	0.705	0.704	0.701
C	0.1	0.698	0.692	0.688	0.675

Alpha

Drawback : GridSearchCV will go through all the intermediate combinations of hyperparameters which makes grid search computationally very expensive.

RandomizedSearchCV

RandomizedSearchCV solves the drawbacks of GridSearchCV,as it goes through only a fixed number of hyperparameter settins. It moves within the grid in random fashion to find the best set hyperparameters. This approach reduces unnecessary computation.