DSC540\_0500\_Topic8\_Ensemble Classifier

Ensemble learning is the process by which multiple models, such as classifiers or experts, are strategically generated and combined to solve a particular computational intelligence problem. Ensemble learning is primarily used to improve the (classification, prediction, function approximation, etc.) performance of a model, or reduce the likelihood of an unfortunate selection of a poor one.

Decision tree methodology is a commonly used data mining method for establishing classification systems based on multiple covariates or for developing prediction algorithms for a target variable. ... When the sample size is large enough, study data can be divided into training and validation datasets.

Bagging (Bootstrap Aggregation) is used when our goal is to reduce the variance of a decision tree. Here idea is to create several subsets of data from the training sample chosen randomly with replacement. Now, each collection of subset data is used to train their decision trees. As a result, we end up with an ensemble of different models. Average of all the predictions from different trees are used which is more robust than a single decision tree.

Boosting is another ensemble technique to create a collection of predictors. In this technique, learners are learned sequentially with early learners fitting simple models to the data and then analyzing data for errors. In other words, we fit consecutive trees (random sample) and at every step, the goal is to solve for net error from the prior tree.

Stacking permits users to expand their network capability while not the trouble of managing multiple devices

Ensemble methods improve classification accuracy

1 Combination of Models are used which increase the accuracy and efficiency.

2 we have k learned Models; we use the combination of these learned models to create an improved model

3 Bagging i.e. averaging the prediction over a collection of classifiers and boosting i.e weighted vote with a collection of classifiers are few Ensemble Methods which are used for improving accuracy.

Ensemble learning is referred to the process with which we can use to combine the features of the multiple machine learning model that all are strategically constructed. That is the main reason why the overall ensemble might work better than the individual classifier. and in the overall ensemble, we can customize the model according to our needs. that why a person should go with the ensemble learning compared to any other classifier.

The ensemble is the art of combining different types of or the set of different individual models that will together improve the machine learning model accuracy and it also provides the stability and predictive power of the model. and one of the important advantages of the Ensemble-based models is that we can work with the large or small both the datasets.

Here is a diagram that is showing the main idea of the ensemble: -

