



## Practical No:-7.

Title:- Boosting

- (a) Cross Validation
- (b) AdaBoost

Objective:- To study and implement.

- (a) Cross Validation
- (b) AdaBoost.

Theory:-

Boosting is a machine learning technique that combines multiple weak learners to create a strong learner. It aims to improve the overall predictive accuracy of the model by focusing on the instances that are difficult to classify correctly.

In Boosting, weak learners are trained iteratively, with each subsequent learner attempting to correct the mistakes made by the previous one. During training, the instances that are misclassified or have higher errors are given more emphasis, while the correctly classified instances receive less attention.

The final prediction is made by aggregating the predictions of all weak learners, often using a weighted voting schema where the learners with better performance have higher influence.

Boosting algorithms, such as AdaBoost, Gradient Boosting, or XGBoost, have shown remarkable success in various domains, achieving state-of-the-art results in tasks such as classification, regression, and ranking. They are known for their ability to handle complex relationships in data and handle both numerical and categorical features effectively.

### Types of Boosting

- (1) Adaboost (Adaptive Boosting)
- (2) Gradient Tree Boosting
- (3) XGBoost.

### (a) Cross Validation

Cross validation is a technique used in Machine Learning to assess the performance of a model on unseen data and to tune its hyperparameters. It helps in estimating how well a model will generalize to new, unseen data.

The process involves splitting the available data into multiple subsets or folds. The model is



then trained on a combination of these folds and evaluated on the remaining fold. This process is repeated on a combination of these folds. The performance metrics obtained from each iteration are averaged to obtain an overall performance estimate. Cross-validation helps in identifying and mitigating issues like overfitting and underfitting, and it aids in selecting the best hyperparameters for a model.

### b) AdaBoost.

AdaBoost, short for Adaptive Boosting, is a popular ensemble learning algorithm that combines weak classifiers to create a strong classifier. The algorithm works by iteratively training a sequence of weak classifiers on weighted versions of the training data. In each iteration, the weights are adjusted to give more importance to the misclassified instances from previous iterations. This adaptive weighting allows AdaBoost to focus on the instances that are difficult to classify correctly. The weak classifiers are then combined into a final strong classifier using a weighted majority voting schema, where the classifiers with better performance have higher weight.

## Conclusion.

In conclusion, boosting is a ML technique that aims to improve the performance of a model by combining multiple weak learners into a strong learner. Two important concepts in boosting are cross-validation & AdaBoost. Cross Validation is a technique used to assess the performance of a model and tune its hyperparameters. By dividing the available data into multiple subsets and iteratively training and evaluating the model on different combinations of these subsets, cross-validation provides a reliable estimate of how well the model will generalize to unseen data.

X<sup>T</sup>  
X  
y