Decision Tree:

⇒ Decision Tree is a flowchart-like tree structure, where internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

⇒ A tree can be "learned" by splitting the source set into subset based on an attribute value test.

⇒ This process is repeated on each derived subset in a recursive manner called recursive partitioning.

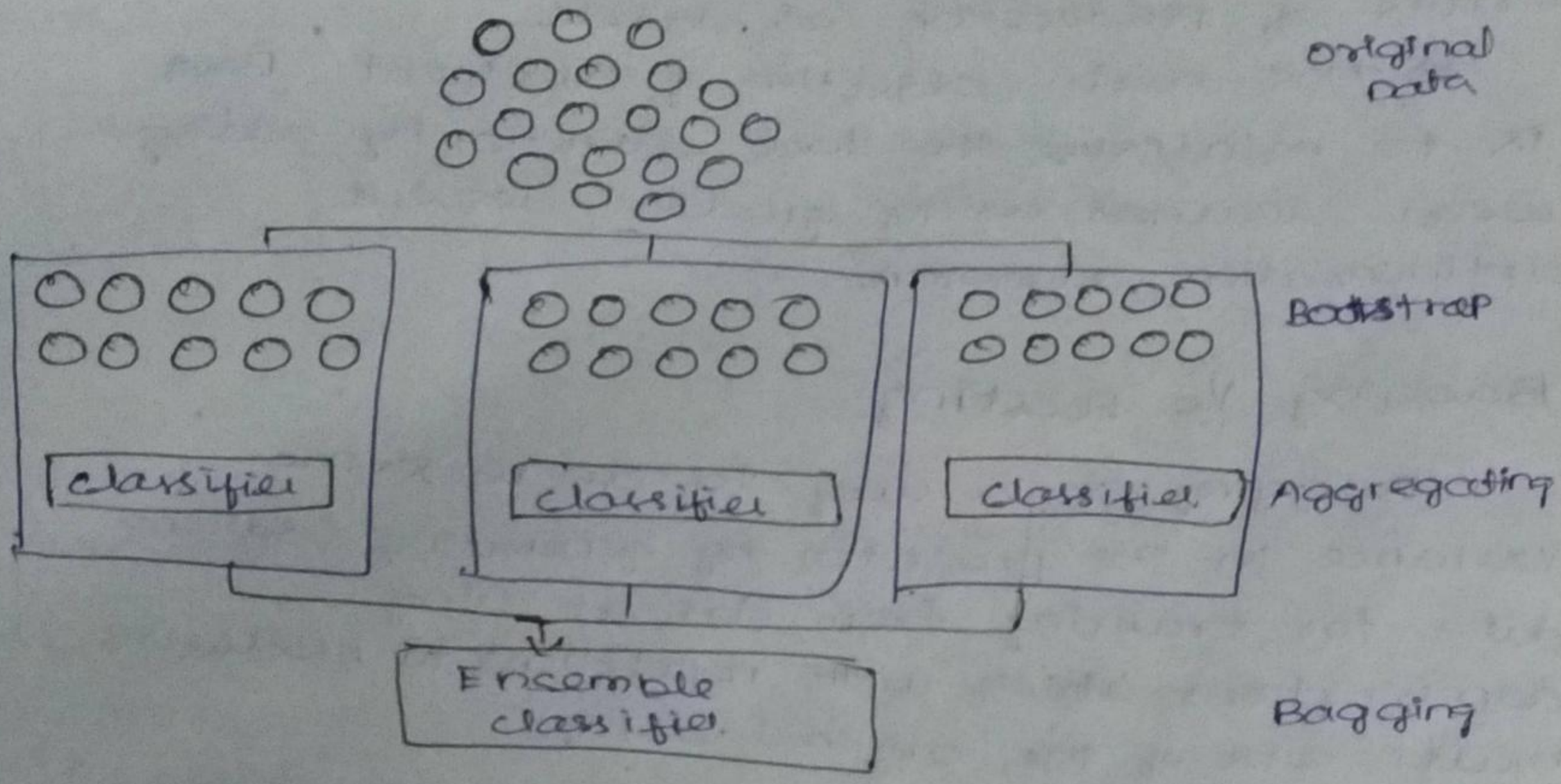
⇒ The recursion is completed when the subset at a node all has the same value of the

target variable, or when splitting no longer adds value to prediction

Bagging:

⇒ In bagging, a certain no. of equally sized subset of a dataset an extracted with replacement.

⇒ Then a ml algorithm is applied to each of these subsets and outputs are ensembled.




Bagging

Random Forest:-

⇒ Every decision tree has high variance, but when we combine all of them together in parallel then the resultant variance is low as each decision tree gets perfectly trained on that particular sample data & hence the output doesn't depend on one decision tree but multiple decision trees.

⇒ In the case of classification pblm, the final output is taken by using the majority voting classifier.

⇒ In the case of regression problem, the final output is the mean of all the outputs. This part is Aggregation.

# Boosting:

⇒ Ensemble modelling technique that attempts to build a strong classifier from the no. of weak classifier.

⇒ It is done by building a model by using weak models in series.

⇒ Firstly, a model is built from the training data. Then the 2nd model is built which tries to correct the errors present in the first model.

⇒ This procedure is continued and models are added until either the complete training data set is predicted correctly or the max no. of models are added

## Gradient Boosting:

⇒ Gradient Boosting is a popular boosting algorithm.

⇒ In gradient boosting, each predictor corrects its predecessor's error.

⇒ In contrast to Adaboost, the weights of the training instances are not tweaked, instead each predictor is trained using the residual errors of predecessor as labels.

⇒ The main objective of Gradient Boost is to minimize the loss function by adding weak learners using gradient descent optimization algorithm.

## Bagging Vs Boosting:

Bagging is a way to decrease the variance in the prediction by generating additional data for training from dataset using combinations with repetitions to produce multi-sets of the original data.

Boosting is an iterative technique which adjusts the weight of an observation based on the last classification.

## XGBoost Algorithm:

⇒ XG Boost is a decision-tree-based ensemble ML algorithm that uses a gradient boosting framework.

⇒ In prediction problems involving unstructured data (images, text, etc.) artificial neural networks tend to outperform all other algorithms or frameworks.

⇒ In this algorithm, decision trees are created in sequential form. Weights play an important role in XGBoost.

⇒ Weights are assigned to all the independent vars which are often fed into the decision tree

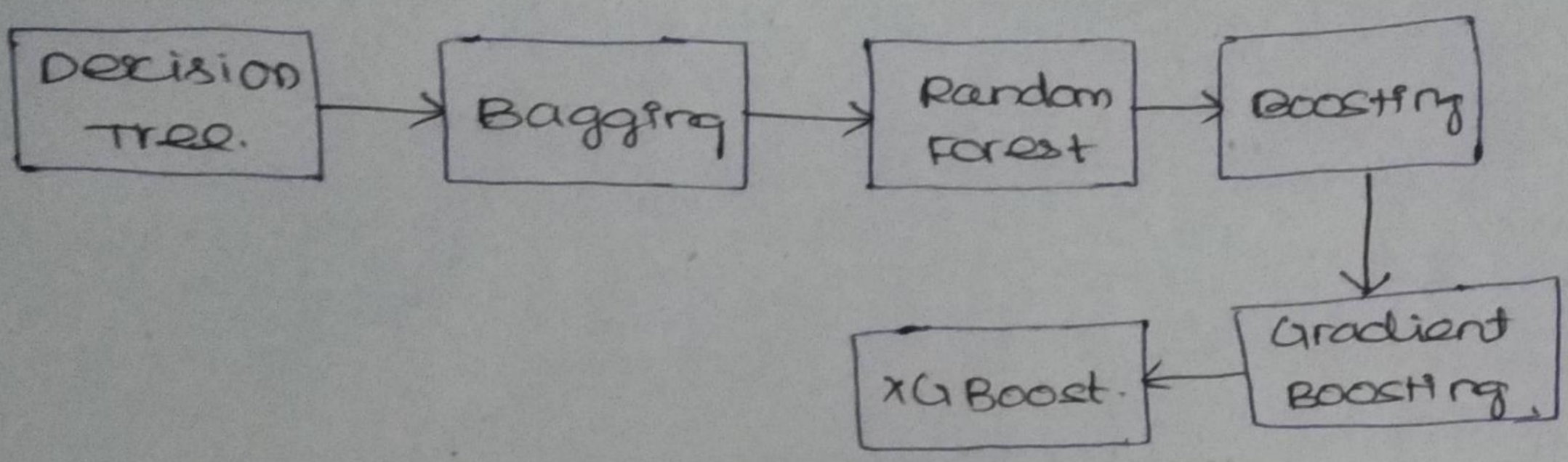
which predicts result.

→ The weight of variables predicted wrong by the tree is increased & those variables are then fed to the second decision tree.

⇒ These individual classifiers/ predictors then ensemble to give a strong and more precise model.

⇒ It can work on regression, classification, ranking, and user defined prediction problems.

Evaluation of Tree-Based Algorithms:-




How XGBoost optimizes standard GBM algorithm;

*Parallelized tree building.
*Tree pruning using depth-first approach.
*Cache awareness and out-of-core computing.
*Regularization for avoiding overfitting.
*Efficient handling of missing data.
*In-built cross-validation capability