Bagging Ensembles

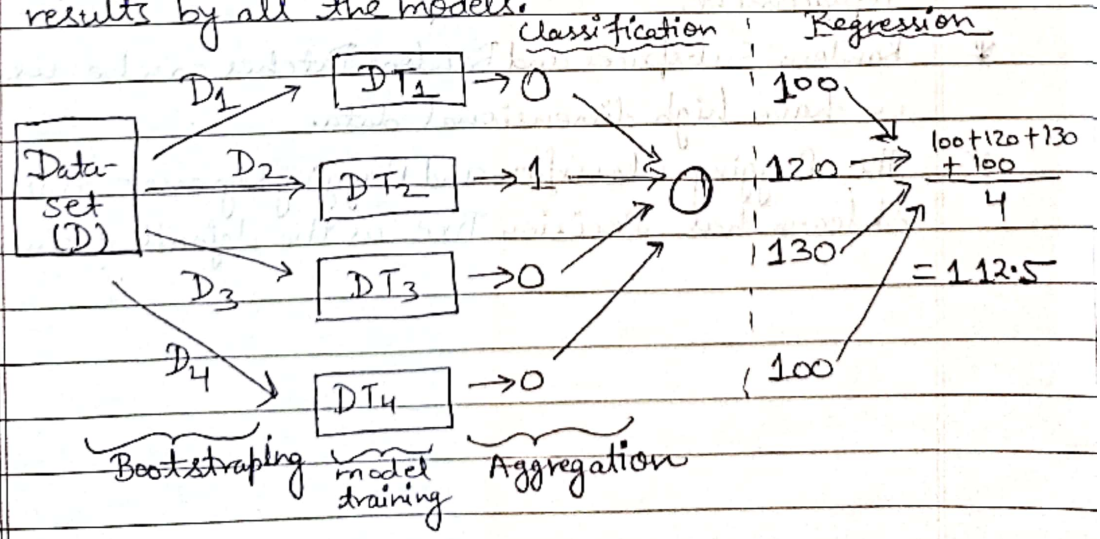
Bagging short for "Bootstrap Aggregation", is an ensemble learning technique that aims to improve the stability and accuracy of models by reducing variance and minimizing overfitting. It is particularly effective with models having low bias and high variance (like decision trees).

→ How Bagging Works :

(a) Bootstrap Sampling : From the original training dataset, multiple new datasets are created by sampling with replacement. Typically, each new dataset we create has the same size as the original dataset but it may contain duplicate examples.

(b) Model Training : In typical bagging, we use the same base model for training based on the bootstraped samples created from the original dataset. Each models differ from each other as they get trained on different samples of data.

(c) Aggregation : In classification problems, final predictions are made after taking a majority vote from prediction of all the individual models.
In regression problems, final predictions are made after combining and taking average of individual results by all the models.



→ **Key Benefits of Bagging :**

(a) Reduced Variance : By averaging the predictions of multiple models, bagging smoothens out fluctuations, thus reducing models variance. This makes the final model less sensitive to noise in the training data.

(b) Improved Accuracy : Because it reduces overfitting, bagging can often improve the accuracy of the model compared to a single model on the original dataset.

(c) Prevents Overfitting : By exposing models to different subsets of data, bagging helps to reduce risk of overfitting.

→ **Types of bagging :**

(a) Standard bootstrapping → In this different models (of same type) are trained using different subsets of data drawn using replacement.

(b) Pasting → In this models are trained using different subsets of original data drawn **without** replacement.

(c) Random Subspaces → Instead of sampling rows, this method samples features/columns. Each base model is trained on a random subset of features on all data points.

(d) Random Patches → In this we can do both row sampling as well as column sampling both with and without replacement.

* Random Subspaces and Random Patches can be used when we have high dimensional data.

* The Bagging Classifier and Bagging Regressor available in sklearn has Decision Tree as the default estimator.