**Lab 4**

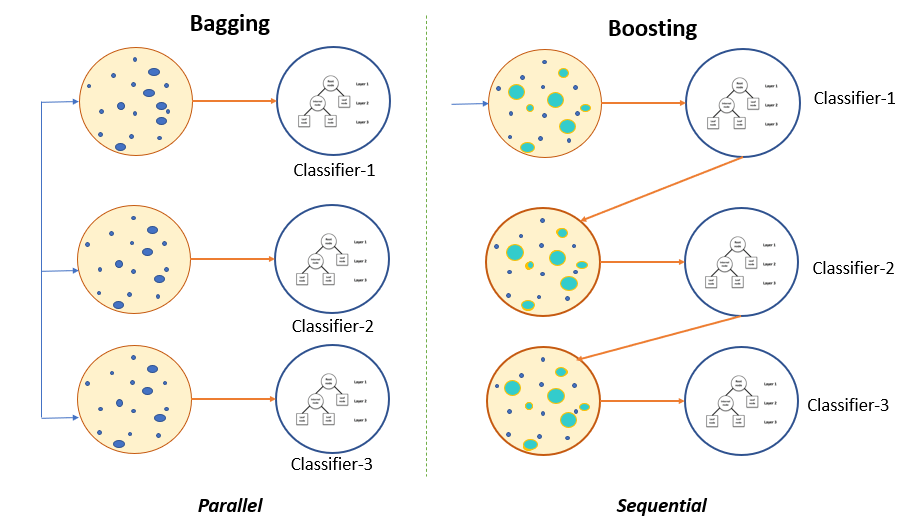
**To implement Ensemble learning (bagging/boosting)**

# Combine Model Predictions Into Ensemble Predictions

The three most popular methods for combining the predictions from different models are:

* **Bagging**. Building multiple models (typically of the same type) from different subsamples of the training dataset.
* **Boosting**. Building multiple models (typically of the same type) each of which learns to fix the prediction errors of a prior model in the chain.

In Bagging, multiple homogenous algorithms are trained independently and combined afterward to determine the model’s average. Boosting is an ensemble technique, where we train multiple homogenous algorithms sequentially. These individual algorithms create a final model with the best results. The performance of one algorithm is influenced by the performance of the previously built algorithm.



### Benefits of using Bagging algorithms

* Bagging algorithms improve the model’s accuracy score.
* Bagging algorithms can handle overfitting.
* Bagging algorithms reduce bias and variance errors.
* Bagging can easily be implemented and produce more robust models.

The Bagging technique is also known as Bootstrap Aggregation and can be used to solve both classification and regression problems. In addition, Bagging algorithms improve a model’s accuracy score. These algorithms prevent model overfitting and reduce variance.

Steps:

1. Read the database and perform pre-processing
2. Perform feature scaling (transforming a dataset to fit within a specific range)
3. Split the database into training and testing set (80:20)
4. Evaluate using single weak learner (Decision tree classifier) with 5-fold cross validation
5. Develop a boosting ensemble classifier using the weak learner (weak learner = Decision tree, number of weak learners used=100, training dataset for resampling =80, apply bootstrapping)
6. Train and evaluate this bagging model
7. Compare the performance of DT and bagging classifier.