Ensemble Techniques-Bagging Vs Boosting

- Used for both Regression and Classification

Ensemble Techniques use wisdom of the Crow.

- Crowd knows more than an individual.

- **Make sure the models used are different.**

Types of Ensemble Learning

1. Voting
2. Bagging
3. Boosting
4. Stacking

Bagging

- **Random Forest**

Boosting

- AdaBoost
- Gradient Boosting
- **XGBoost** (VIMP)
- CATBoost

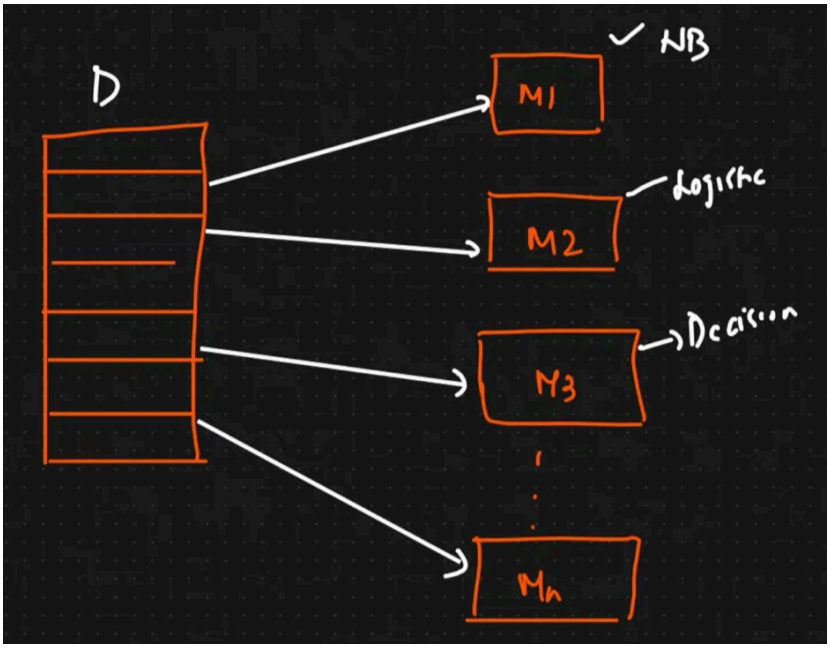
Bagging

- Bootstrap Aggregation
- **Reduces overfitting** (**reduces variance**)
- It trains various models

💡 Ensemble of **DTs** is called **Random Forest.**

Ensemble of other models is called **Bagging**.

💡 📌 **Key Idea:** Train multiple models independently on different random subsets of data and average their predictions to reduce variance.



Bootstrap Sampling: Randomly sample the dataset **with replacement** to create multiple training sets.

- You'll get an accuracy
- Calculate **mean of the accuracies** from all models.
- Classification:
  - Prediction is calculated based on highest number (**like voting**)
  - **Maximum Voting Classifier**
    - If 3 models say yes and 1 model says No, the result will be **Yes**.
- Regression:
  - Average value of all the models

- **For regression: Average the predictions of all models.**
- **For classification: Take a majority vote of all models.**

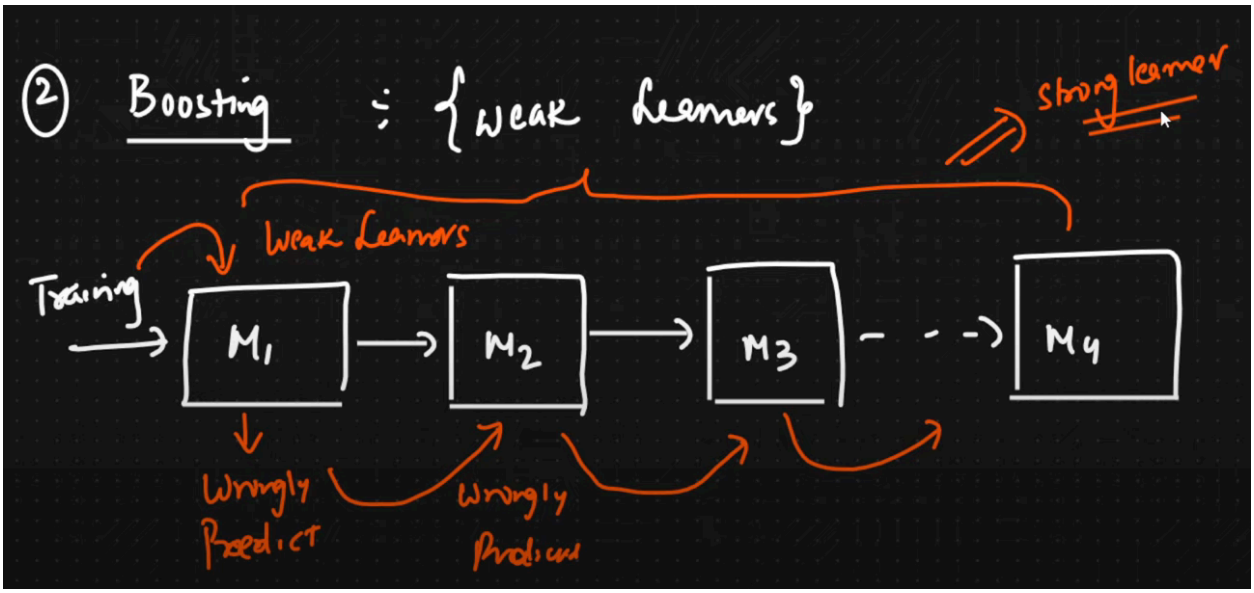
📌
Goal: Reduce variance and prevent overfitting.

Example Algorithm:

- **Random Forest** (Bagging with Decision Trees)

Boosting

- **Weak learners**
- Models are arranged sequentially



- Each model tries to correct the mistakes of the previous model.
- The final prediction is a **weighted sum** of the predictions of all models.

How Does Boosting Work?

1. **Train the First Model**:
   - Train a weak model (e.g., a shallow Decision Tree) on the training data.
2. **Calculate Errors**:
   - Measure the errors made by the model.
3. **Train the Next Model**:
   - Train a new model that focuses on the errors made by the previous model.
   - This is done by giving more weight to the misclassified samples.
4. **Repeat**:
   - Repeat the process for a fixed number of iterations or until the errors are minimized.
5. **Combine Predictions**:
   - Combine the predictions of all models using a weighted sum.

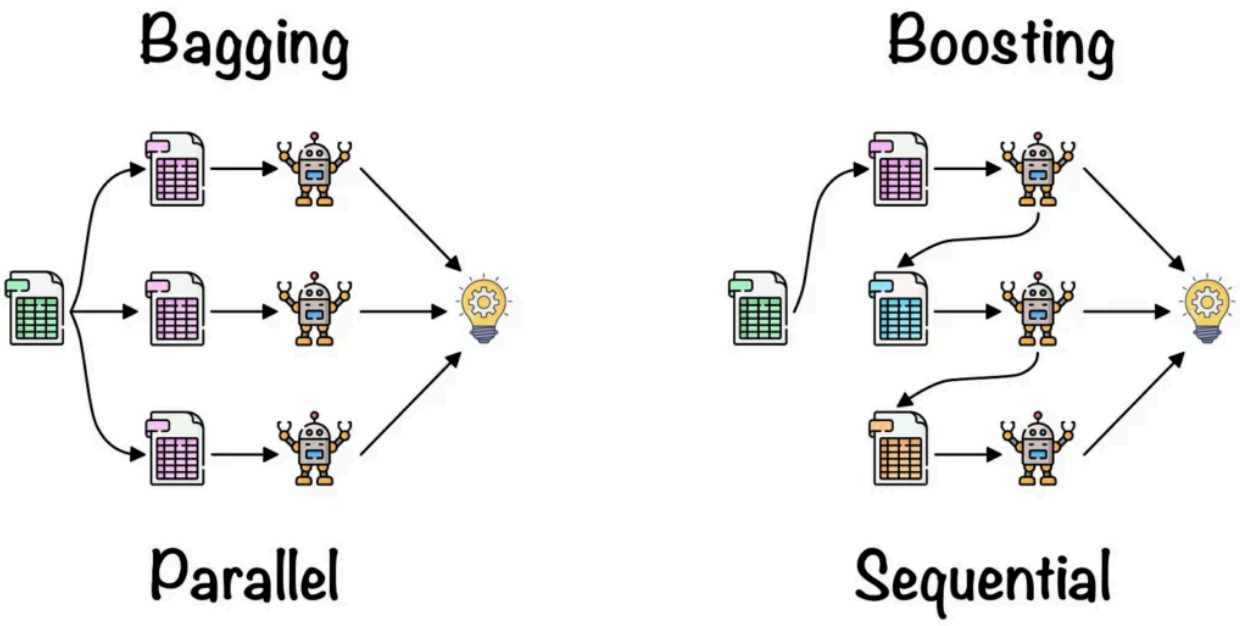
Weighted Samples: In each round of training, more weight is given to the misclassified data points from the previous model, so the next model focuses more on these difficult cases.

Key Features of Boosting

- **Reduces Bias**: Boosting focuses on correcting errors, which helps reduce bias.
- **Sequential Training**: Models are trained one after the other, so boosting cannot be parallelized.
- **Works Well with Weak Models**: Boosting is effective for models that perform slightly better than random guessing.

Bagging vs. Boosting

Aspect	Bagging	Boosting
Goal	**Reduce variance.**	**Reduce bias.**
Training	Models are trained independently.	Models are trained sequentially.
Parallelization	Can be parallelized. (independent models)	Cannot be parallelized. (models learn from each other)
Focus	Averages predictions.	Corrects errors from previous models.
Example Algorithms	Random Forest, Bagging Classifier.	AdaBoost, Gradient Boosting, XGBoost.
Best For	High-variance models (e.g., Decision Trees).	Weak models (e.g., Decision Stumps).



When to Use Bagging vs. Boosting?

- **Use Bagging**:
  - When your model is overfitting (high variance).
  - When you want to improve the stability and accuracy of high-variance models like Decision Trees.

✅
Use Bagging when you have a **high variance** model (like deep decision trees) and want to stabilize predictions.

- **Use Boosting**:
  - When your model is underfitting (high bias).
  - When you want to improve the accuracy of weak models.

✅

Use Boosting when you have **high bias** (like weak models) and want to make them stronger.