Ensembles and its techniques

* Till now we have used only one ML Algorithm / model.

① * data → Model L → train – Predict

② * data
- → M_1
- → M_2
- → M_3
- → M_4
- → M_5

→ Combine prediction.

Analogy

DS Jobs — C (ML), B (ML), A (ML)

Ajay (data Analytics)

→ one person might give you wrong advice.

→ you will connect to multiple mentors.

→ Chances of getting wrong is minimized.

* Ensembles : Combine multiple Models

: Prediction which is more stable and accurate as compared to individual Models.

Combine Multiple models

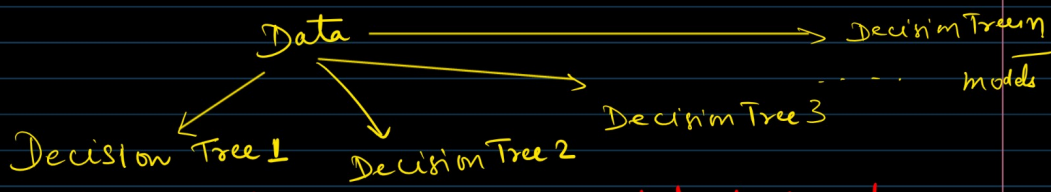
of Same Algorithm	of different algorithms
DT_1 (max depth : 5)	Logistic Regression
DT_2 (max depth : 10)	SVC
DT_3 (max depth : 12)	DTC

* Ensemble :– Not necessarily only one type of algorithm.

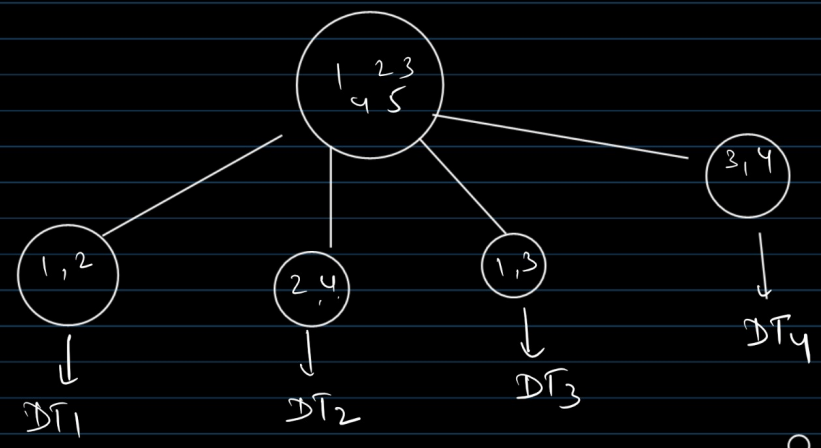
Ensemble technique

- Parallel technique (Bagging)
- Sequential technique (Boosting)
- parallel + sequential technique (Stacking)

Parallel technique of Ensembles



→ All the models are built parallely and independent of each other.



Classification: 1 1 1 0

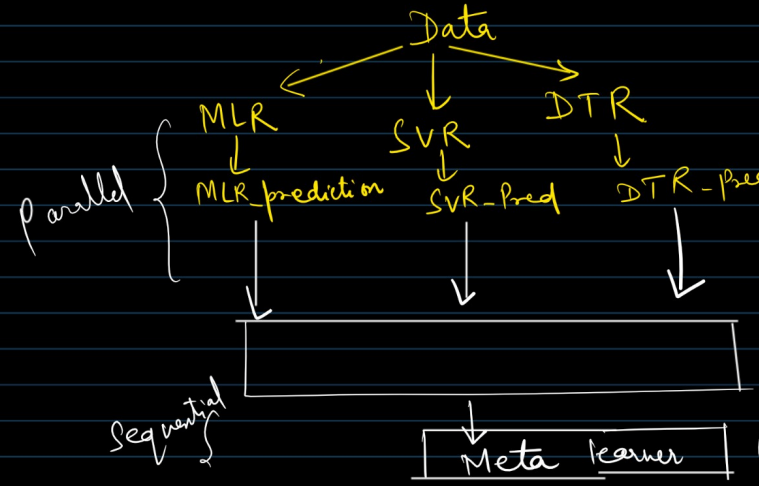
final prediction: 1

* Sequential technique of Ensembles

→ All the models are built sequentially and dependent on each other.

→ learning from mistakes of previous model.

* Parallel + Sequential technique of Ensembles



Predict the final output

Ensemble Technique

- Parallel
- Sequential
- Parallel + Sequential

Ensemble Models

- Bagging (parallel technique)
- Boosting (sequential technique)
  - Only one type of Algorithm is used
- Stacking (both technique)
  - Multiple type of algorithms

* Bagging :→ Random forest → Regressor / Classifier.
  → Custom Bagging (Different Algorithms)

* Boosting :-
  1. AdaBoost
  2. Gradient Boost
  3. XG Boost
  4. Cat Boost.

* Stacking :- We will see example of different models.