Ensemble Methods :

→ Strong learner

⇒ Combine multiple models to improve overall performance

⇒ The main idea is that using several models and combining their outputs (averaging - majority voting ...)

— individual model
(LR, SVM, DT) ⇒ Single Learner (weak learner)

⇒ Ensemble Techniques :

① Voting : Combine the predictions of multiple models by taking a majority vote (for classification) or averaging (for Regression)

→ for classification
- diff. models - same dataset.

→ hard voting (voting (mode))

↳ soft voting (avg of probabilities)

② Bagging (Bootstrap Aggregation) : Using multiple models of the same type (DT) and each one is trained on a different subset of dataset in Parallel.

→ ex. Randomforest

③ Boosting : models are built Sequentially, with each new model focusing on Correcting the errors made by the previous ones (same model - diff subsets)

→ ex. Gradient Boosting ⟶ sklearn
      AdaBoost ⟶

      Xgboost ⟶ Xgboost Library