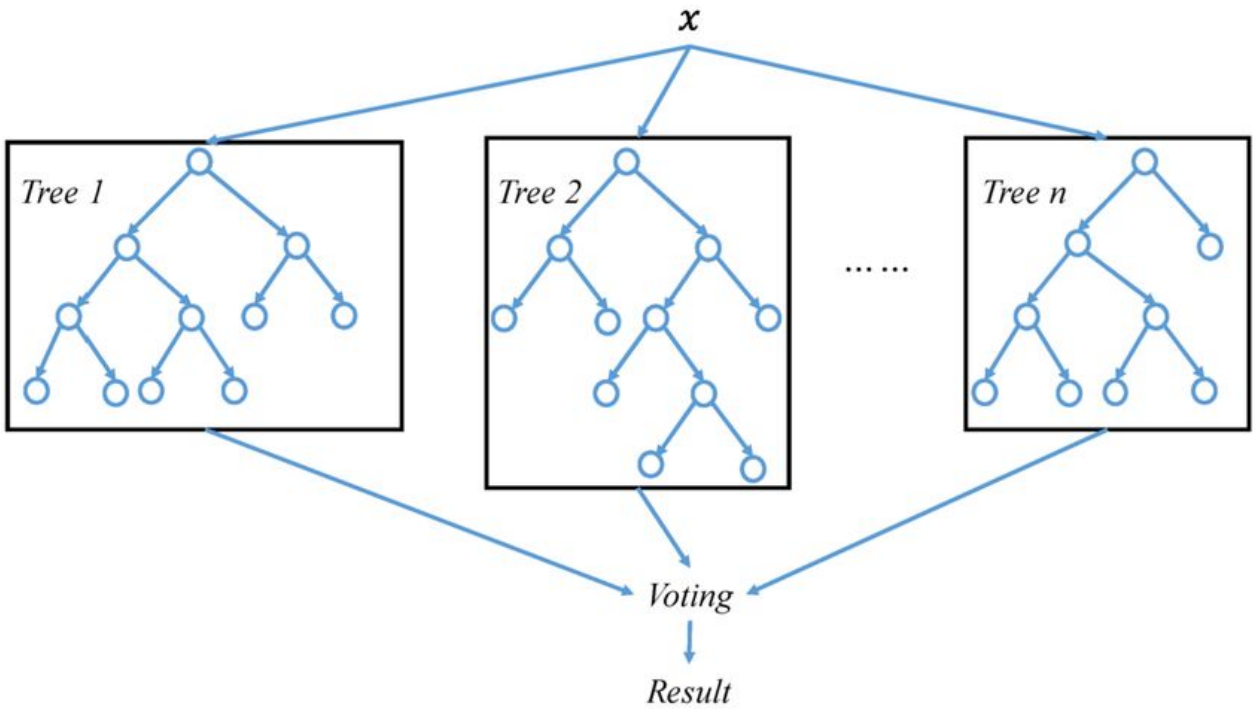
Decision Tree Ensembles

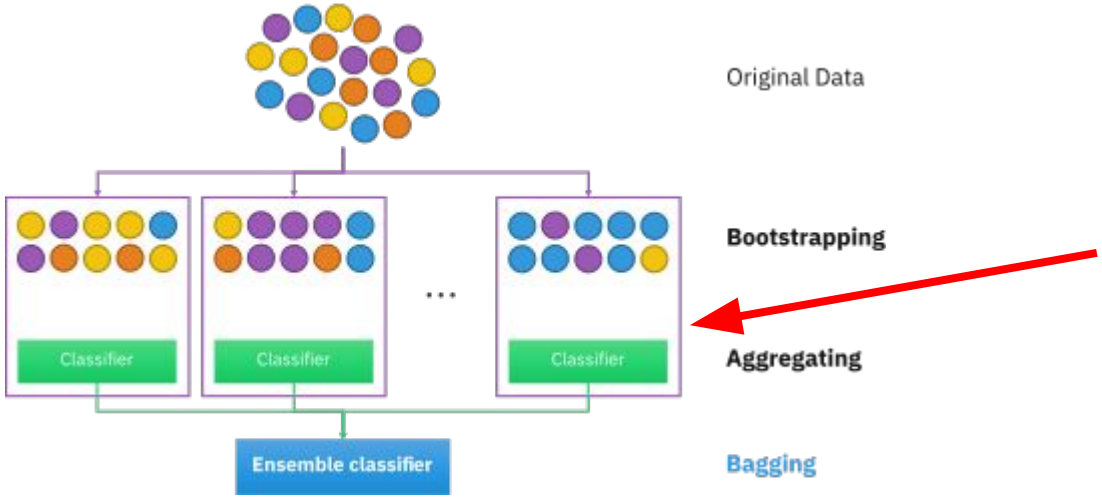
Supplementary Material

Random Forest - Parallel Ensemble Architecture



https://www.almabetter.com/bytes/tutorials/data-science/adaboost-algorithm

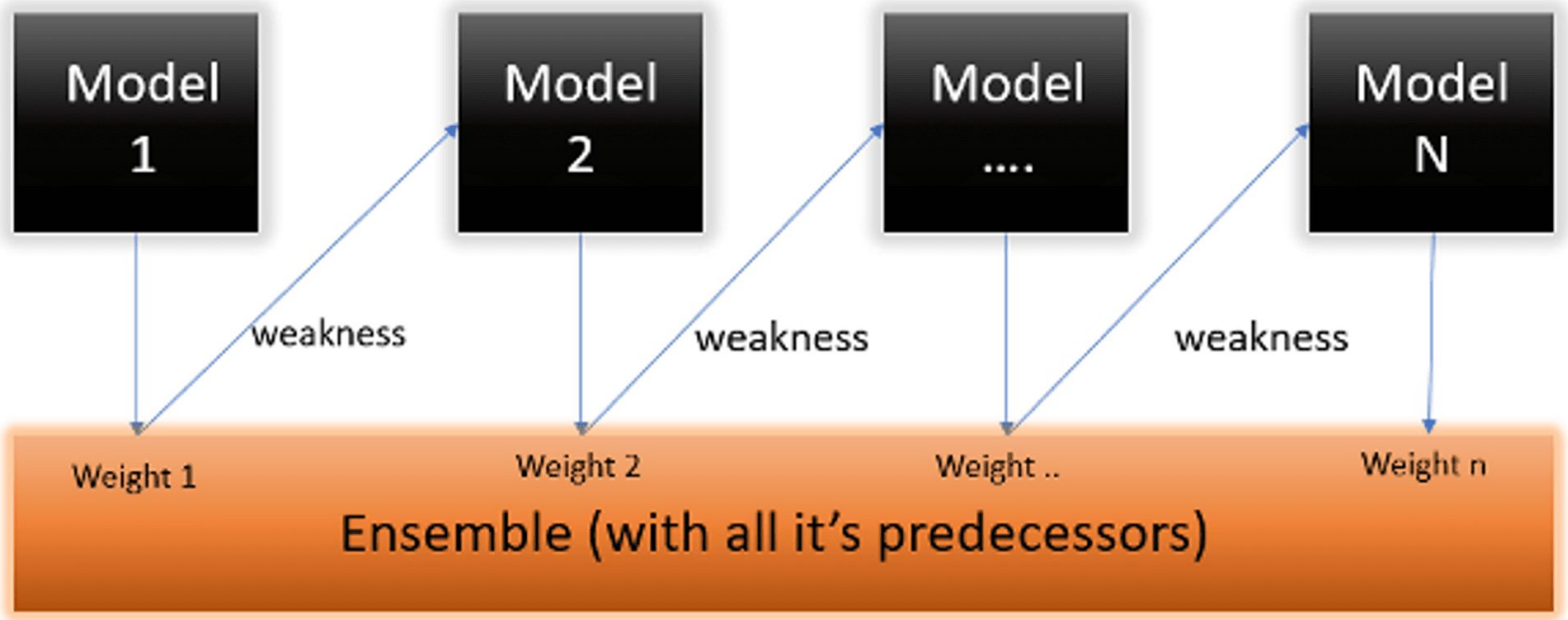
Random Forest - Training Approach



For each classifier:
1. Random sampling with replacement: unique datasets
2. Subspace sampling (random selection of features): unique features

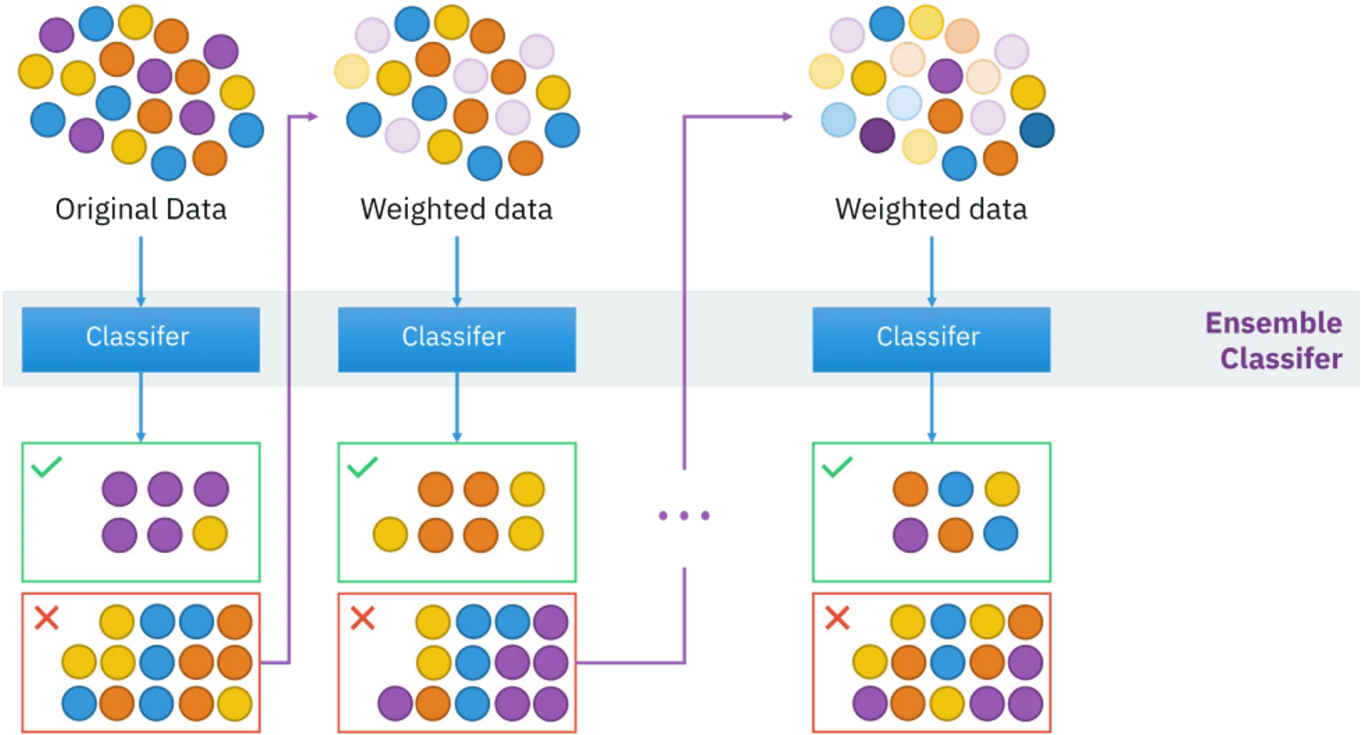
https://www.almabetter.com/bytes/tutorials/data-science/adaboost-algorithm

AdaBoost - Sequential Ensemble Architecture



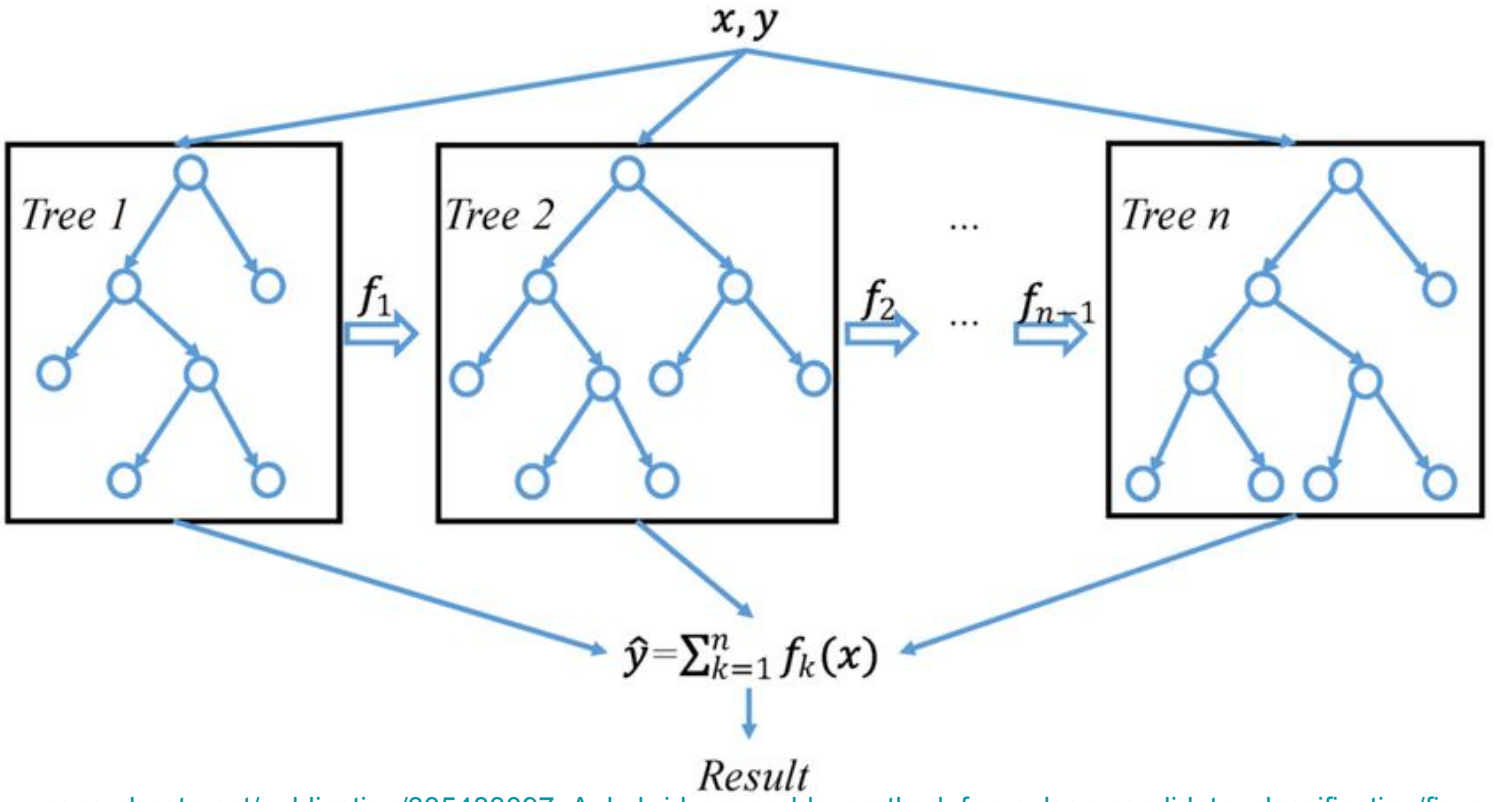
https://www.almabetter.com/bytes/tutorials/data-science/adaboost-algorithm

AdaBoost - Training Approach



https://www.almabetter.com/bytes/tutorials/data-science/adaboost-algorithm

Gradient Boosting - Sequential Ensemble Architecture



https://www.researchgate.net/publication/335483097_A_hybrid_ensemble_method_for_pulsar_candidate_classification/figures?lo=1

Gradient Boosting - Training Approach

Sequentially add weak learners to the model ensemble. Each new learner focuses on correcting the errors of the existing ones — it does this via gradient descent to minimize the specified loss function (fits the gradient of the loss function

“Gradient Boosting Example XLSX” is a representative approach (directly fits residuals)