BATCH : **B 84** Data Science

LESSON : **Machine Learning**

DATE : 07.10.2022

SUBJECT : **Supervised Learning**

AdaBoost-Gradient Boost - XGBoost

techproeducation

techproeducation

techproeducation

techproeducation

techproedu

MACHINE LEARNING - 5

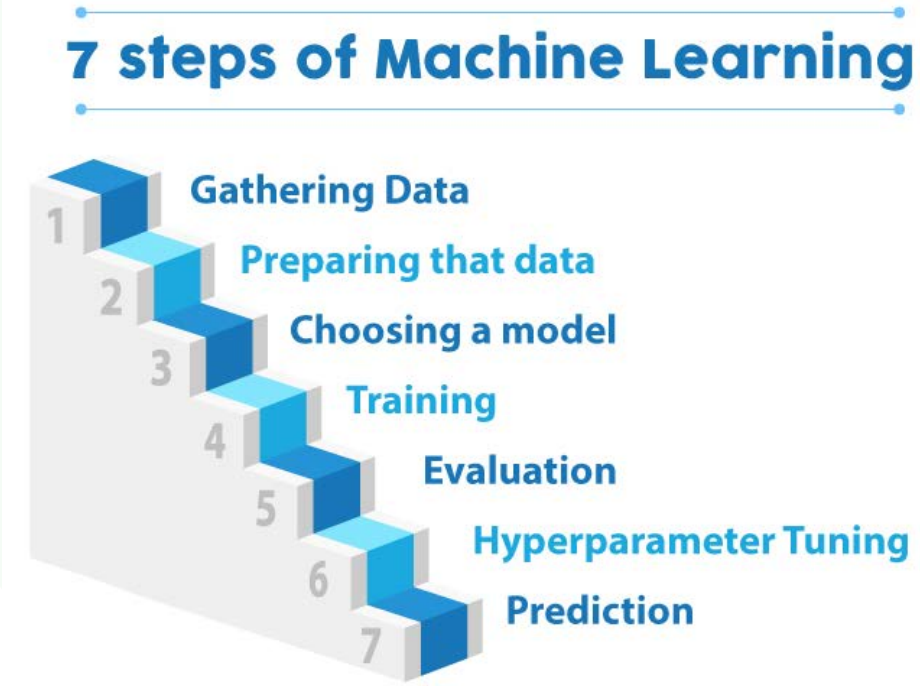
Machine Learning

Makine Öğrenmesi – 5

Overall Table of Contents

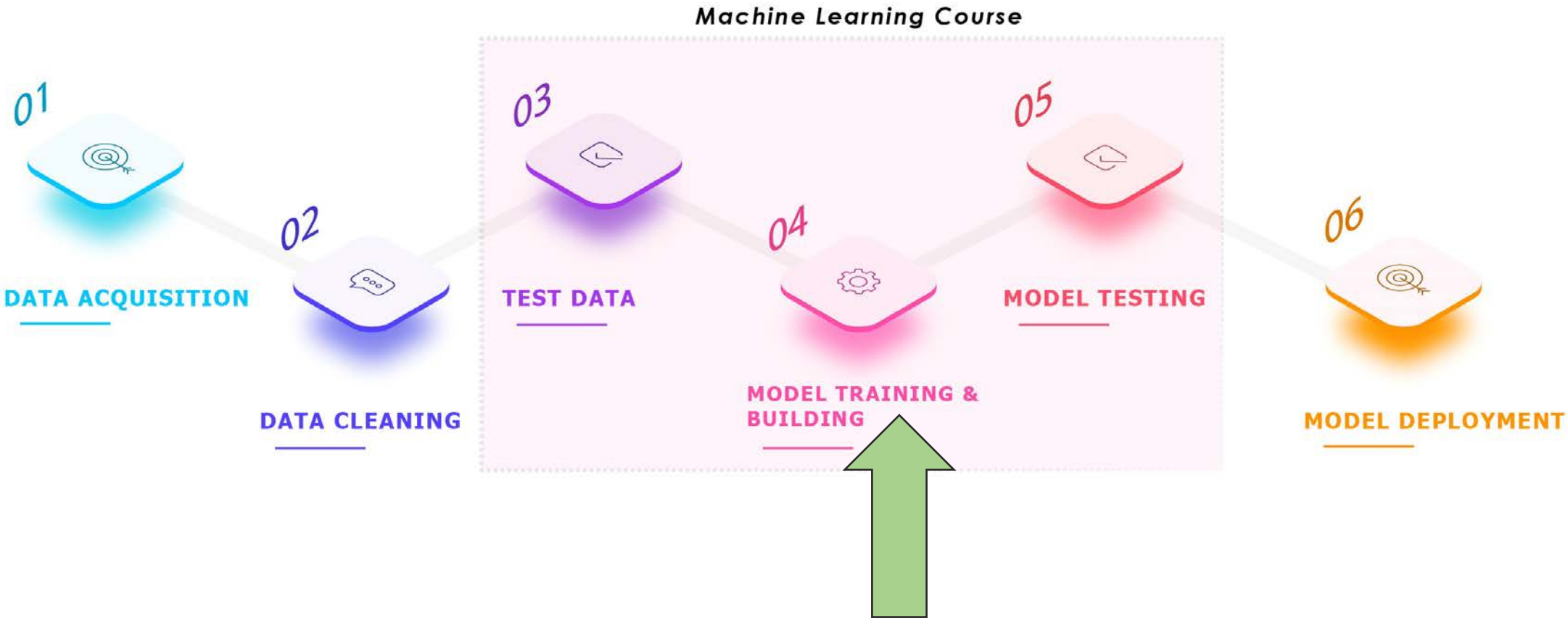
General Content

- Supervised Learnig Algorithm – **Bagging – Boosting Methods**
- Supervised Algorithm practices Python application
- Projects Solutions

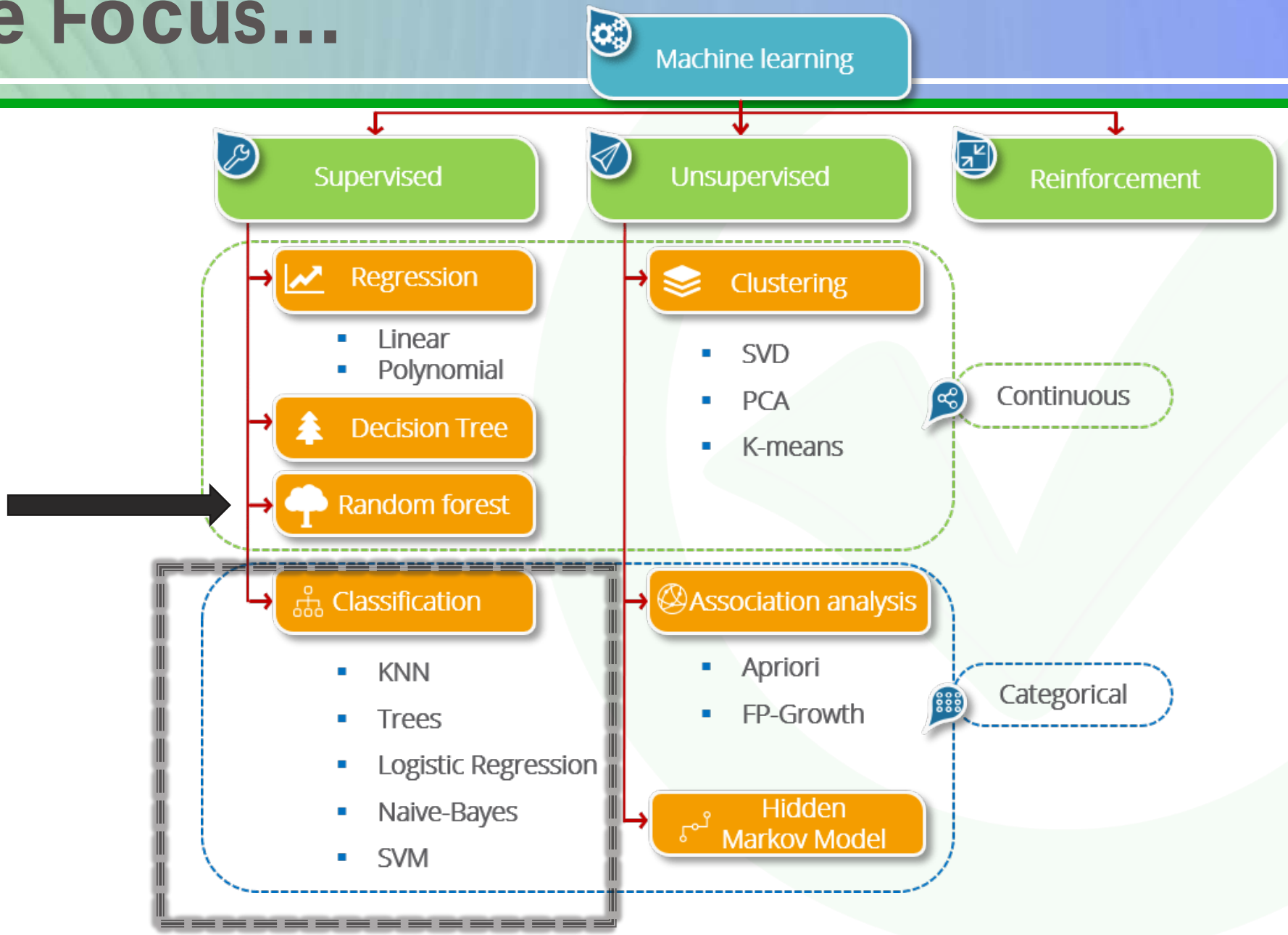


ENSEMBLE LEARNING

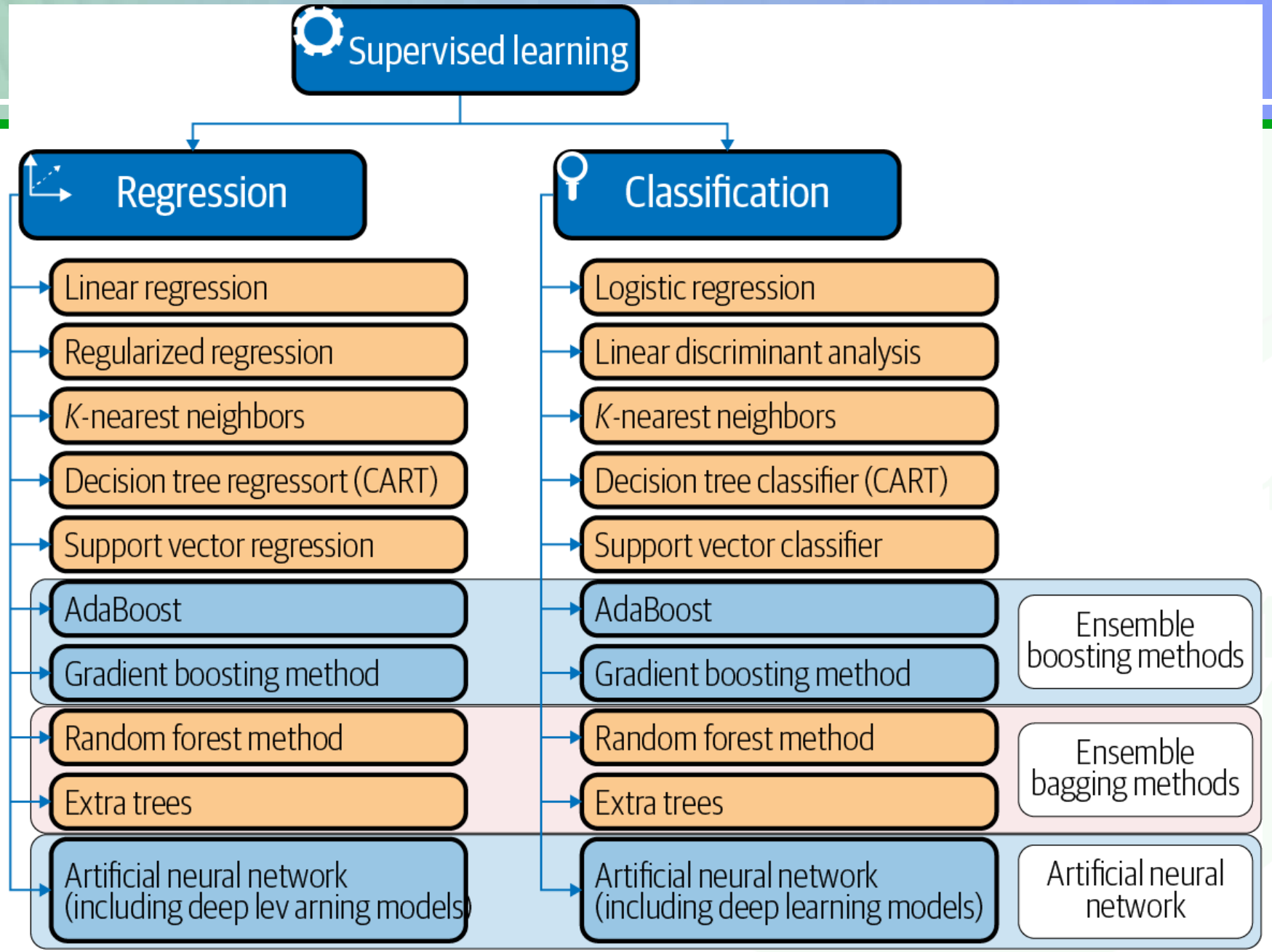
Where are we?


Machine Learning Course
01
DATA ACQUISITION
02
DATA CLEANING
03
TEST DATA
04
MODEL TRAINING & BUILDING
05
MODEL TESTING
06
MODEL DEPLOYMENT

More Focus...

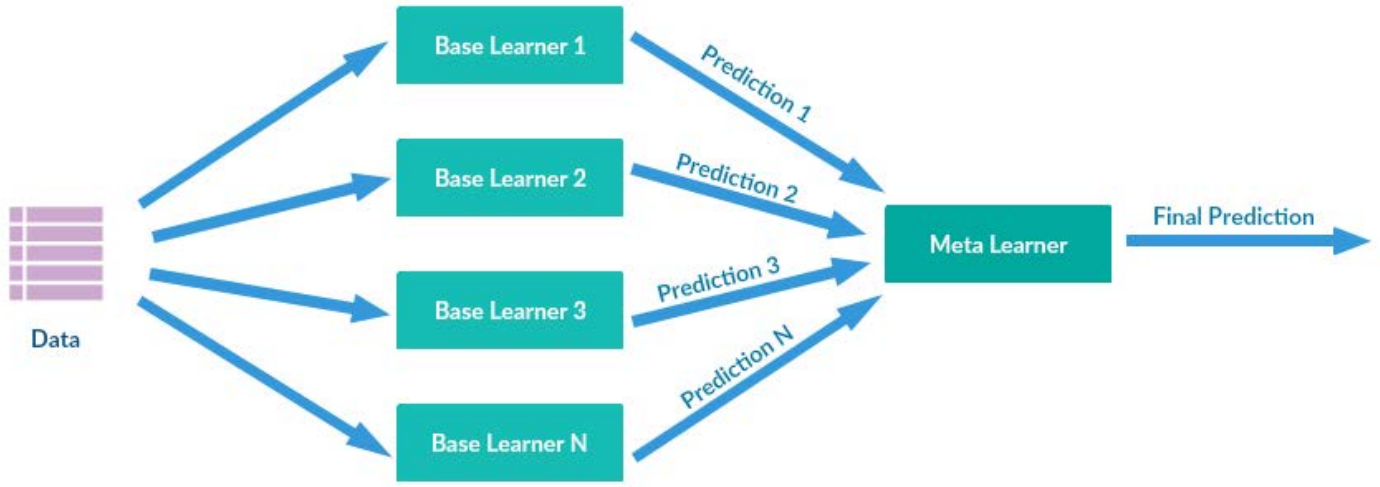

Machine learning
Supervised
Unsupervised
Reinforcement
Regression
Linear
Polynomial
Decision Tree
Random forest
Clustering
SVD
PCA
K-means
Continuous
Classification
KNN
Trees
Logistic Regression
Naive-Bayes
SVM
Association analysis
Apriori
FP-Growth
Categorical
Hidden Markov Model

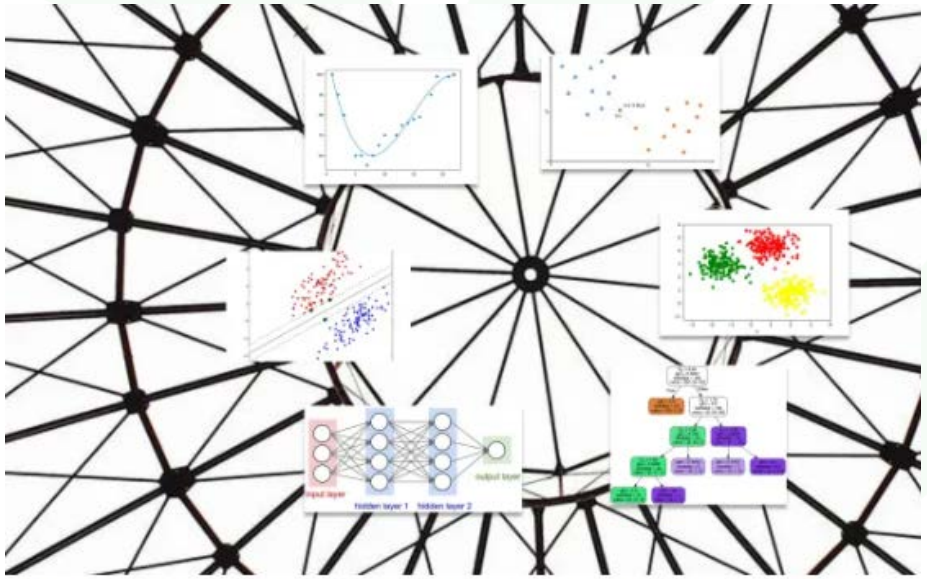
More..


Supervised learning
Regression
Classification
Linear regression
Regularized regression
K-nearest neighbors
Decision tree regressort (CART)
Support vector regression
AdaBoost
Gradient boosting method
Random forest method
Extra trees
Artificial neural network (including deep lev arning models)
Logistic regression
Linear discriminant analysis
K-nearest neighbors
Decision tree classifier (CART)
Support vector classifier
AdaBoost
Gradient boosting method
Random forest method
Extra trees
Artificial neural network (including deep learning models)
Ensemble boosting methods
Ensemble bagging methods
Artificial neural network

ENSEMBLE LEARNING METHODS (Kolektif Öğrenme Metotları)

- Farklı makine öğrenmesi algoritmalarını daha iyi bir tahmin için birlikte kullanmak mümkün mü?

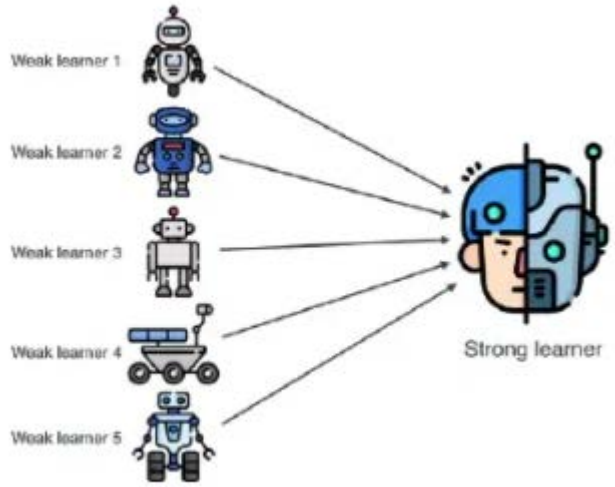




Ensemble methods combine several machine learning models to improve results

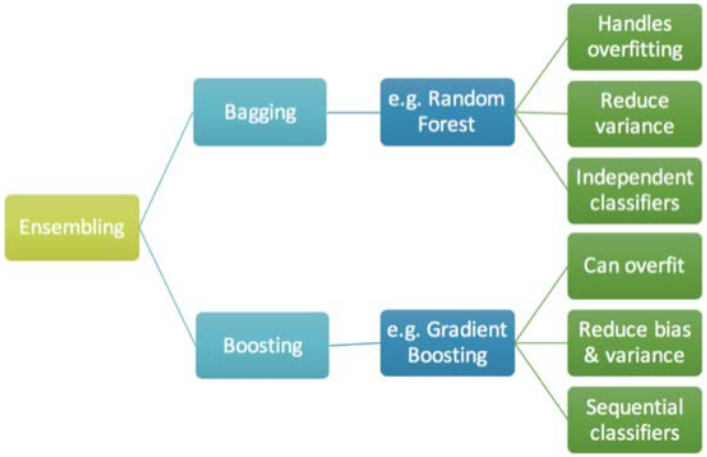
Supervised Learning

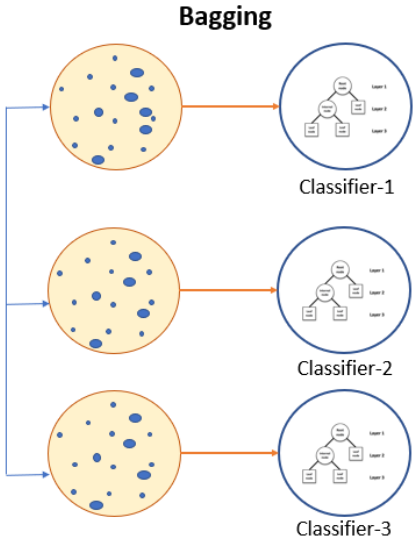
Özetle…

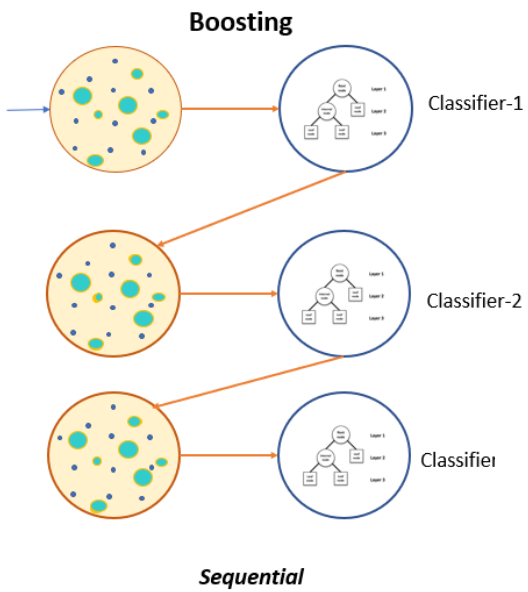


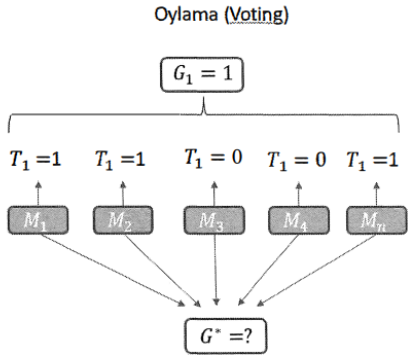
Supervised L. - ENSEMBLE LEARNING METHODS

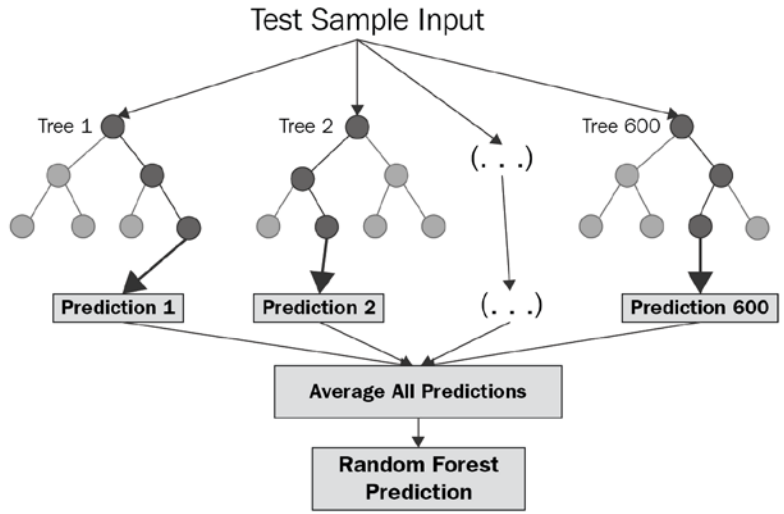
- ✓ Bagging (Bootstrap Aggregating)
- ✓ Boosting

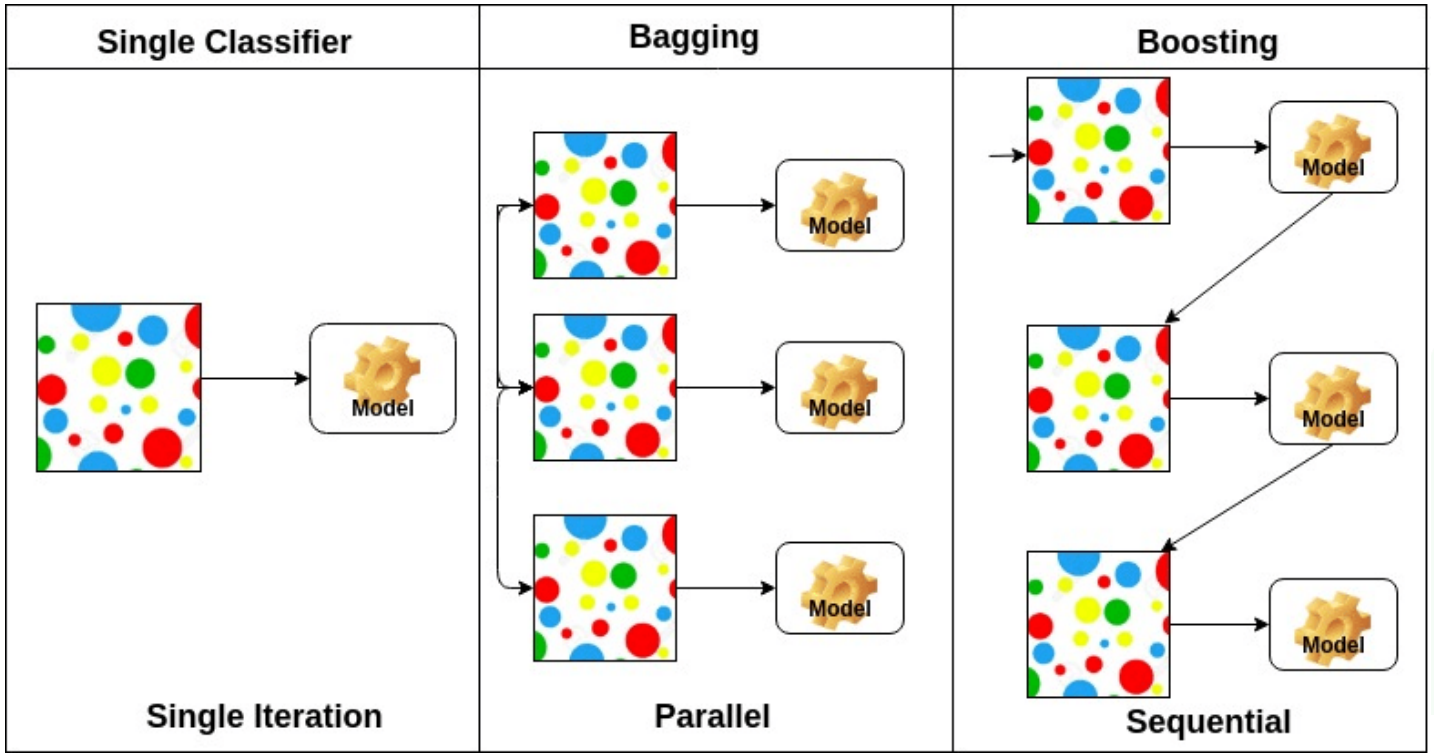


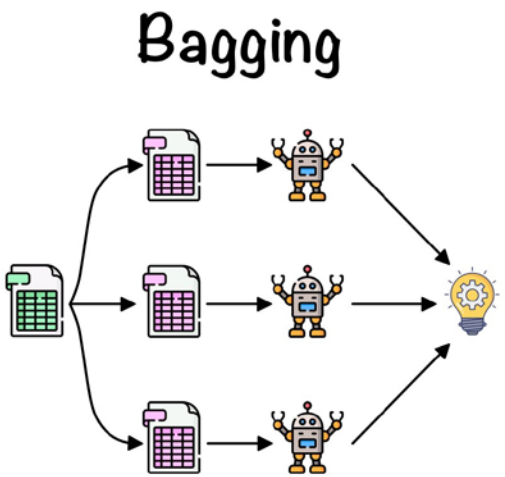


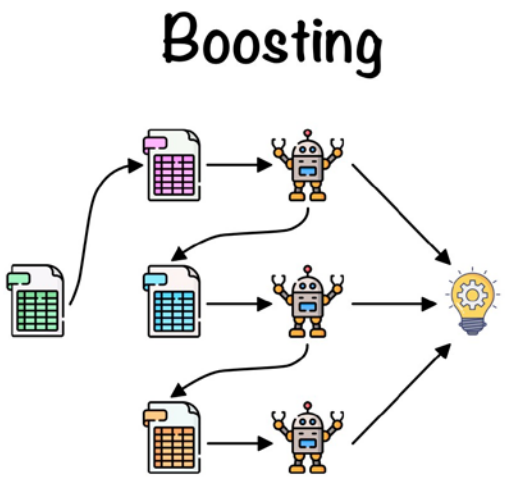






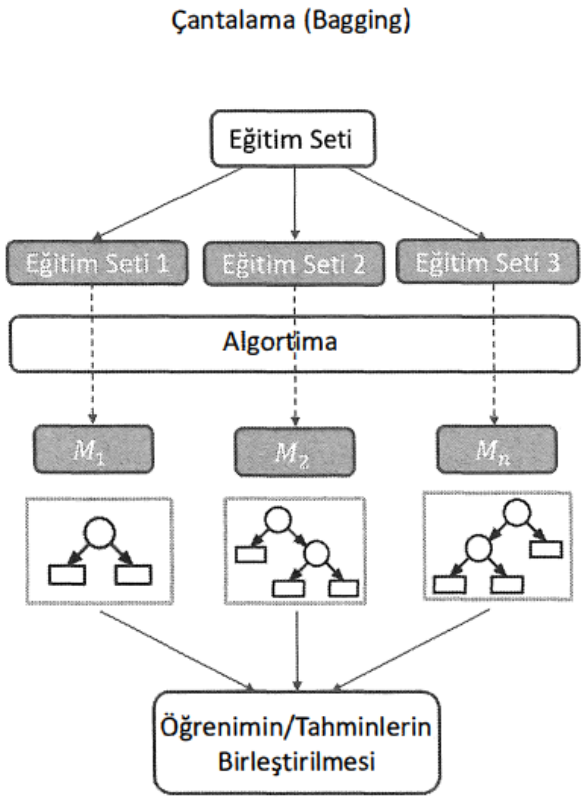

Single Classifier
Bagging
Boosting
Model
Model
Model
Model
Model
Model
Model
Single Iteration
Parallel
Sequential

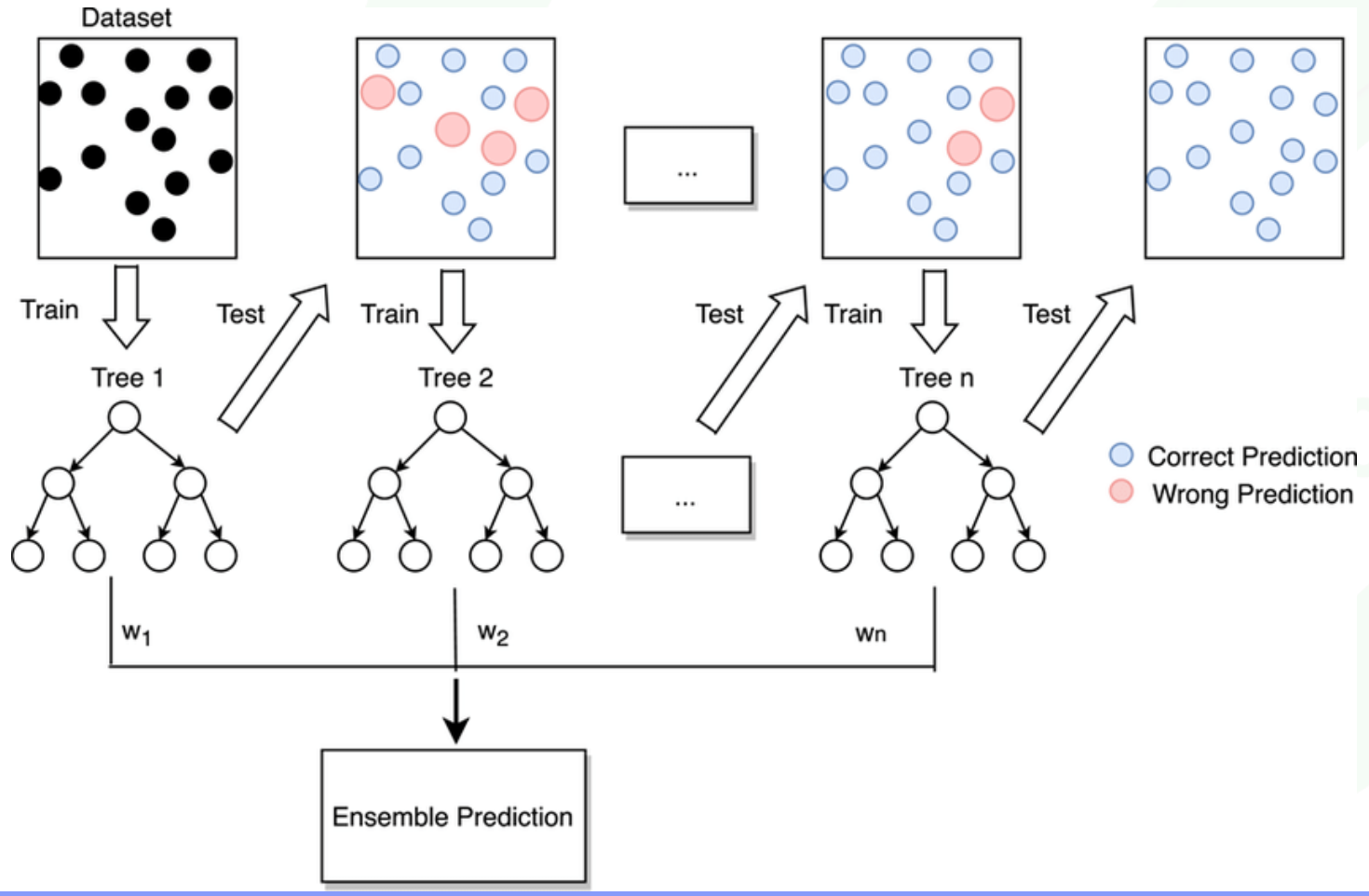

Bagging
Parallel


Boosting
Sequential

Supervised L. - ENSEMBLE LEARNING METHODS

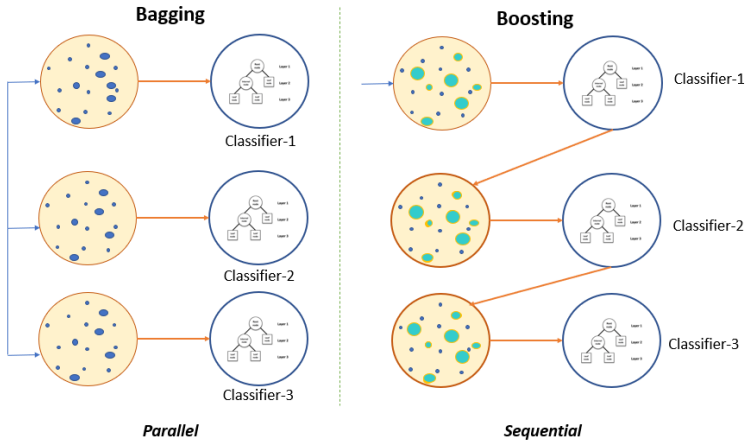
✓ Bagging (Random Forest-RF)

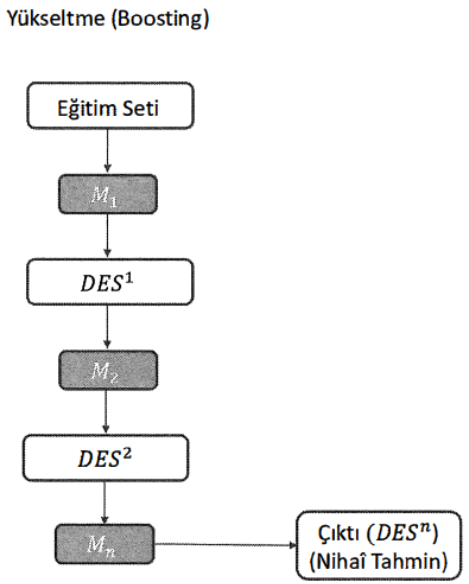


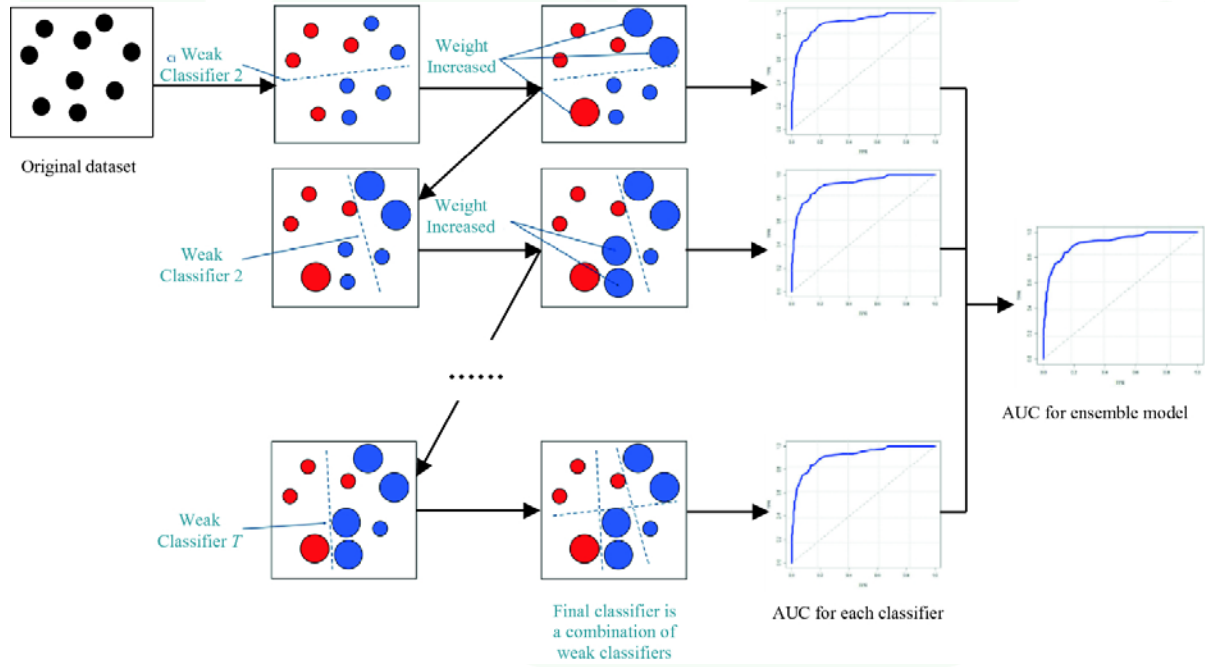


Supervised L. - ENSEMBLE LEARNING METHODS

- ✓ Bagging
- ✓ Boosting (Ada Boost-Gradient Boost-XGBoost)

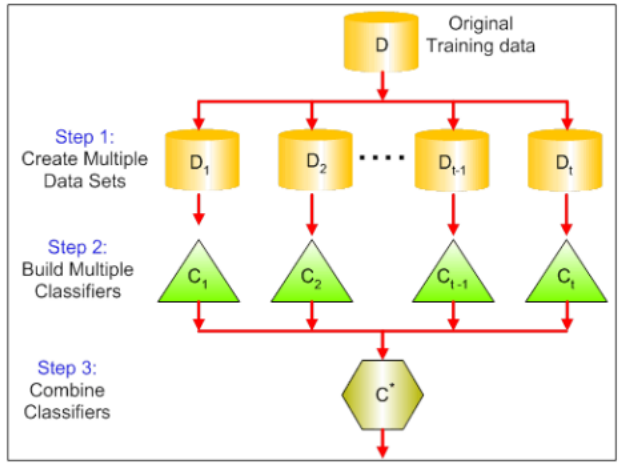






Bagging

Building multiple models (typically of the same type) from different subsamples of the training dataset.



	Bagging	Boosting
Similarities	• Uses voting • Combines models of the same type	
Differences	Individual models are built separately	Each new model is influenced by the performance of those built previously
	Equal weight is given to all models	Weights a model's contribution by its performance

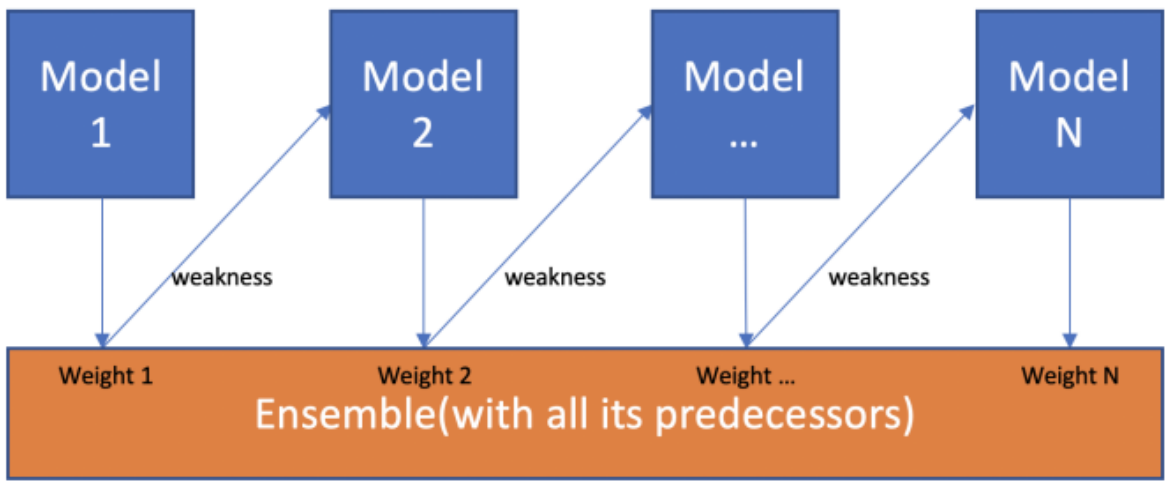
⇨ BOOSTING -->AdaBoost

⇨ BAGGING --> R F

Boosting

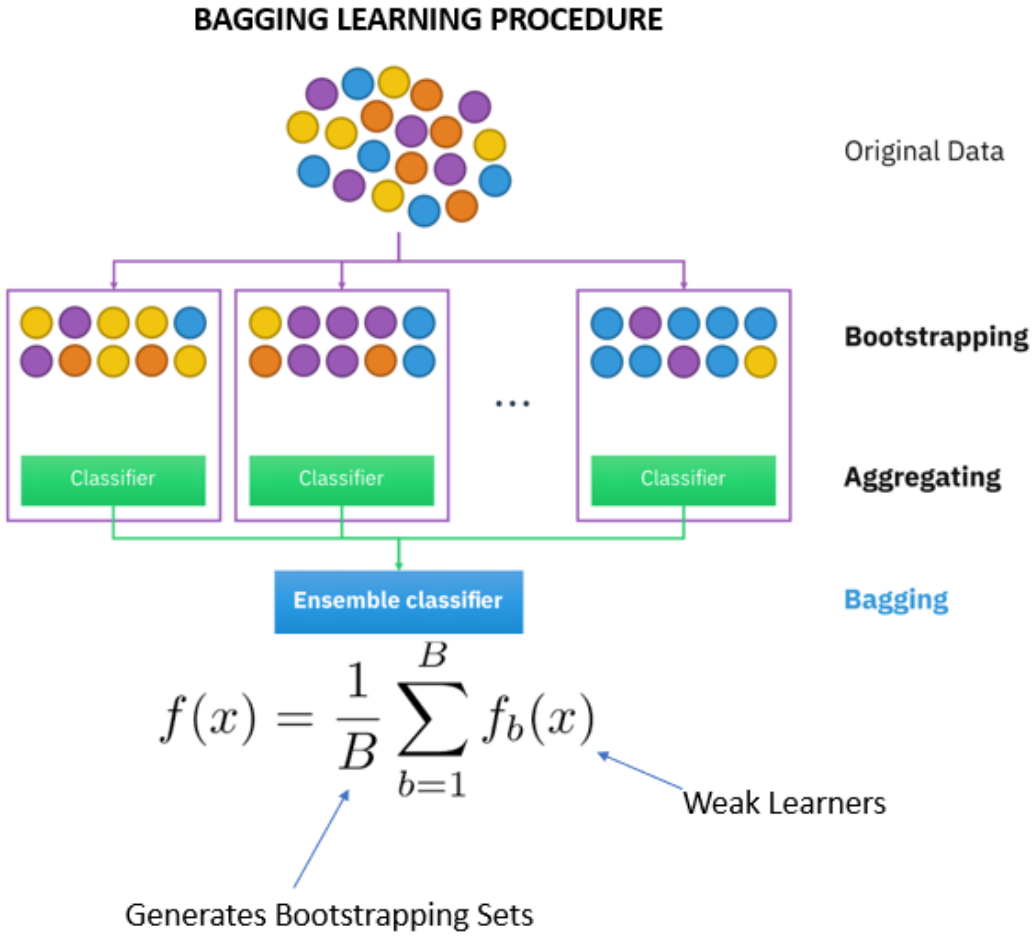
Building multiple models (typically of the same type) each of which learns to fix the prediction errors of a prior model in the chain.

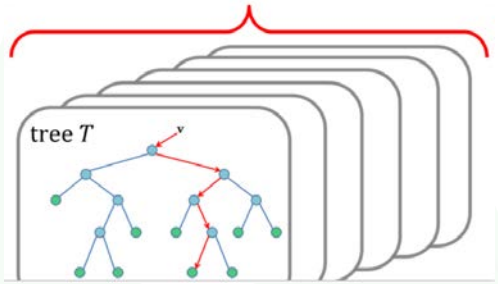
Model 1,2,..., N are individual models (e.g. decision tree)

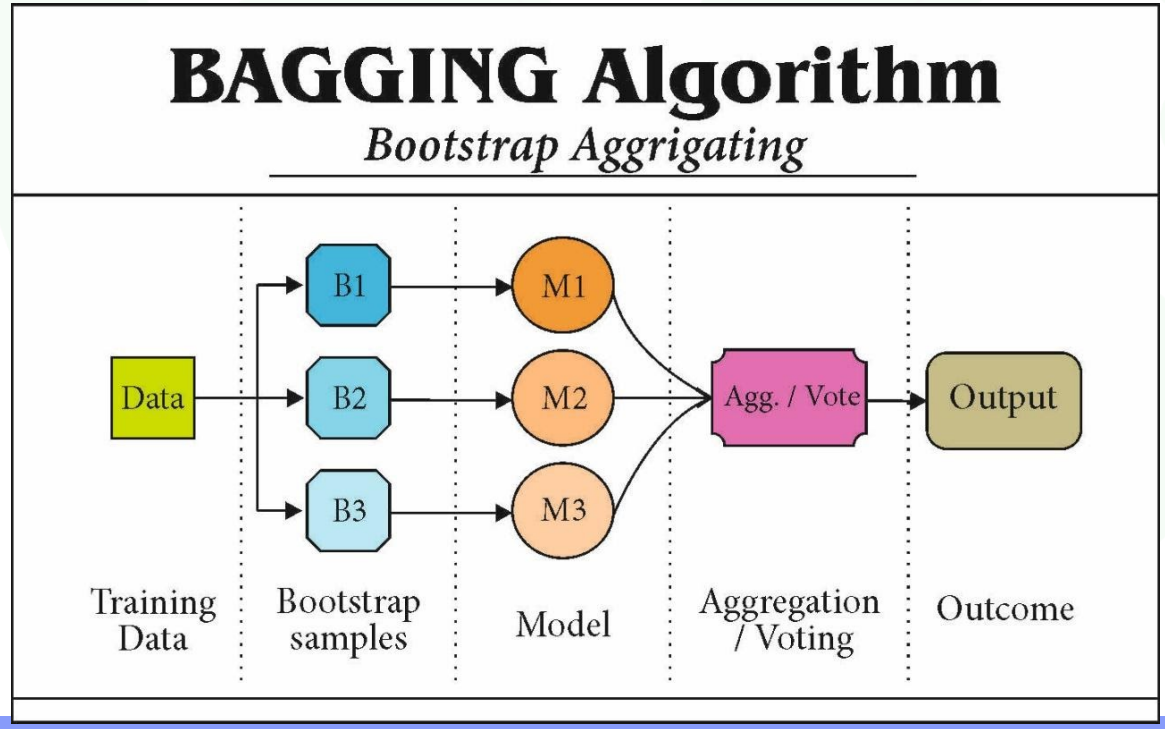


Supervised L. - ENSEMBLE LEARNING METHODS

✓ Bagging (Bootstrap Aggregating)

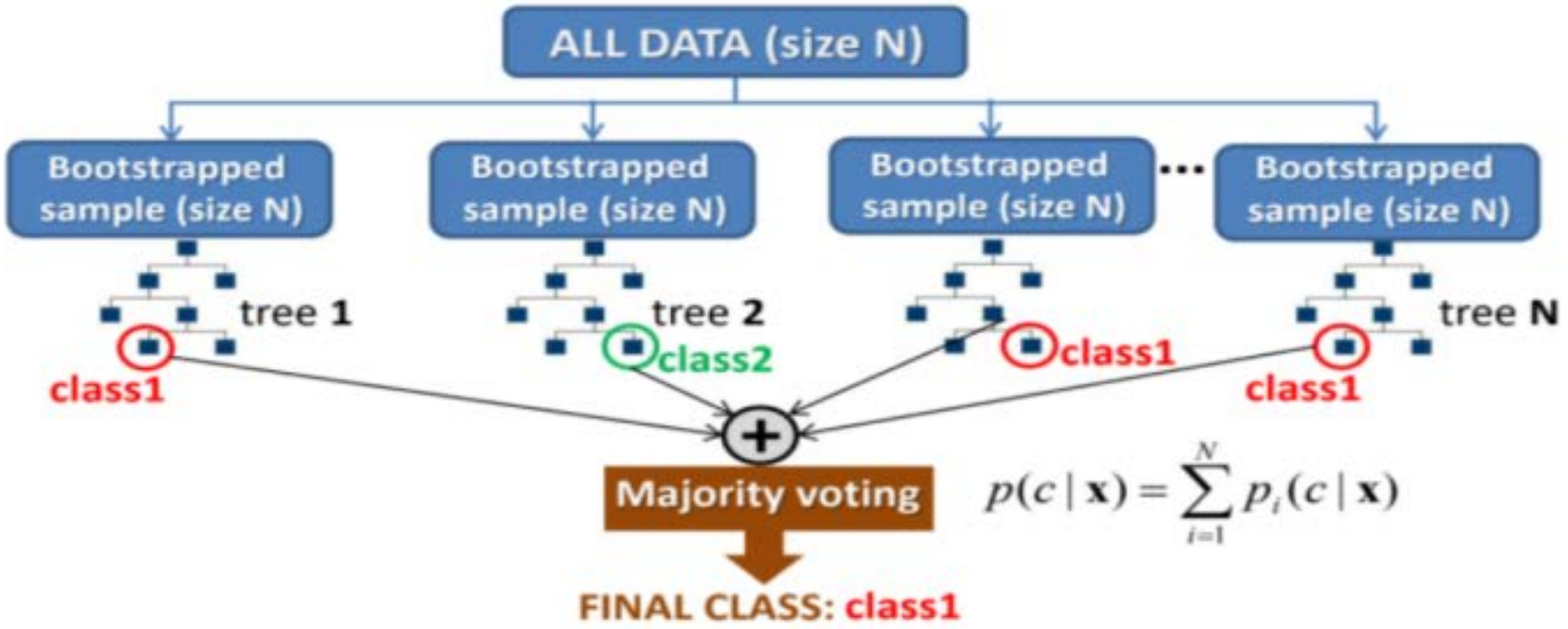






Supervised L. - ENSEMBLE LEARNING METHODS

✓ Random Forest (RF)



Hyperparameters:

"n_estimators" *parameter:* (default=100)

The number of trees in the forest.

The more # of trees, the better accuracy. But CPU intensive.

"max_depth" *parameter:* (default=None)

The maximum depth of the tree.

If None, then nodes are expanded until all leaves are pure.

"max_features" *parameter:*

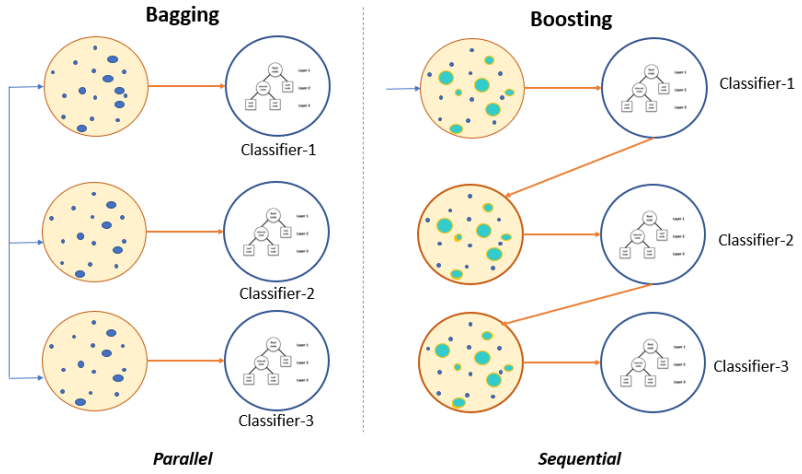
Number of features to consider when looking for the best split.

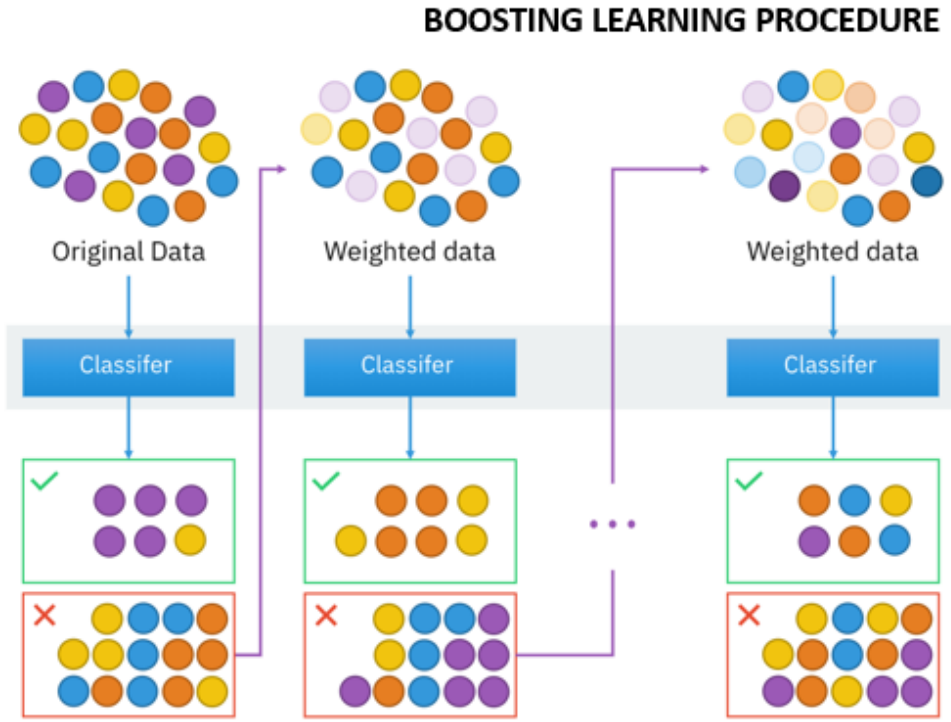
Increase will improve the performance but results in a correlation between the trees.

Avantaj - Dezavantajları

Supervised L. - ENSEMBLE LEARNING METHODS

✓ Boosting





Strong Learner — Weak Learners

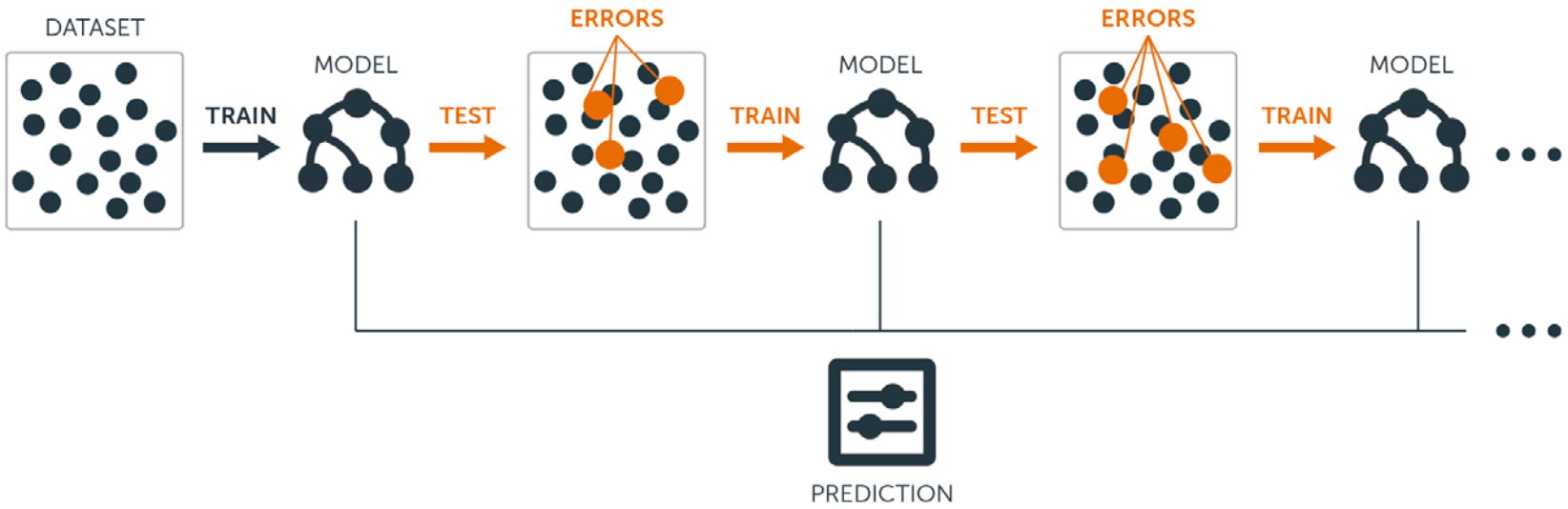
$$f(x) = \sum_t \alpha_t h_t(x)$$

Weight calculated by considering the last iteration's error

MasterChef

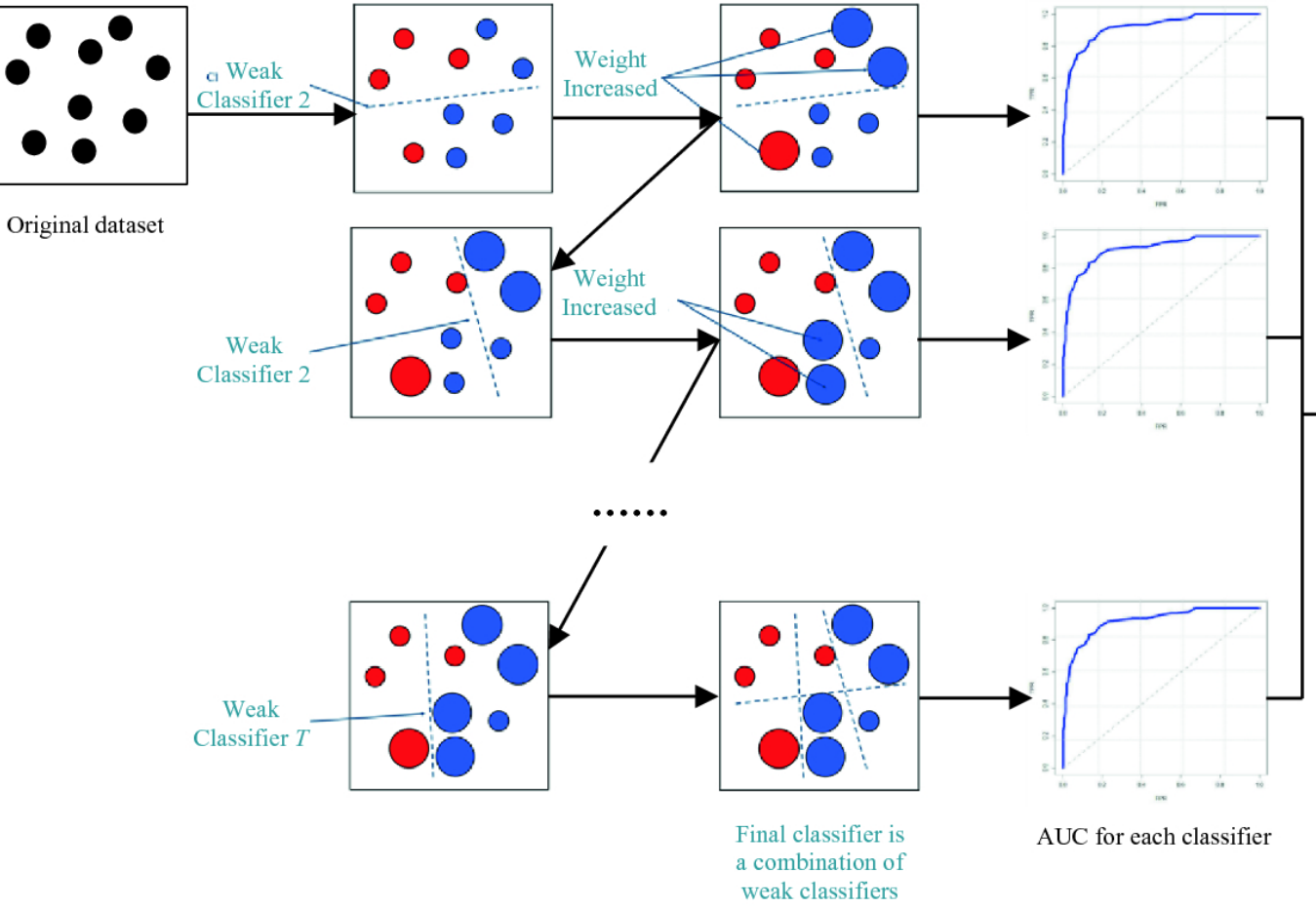
Supervised L. - ENSEMBLE LEARNING METHODS

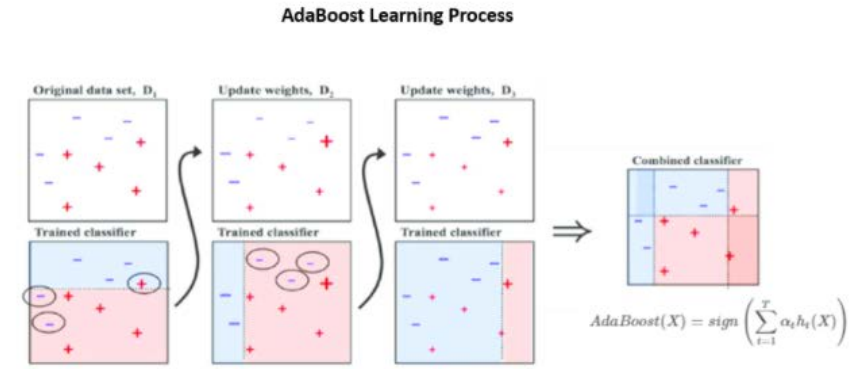
✓ Boosting (Yükseltme)



Supervised L. - ENSEMBLE LEARNING METHODS

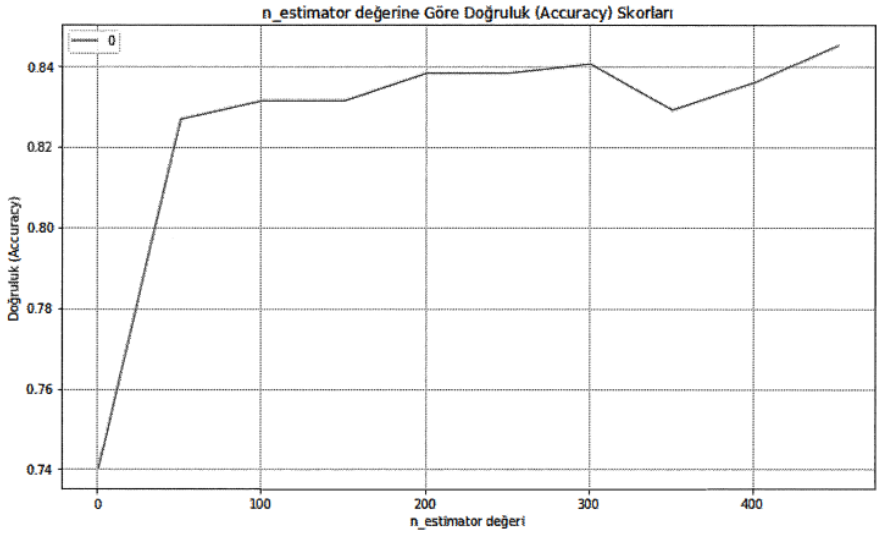
✓ Ada Boosting (Adaptive Boosting)

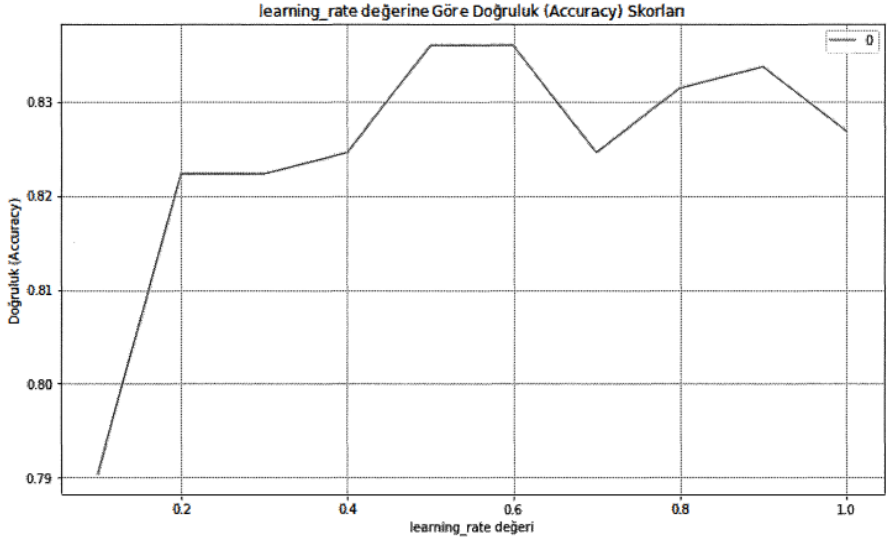




Supervised L. - ENSEMBLE LEARNING METHODS

✓ Ada Boosting (Adaptive Boosting)






```
188 # Varyans sorunu var mı?
189 from sklearn import metrics
190 # Modelin eğtim seti  için doğruluk oranı:
191 print('Accuracy:%0.3f'%metrics.accuracy_score(y_train, y_pred_train))
192 # Modelin test seti için doğruluk oranı:
193 print('Accuracy:%0.3f'%metrics.accuracy_score(y_test, y_pred_test))
194
```

```
Accuracy:0.863
Accuracy:0.827
```

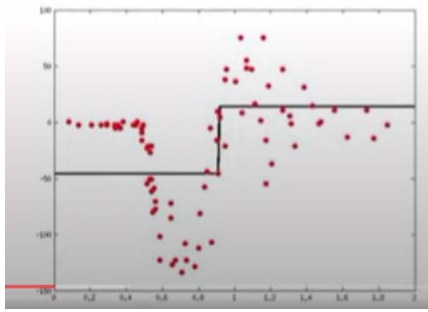
Supervised L. - ENSEMBLE LEARNING METHODS

✓ Gradient Boosting

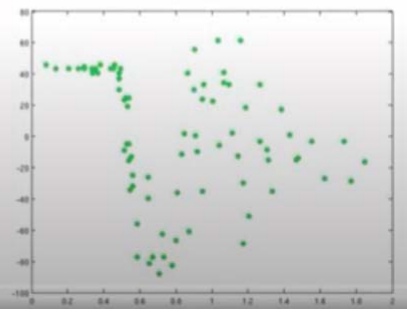
Gradient Boosting

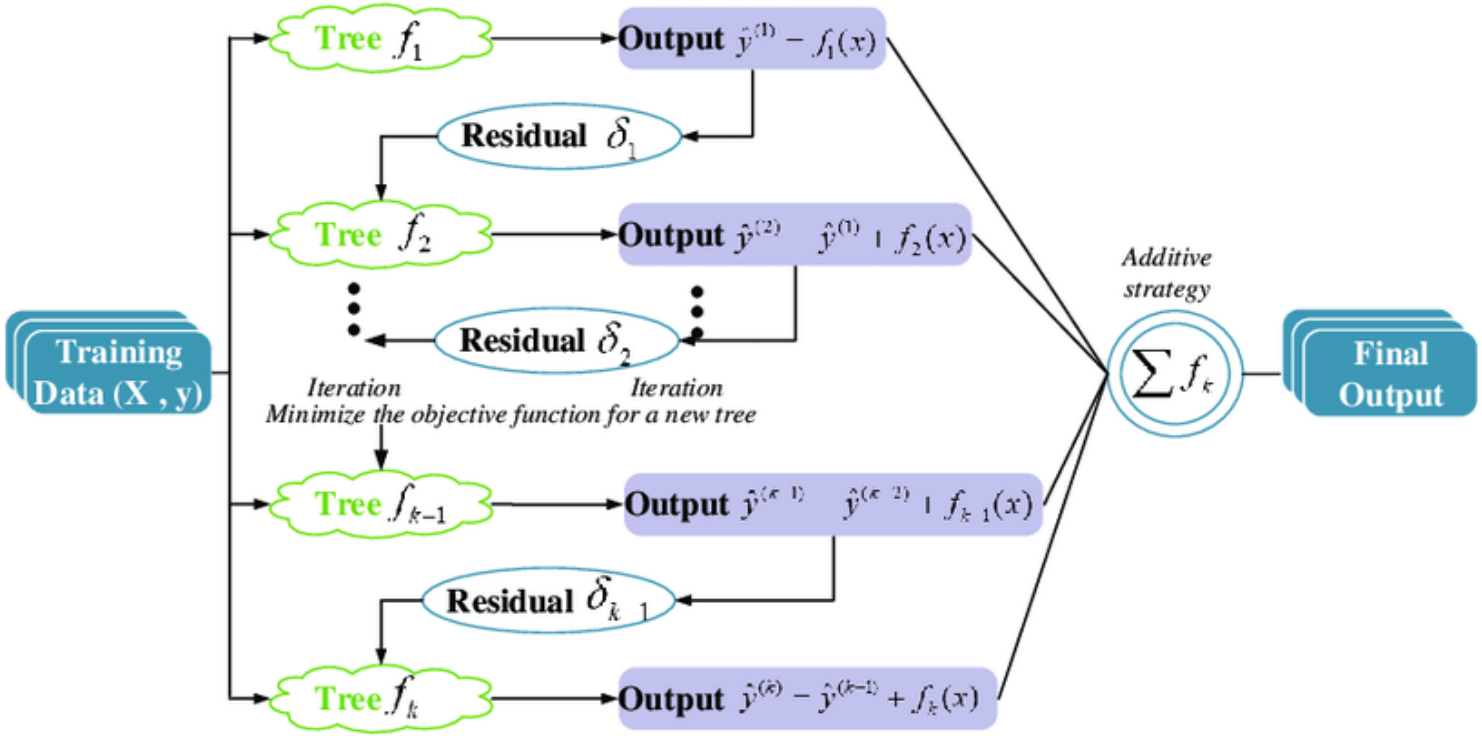
- Learn a regression predictor
- Compute the error residual
- Learn to predict the residual

Learn a simple predictor…



Then try to correct its errors



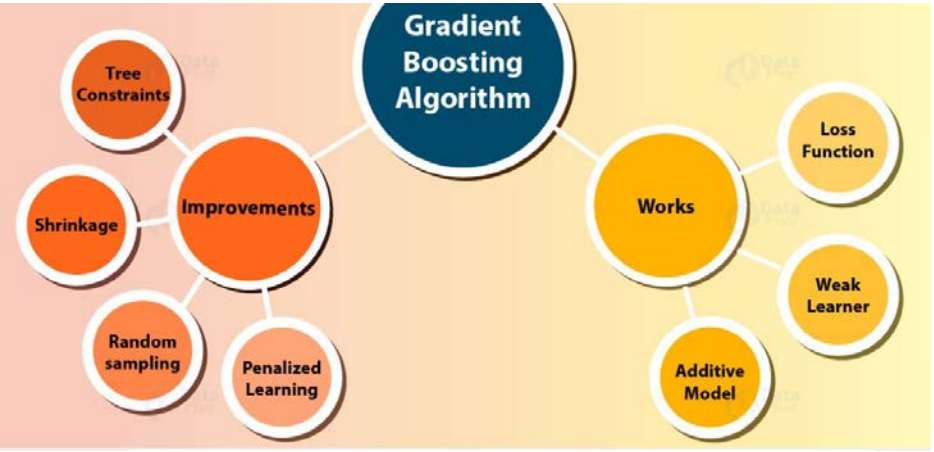


Average price: 688 + Learning Rate = 0.1 × Residual predicted by decision tree: -338

https://youtu.be/3zEqUSf5duw

Supervised L. - ENSEMBLE LEARNING METHODS

✓ Gradient Boosting



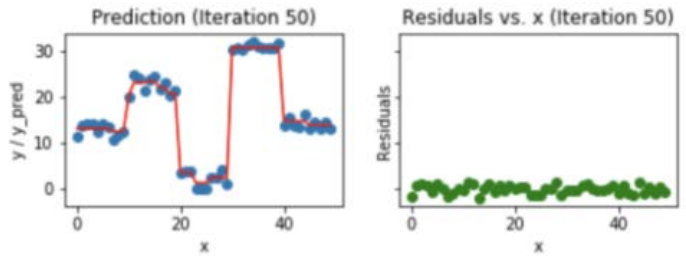


Fig 7. Visualization of gradient boosting prediction (iteration 50th)

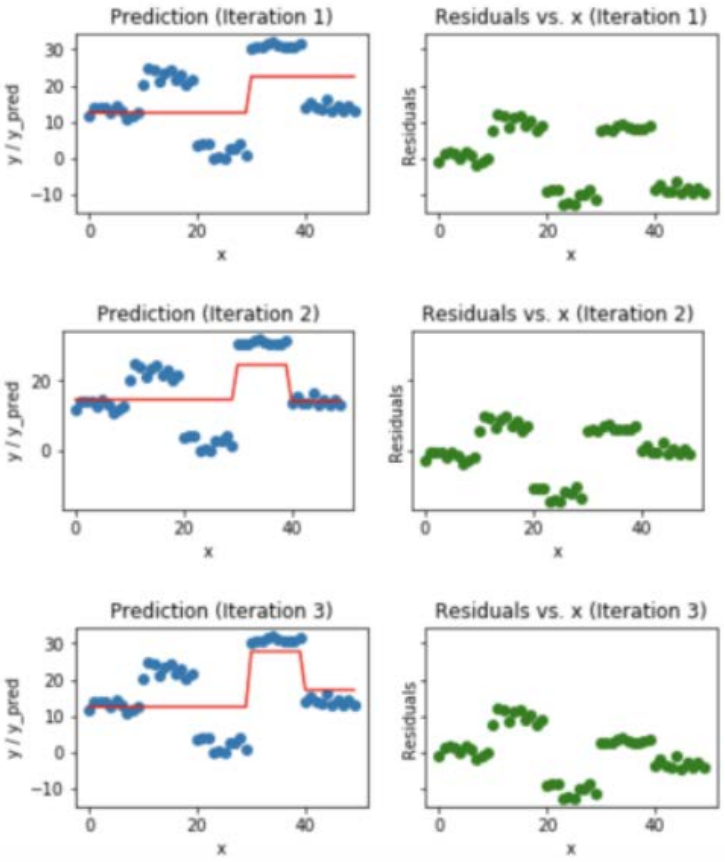


Fig 5. Visualization of gradient boosting predictions (First 4 iterations)

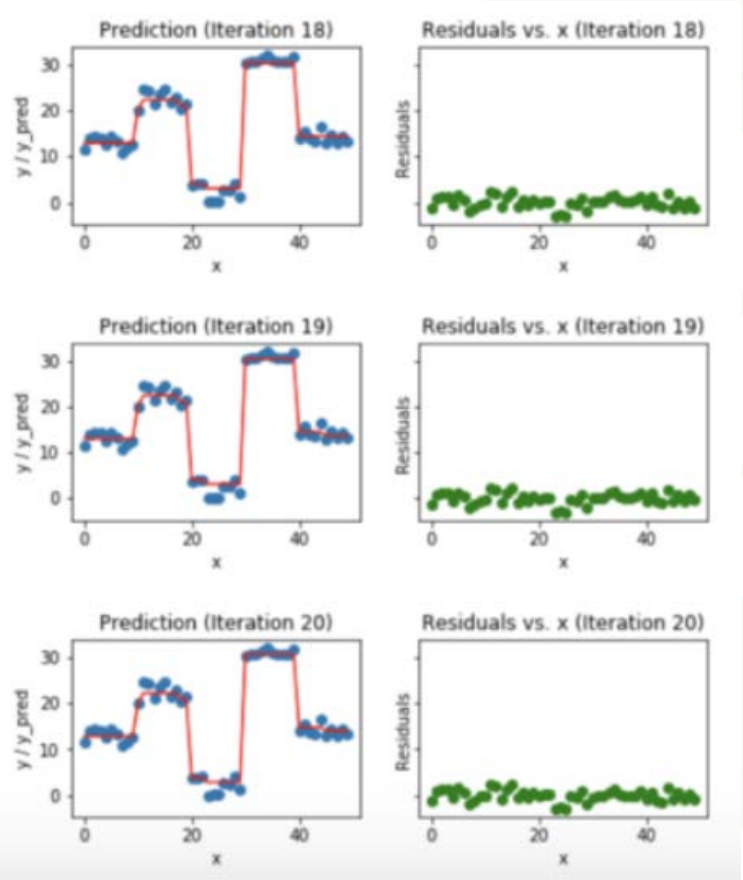


Fig 6. Visualization of gradient boosting predictions (18th to 20th iterations)

https://blog.mlreview.com/gradient-boosting-from-scratch-1e317ae4587d

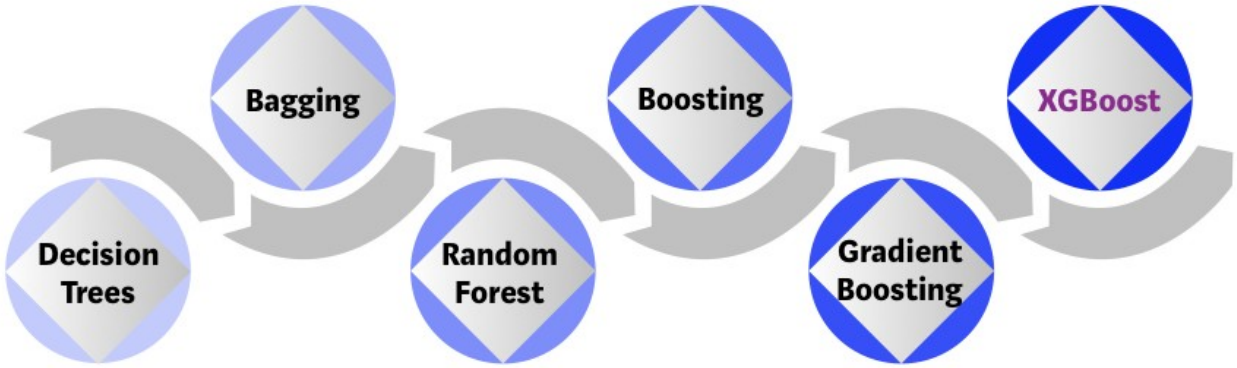
Supervised L. - ENSEMBLE LEARNING METHODS

XG Boost (EXtreme Gradient Boosting)

Bootstrap aggregating or Bagging is a ensemble meta-algorithm combining predictions from multiple-decision trees through a majority voting mechanism

Models are built sequentially by minimizing the errors from previous models while increasing (or boosting) influence of high-performing models

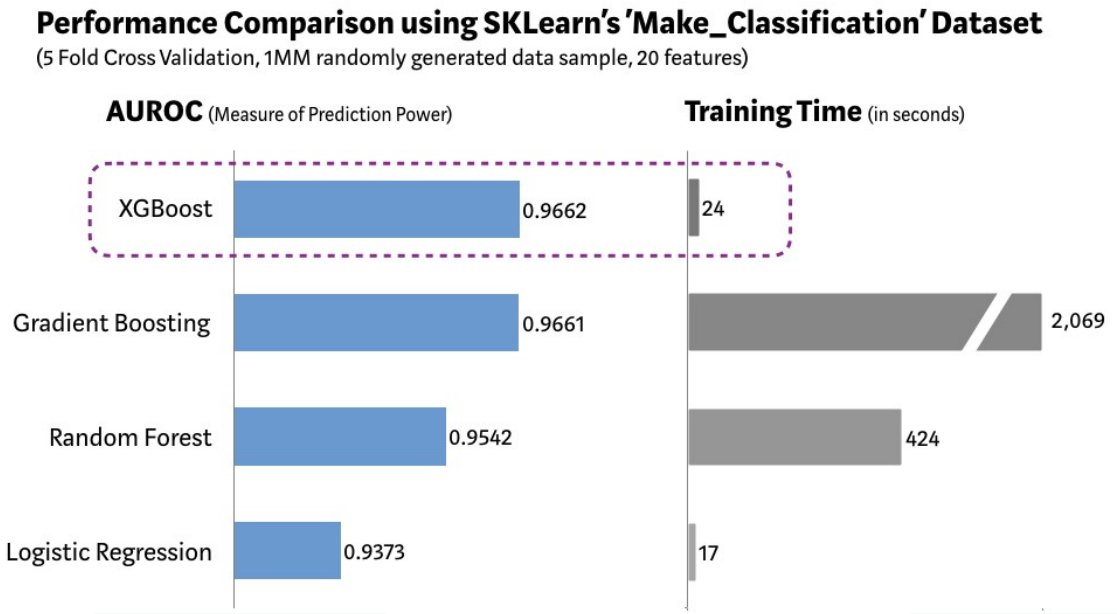
Optimized Gradient Boosting algorithm through parallel processing, tree-pruning, handling missing values and regularization to avoid overfitting/bias



A graphical representation of possible solutions to a decision based on certain conditions

Bagging-based algorithm where only a subset of features are selected at random to build a forest or collection of decision trees

Gradient Boosting employs gradient descent algorithm to minimize errors in sequential models



"n_estimators" *parameter:* (default=100)

The number of boosting stages to perform.

"max_depth" *parameter:* (default=6)

*Max. depth of the individual estimators. This parameter **limits the number of nodes** in the tree.*

"subsample" *parameter:* (default=1.0)

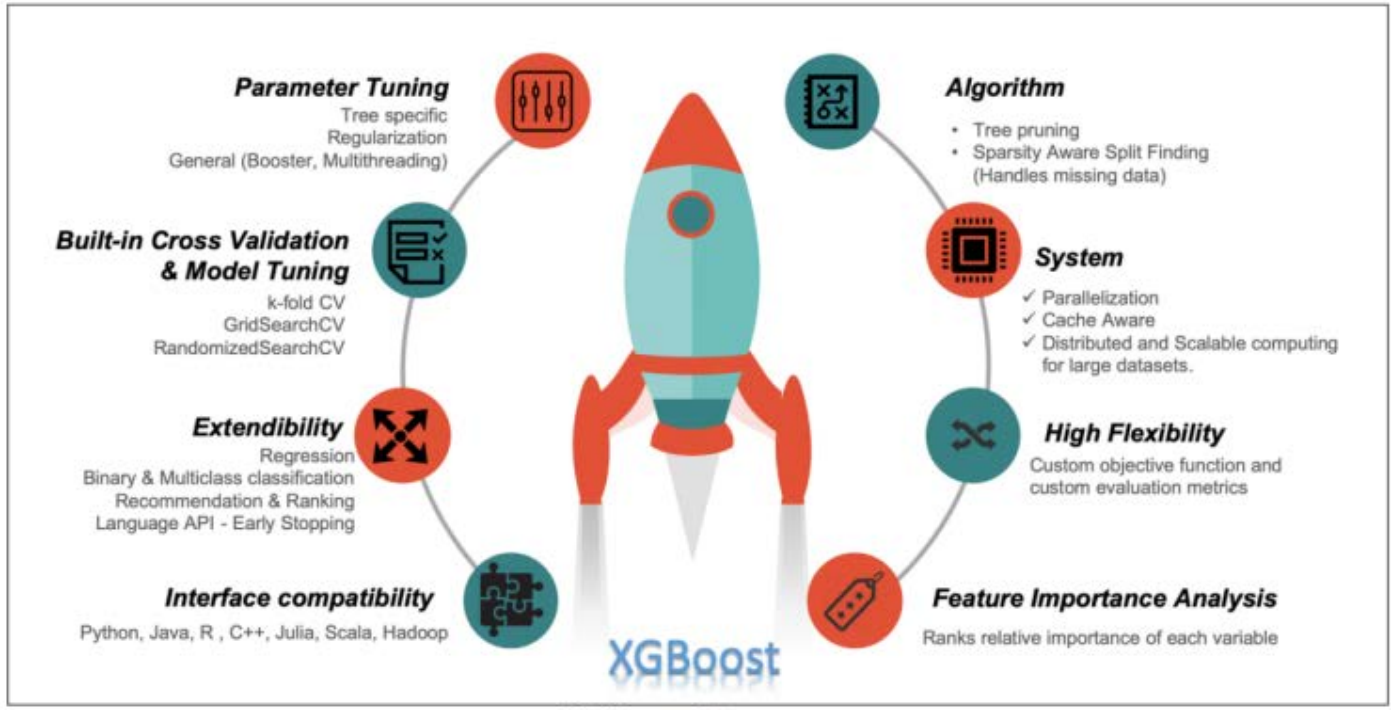
The fraction of samples to be used for fitting the individual base learners. Choosing subsample < 1.0 leads to a reduction of variance and an increase in bias.

"learning_rate" *parameter:* (default=0.3)

Learning rate shrinks the contribution of each tree by learning_rate. There is a trade-off between learning_rate and n_estimator

Supervised L. - ENSEMBLE LEARNING METHODS

- ✓ **XG Boost (EXtreme Gradient Boosting)**
- ✓ **XGB neden diğerlerinden daha iyi?**
- ✓ ***Regularization***
- ✓ ***Parallel Processing***
- ✓ ***High Flexibility***
- ✓ ***Handling Missing Values***
- ✓ ***Tree Pruning***
- ✓ ***Built-in Cross-Validation***



XGBoost Features

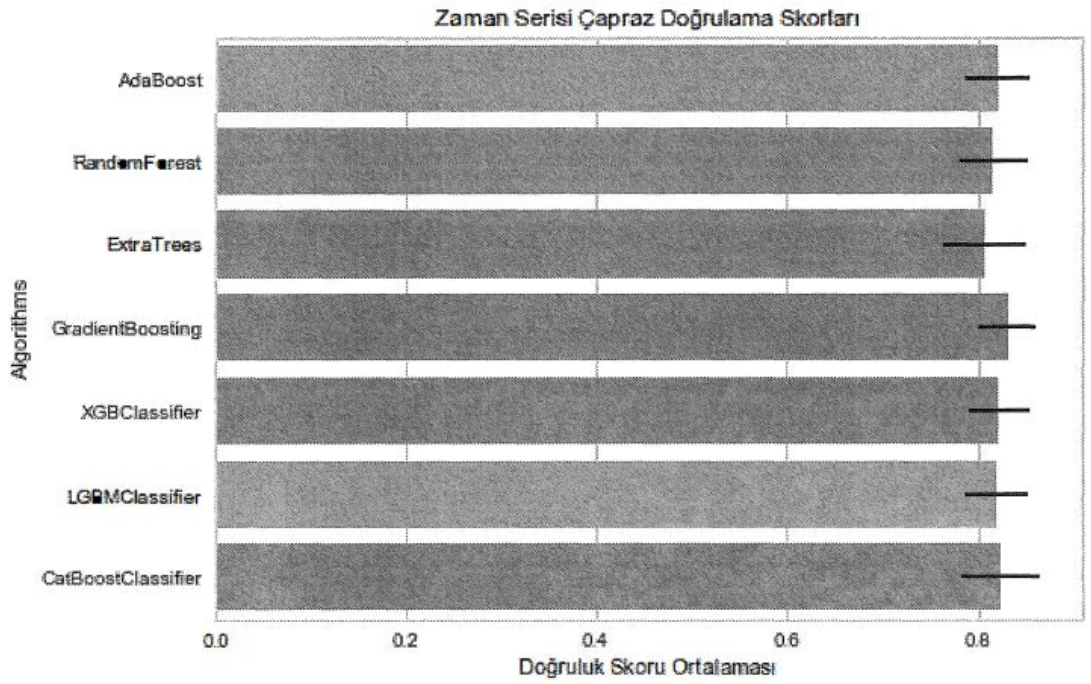
✓ ENSEMBLE KOLEKTİF ÖGRENME YÖNTEMLERİ ARASINDAN MODEL SEÇİMİ

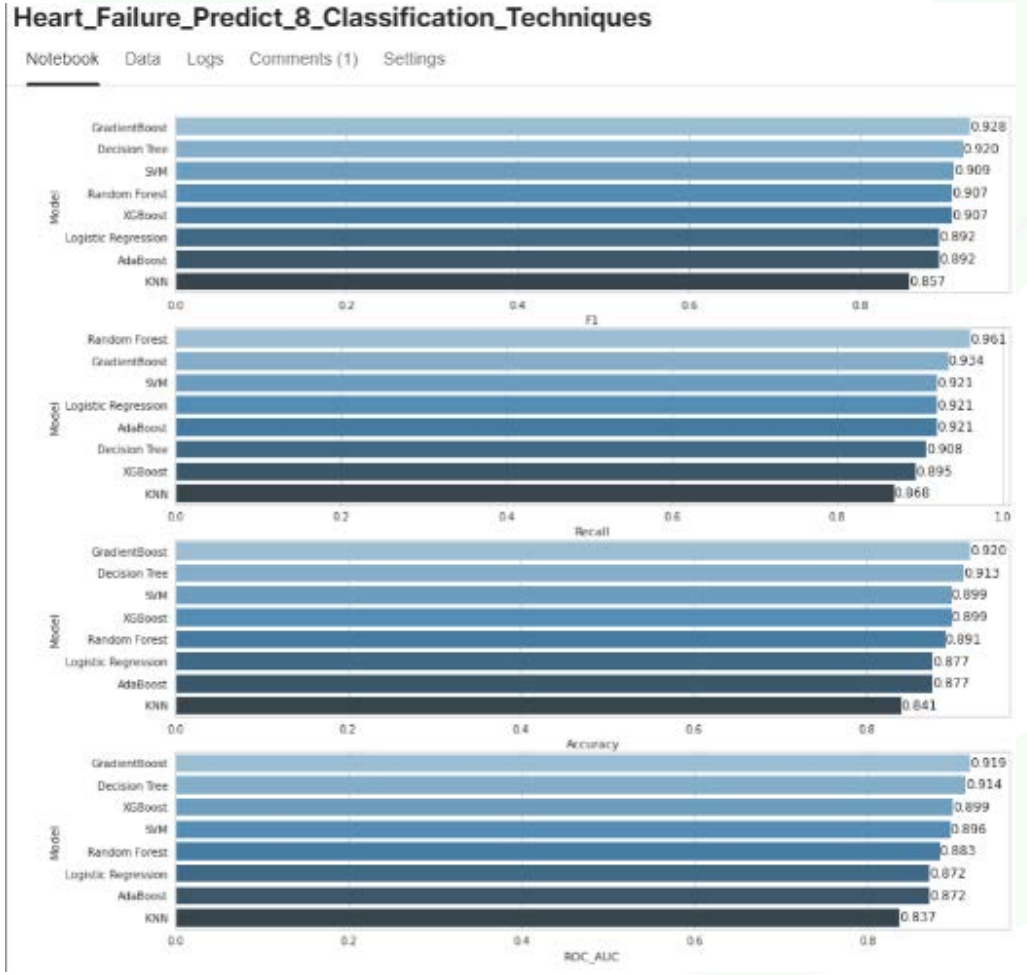
✓ Zaman-performans ikilemi

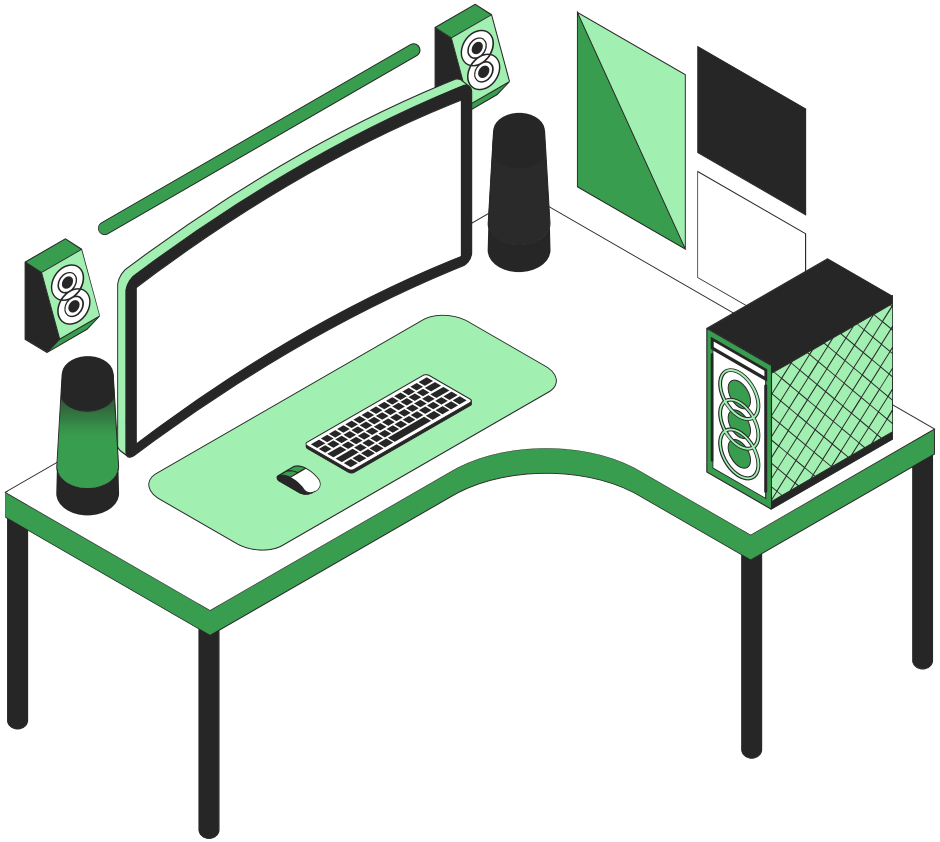
- ✓ **i)** ortaya çıkacak performans bir tek algoritmanın ürettiği modelinkinden daha yüksek olmalı;
- ✓ **ii)** işlem süresi algoritmaların tek tek uygulanmasından daha kısa sürmeli.

Supervised L. - ENSEMBLE LEARNING METHODS

✓ XG BOOST (eXtram Gradient Boosting)







Do you have any questions?

Send it to us! We hope you learned something new.

TECHPROED