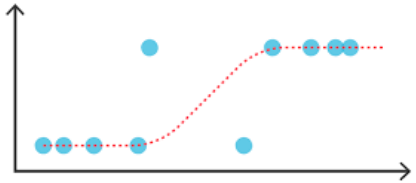
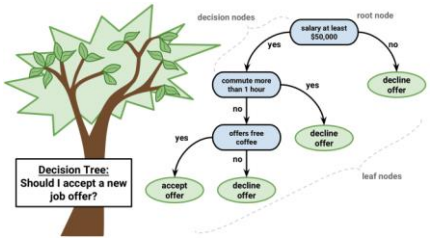
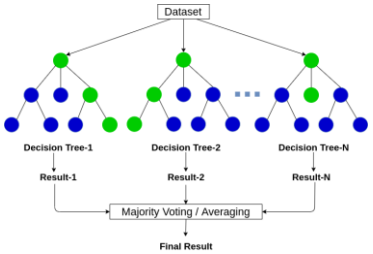
Uygulama -1

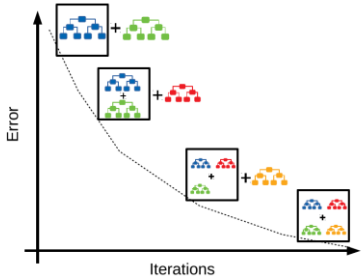
Algoritmalar

- **Lojistik Regresyon**
- Karar Ağacı
- Random Forest
- Gradient Boosting Tree





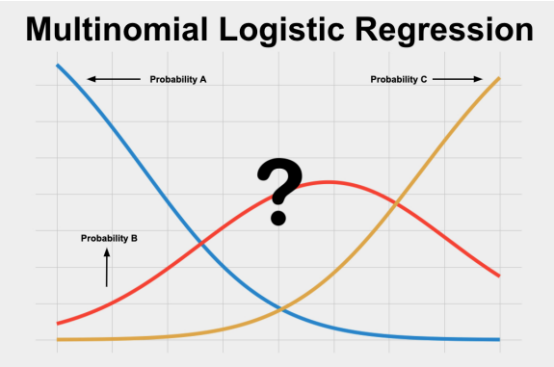


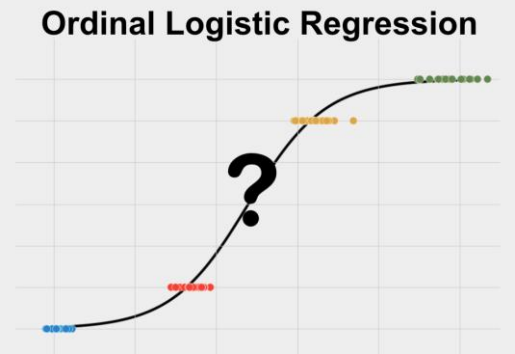


https://www.tibco.com/reference-center/what-is-logistic-regression
https://mychartguide.com/decision-tree/
https://www.tibco.com/reference-center/what-is-a-random-forest
https://morioh.com/p/e108a4521555

Lojistik Resgresyon

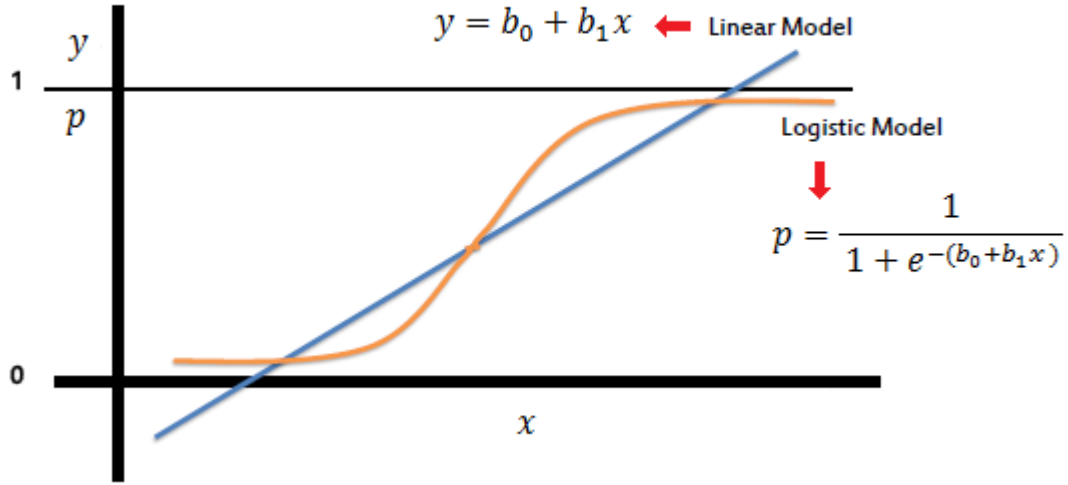
- Binomial Logistic Regression
- Multinominal Logistic Regression
- Ordinal Logistic Regression





https://www.statstest.com/multinomial-logistic-regression/
https://www.statstest.com/ordinal-logistic-regression/

Lojistik Resgresyon



$$\frac{p}{1-p} = \exp(b_0 + b_1 x)$$

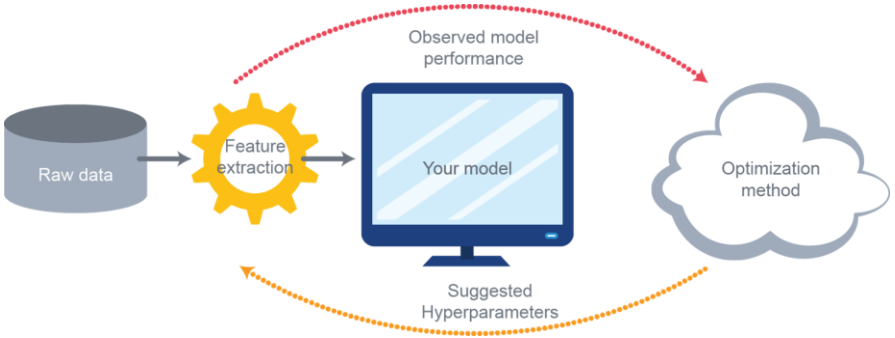
$$ln\left(\frac{p}{1-p}\right) = b_0 + b_1 x$$

$$p = \frac{1}{1 + e^{-(b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p)}}$$

https://www.saedsayad.com/logistic_regression.htm

Hiperparametre Optimizasyonu
(Logistic Regression)

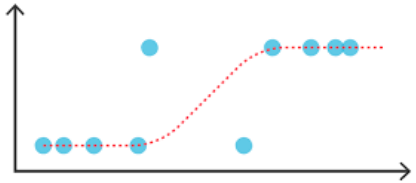
- penalty
- class_weight
- solver
- max_iter
- multi_class
- ...

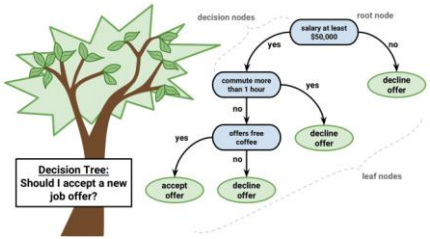


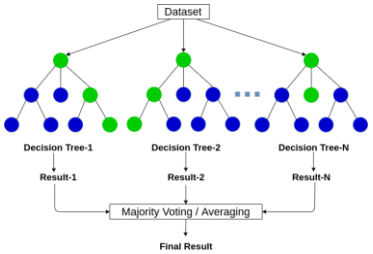
https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html
https://sigopt.com/blog/common-problems-in-hyperparameter-optimization/

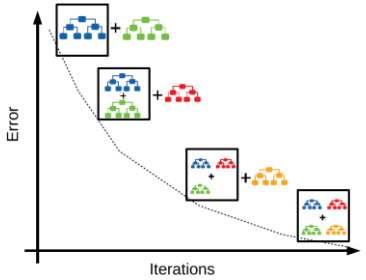
Algoritmalar

- Lojistik Regresyon
- **Karar Ağacı**
- Random Forest
- Gradient Boosting Tree





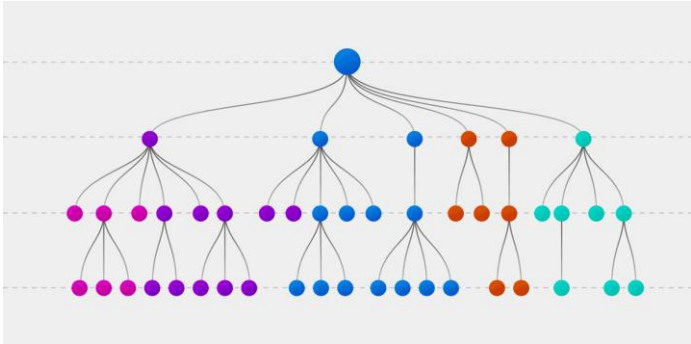




https://www.tibco.com/reference-center/what-is-logistic-regression
https://mychartguide.com/decision-tree/
https://www.tibco.com/reference-center/what-is-a-random-forest
https://morioh.com/p/e108a4521555

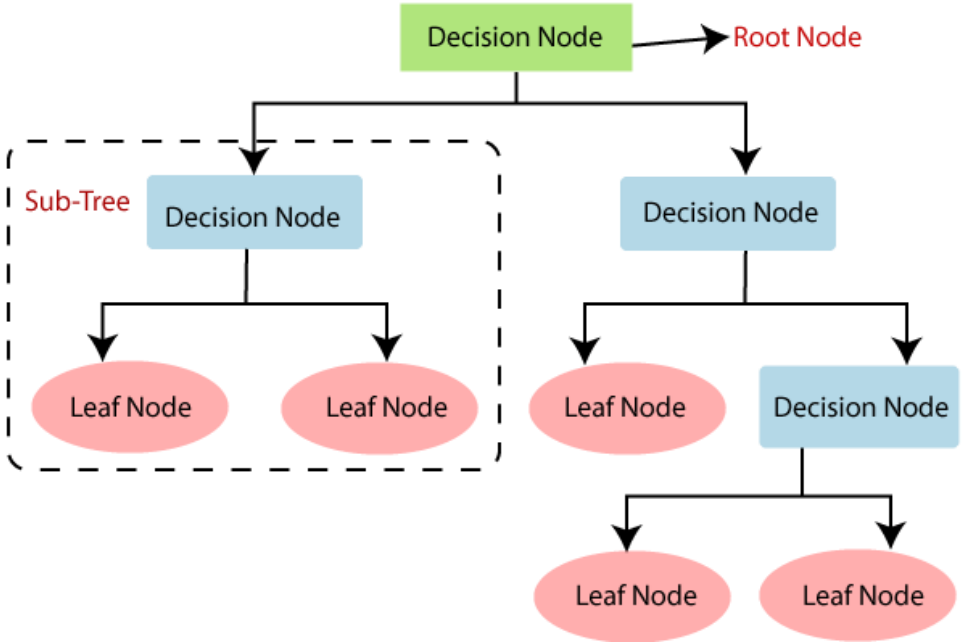
Karar Ağacı

- Karar ağaçları veri sınıflandırma probleminin bir çözümüdür.
- Karar ağacının elemanlarını dallar ve yapraklar oluşturur.
- Ağacın en üst düğümü kök düğümdür.
- En altta kalan yapılar yaprak ve kök-yaprak arasında kalan yapılar dallardır.
- Dallanma kriterleri
  - Entropi
  - **Sınıflandırma ve Regresyon (CART)**
  - Bellek Tabanlı



Karar Ağacı (CART)

- Her bir karar düğümünden itibaren ağaç iki dala ayrılır.
- En bilinen algoritmaları
  - Twoing Algoritması
  - **Gini Algoritması**



https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm

Karar Ağacı (CART / Gini Algoritması)

Adım 1: Her nitelik değerleri ikili olacak şekilde gruplandırılır.

Adım 2: Her nitelik için Sol ve Sağ taraftaki bölünmelere ait $Gini_{sol}$ ve $Gini_{sağ}$ hesaplanır.

$$Gini_{sol} = 1 - \sum_{i=1}^{k} \left[\frac{L_i}{|T_{sol}|}\right]^2 \qquad Gini_{sağ} = 1 - \sum_{i=1}^{k} \left[\frac{R_i}{|T_{sağ}|}\right]^2$$

Adım 3: Her nitelik için Gini değeri hesaplanır.

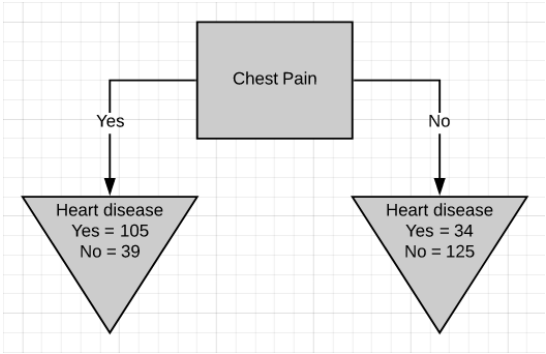
$$Gini_j = \frac{1}{n}(|T_{sol}|Gini_{sol} + |T_{sağ}|Gini_{sağ})$$

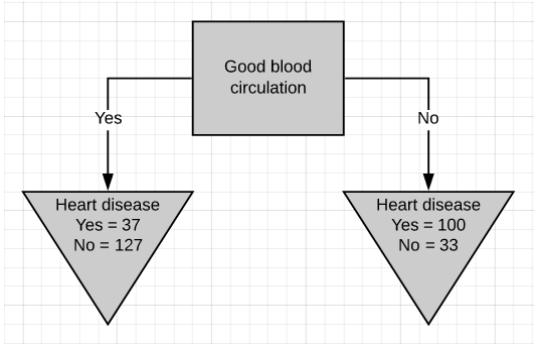
Adım 4: En küçük Gini değerine sahip nitelik seçilir ve bölünme bu düğüm üzerinden gerçekleştirilir.

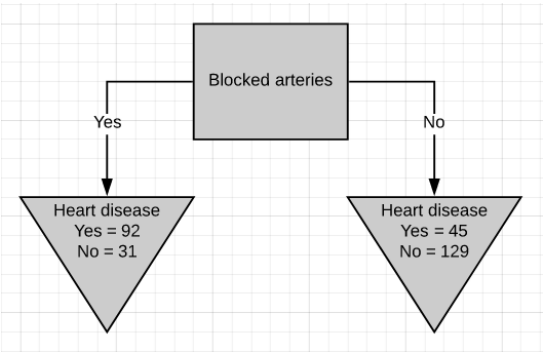
Adım 5: En baştaki adıma dönülerek yeniden dallanma gerçekleştirilir.

Karar Ağacı (CART)

Chest Pain	Good Blood Circulation	Blocked Arteries	Heart Disease
NO	NO	NO	NO
YES	YES	YES	YES
YES	YES	NO	NO
YES	NO	YES	YES
etc.	etc.	etc.	etc.

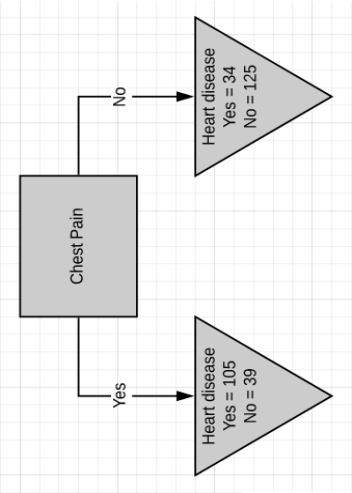






https://towardsdatascience.com/decision-trees-a-step-by-step-approach-to-building-dts-58f8a3e82596

Karar Ağacı (CART)




```
Gini impurity = 1 - (probability of 'yes')² - (probability of 'no')²
              = 1 - (34/34+125)² - (125/34+125)²
Gini impurity = 0.336
```

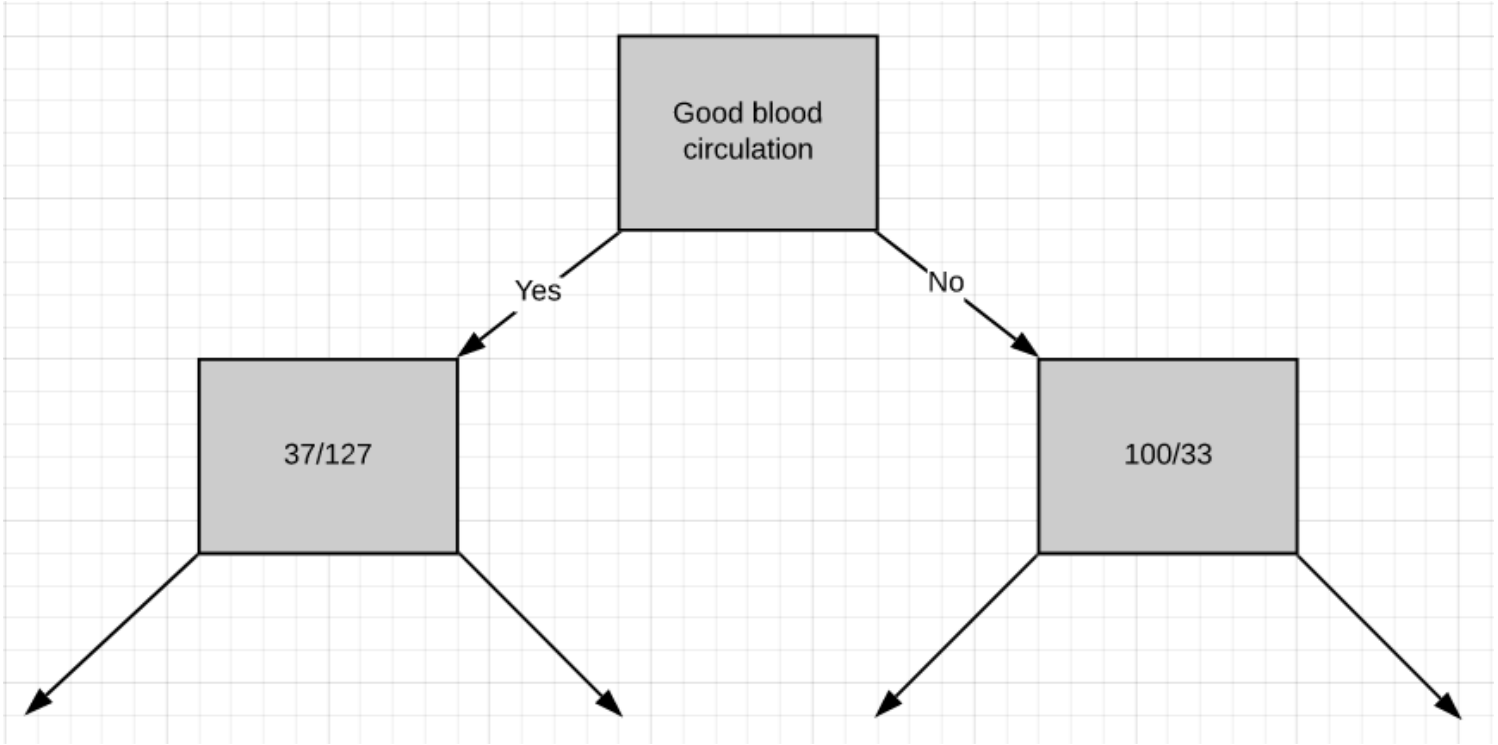
```
Gini impurity = 1 - (probability of 'yes')² - (probability of 'no')²
              = 1 - (105/105+39)² - (39/105+39)²
Gini impurity = 0.395
```

```
Gini impurity = (144/144+159)*0.395 + (159/144+159)*0.336
              = 0.364
```

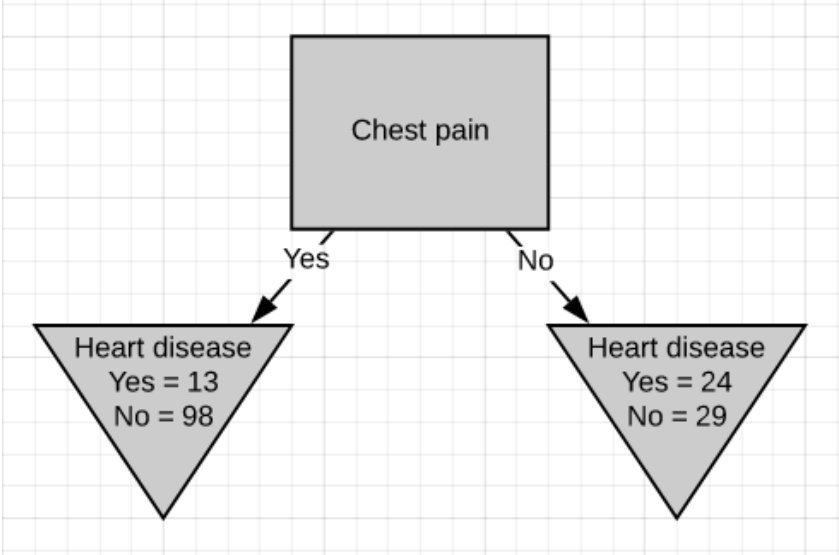
```
Gini impurity for 'good blood circulation' = 0.360
Gini impurity for 'blocked arteries' = 0.381
```

https://towardsdatascience.com/decision-trees-a-step-by-step-approach-to-building-dts-58f8a3e82596

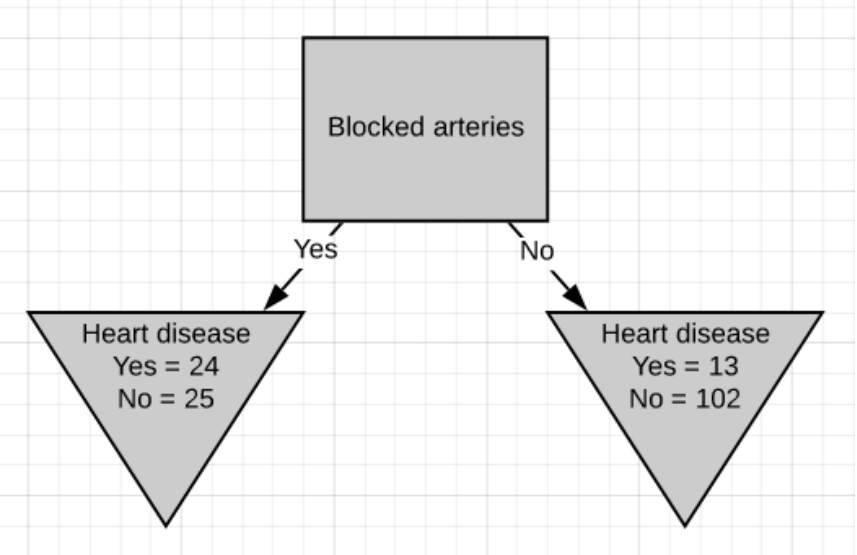
Karar Ağacı (CART)



Karar Ağacı (CART)

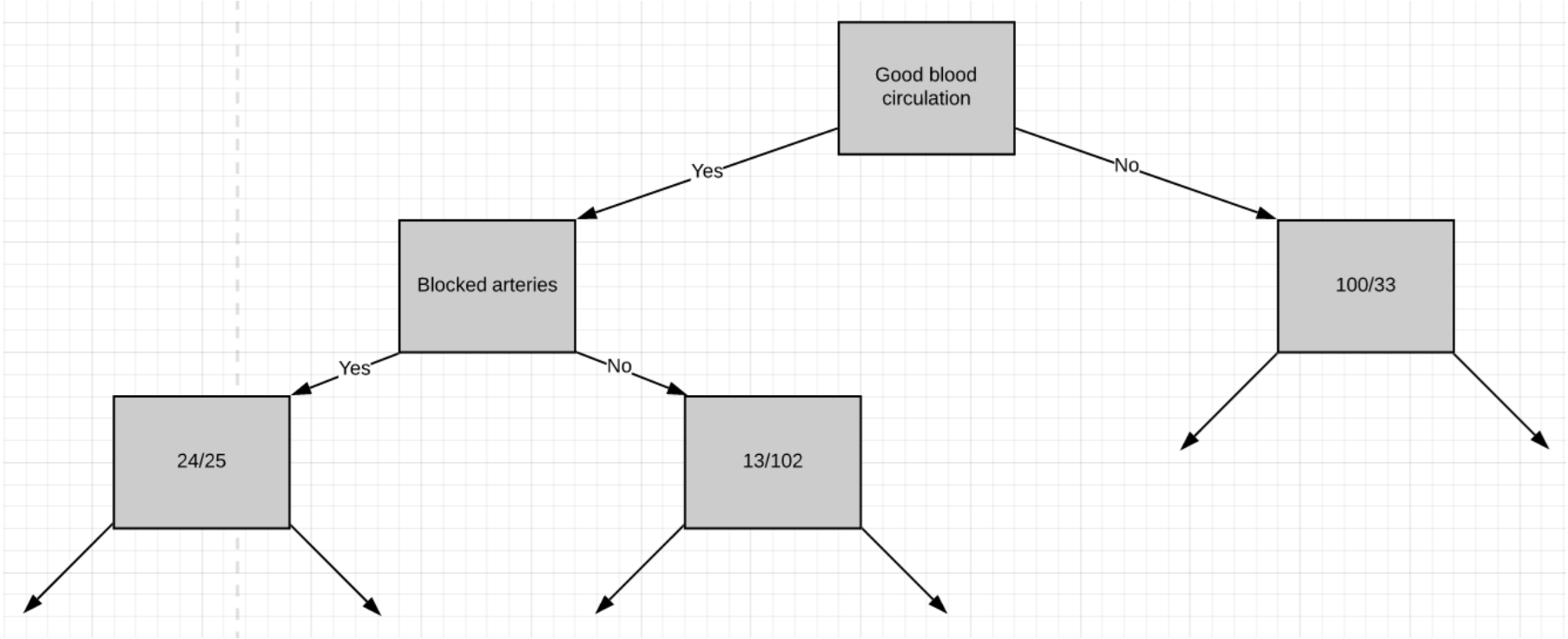


Gini Impurity: 0.3

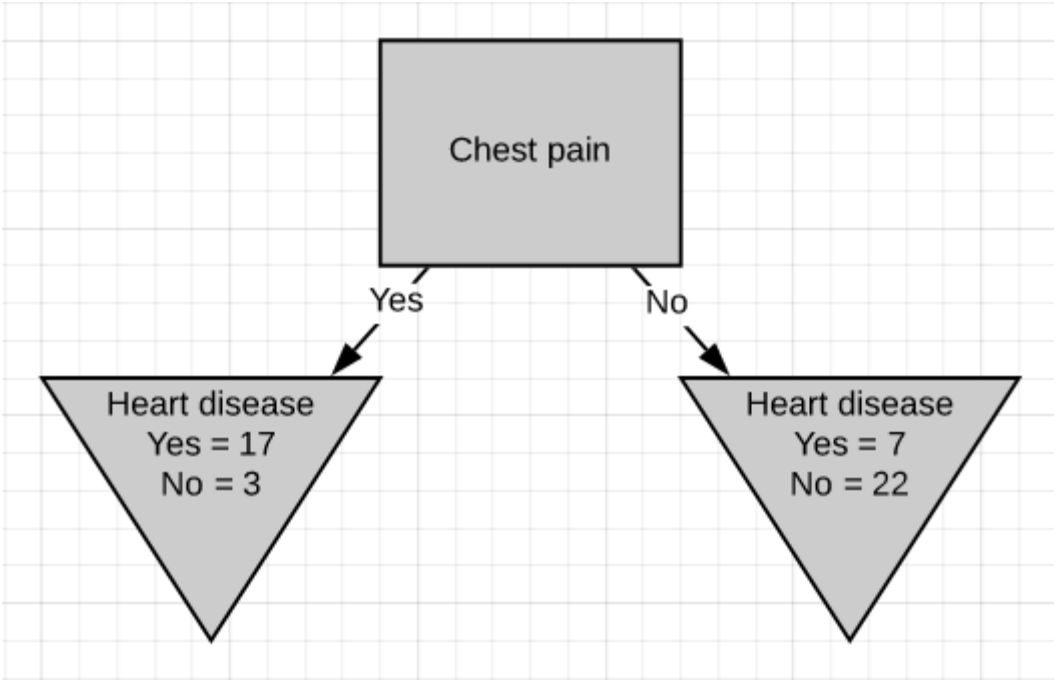


Gini Impurity: 0.29

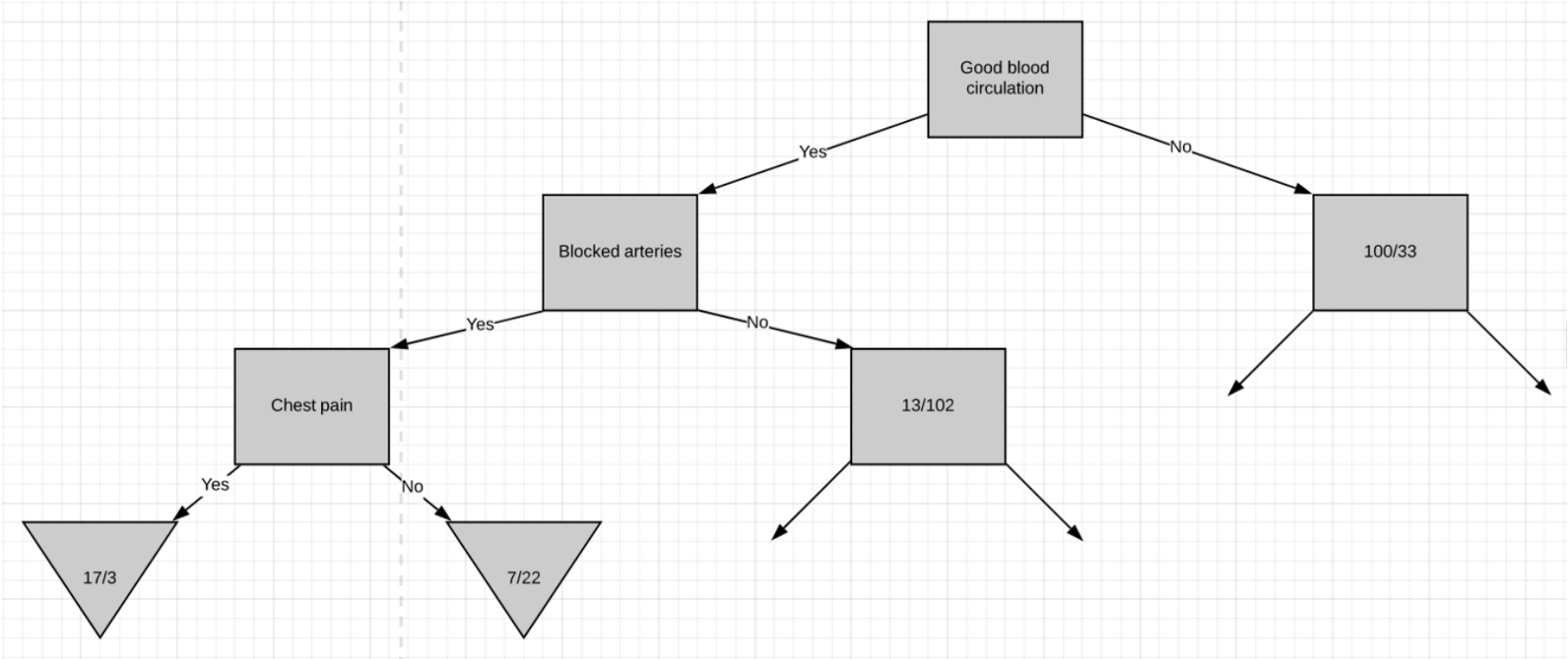
Karar Ağacı (CART)



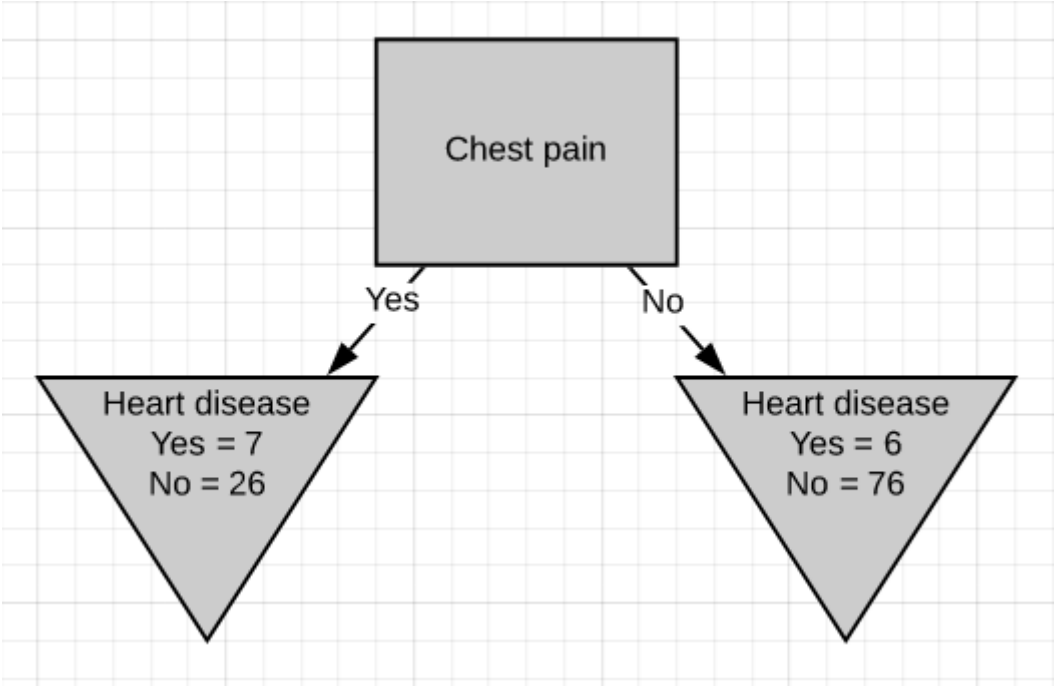
Karar Ağacı (CART)



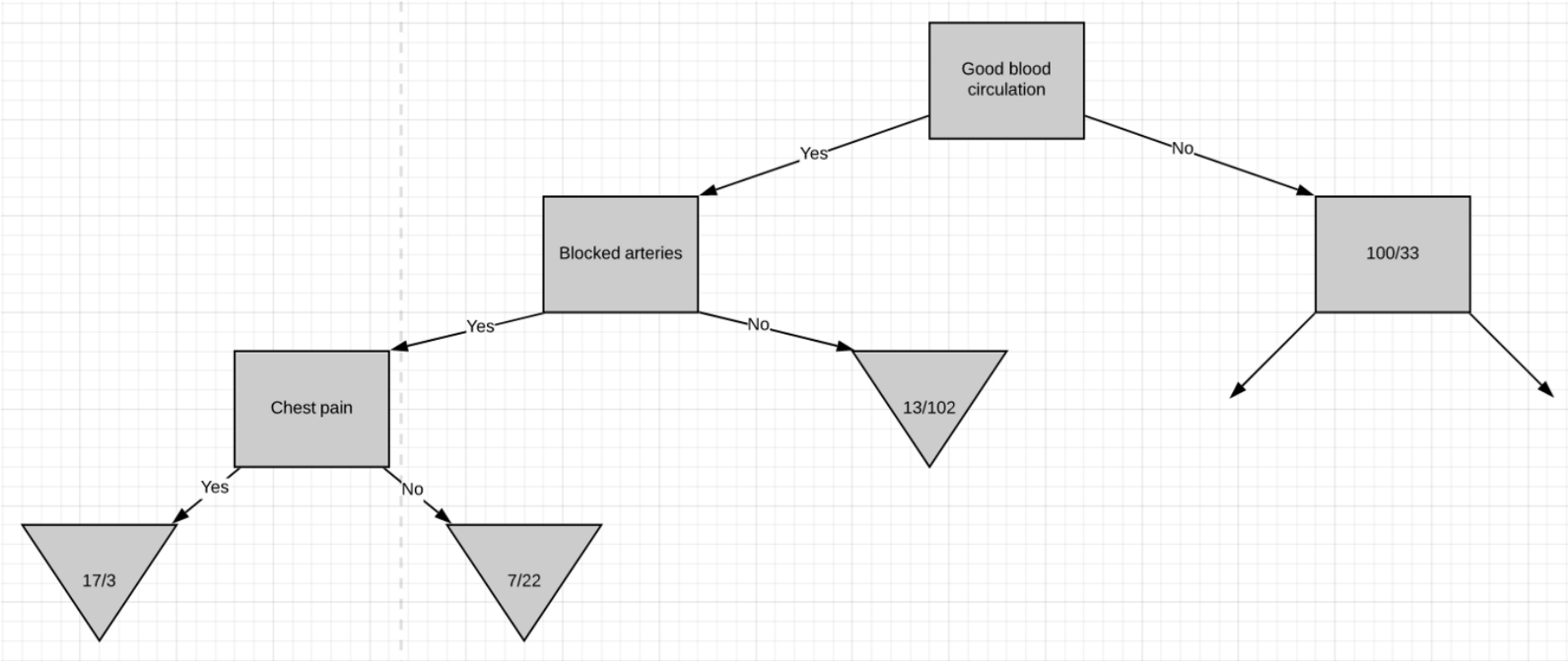
Karar Ağacı (CART)



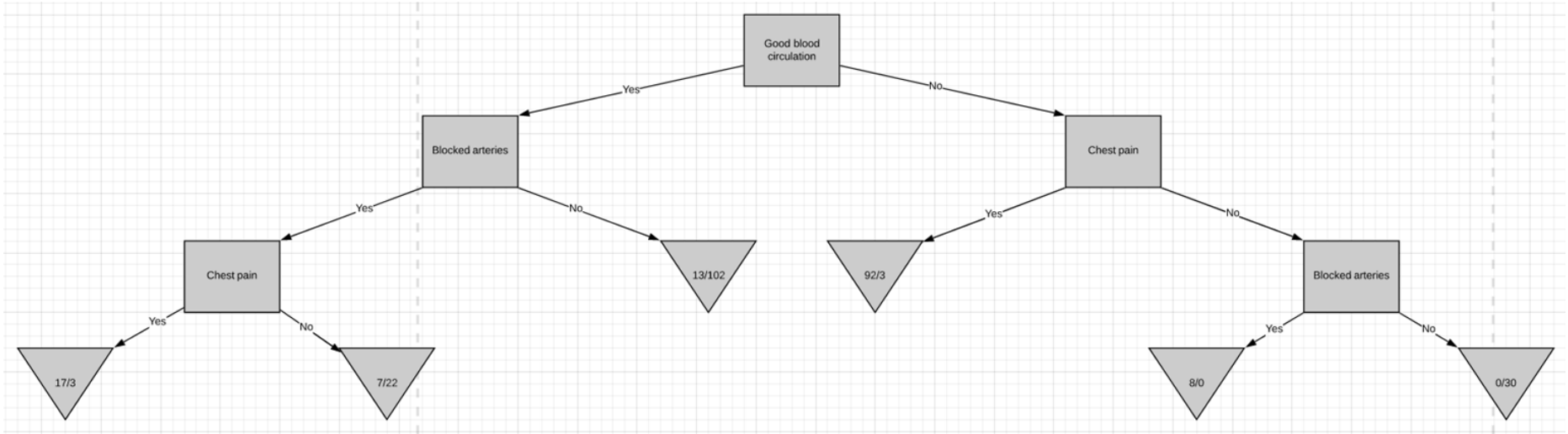
Karar Ağacı (CART)



Karar Ağacı (CART)

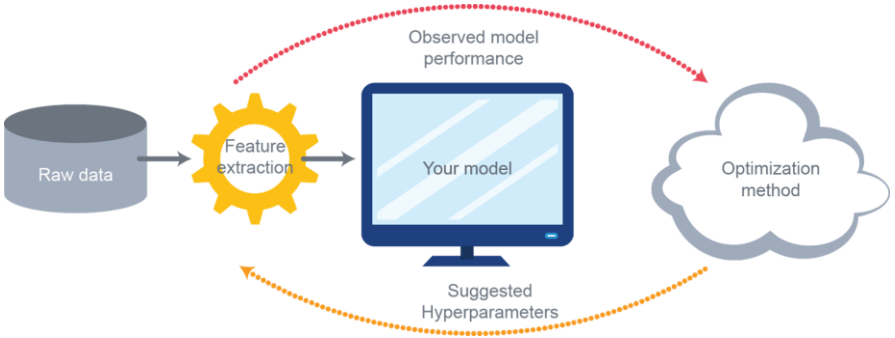


Karar Ağacı (CART)



Hiperparametre Optimizasyonu
(Karar Ağacı)

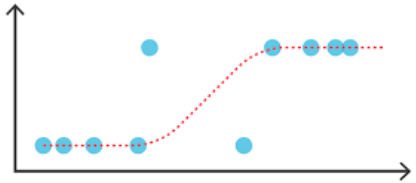
- criterion
- splitter
- max_depth
- min_samples_split
- max_leaf_nodes
- class_weight
- ...

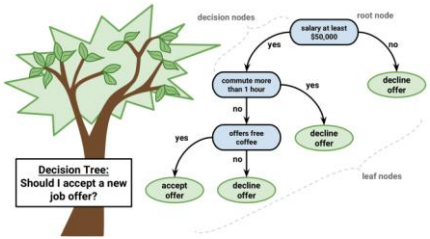


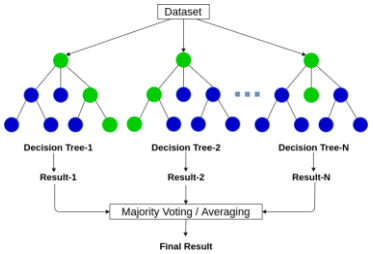
hhttps://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html
https://sigopt.com/blog/common-problems-in-hyperparameter-optimization/

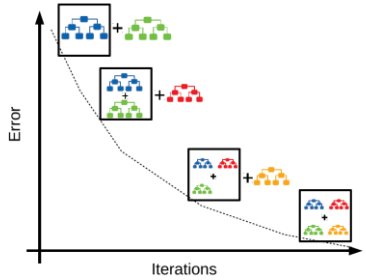
Algoritmalar

- Lojistik Regresyon
- Karar Ağacı
- **Random Forest**
- Gradient Boosting Tree





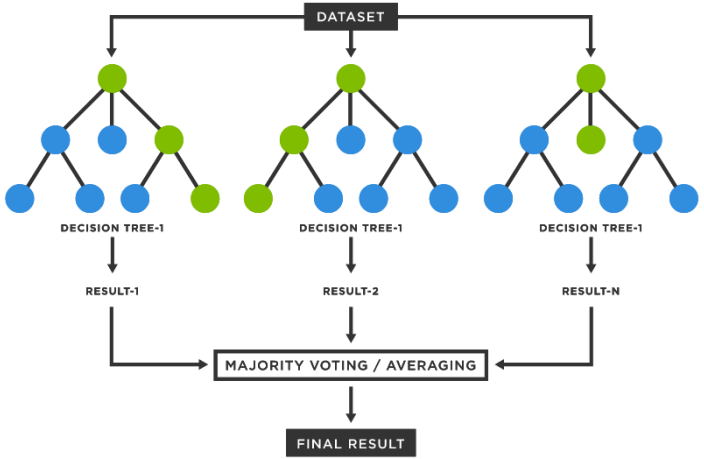




https://www.tibco.com/reference-center/what-is-logistic-regression
https://mychartguide.com/decision-tree/
https://www.tibco.com/reference-center/what-is-a-random-forest
https://morioh.com/p/e108a4521555

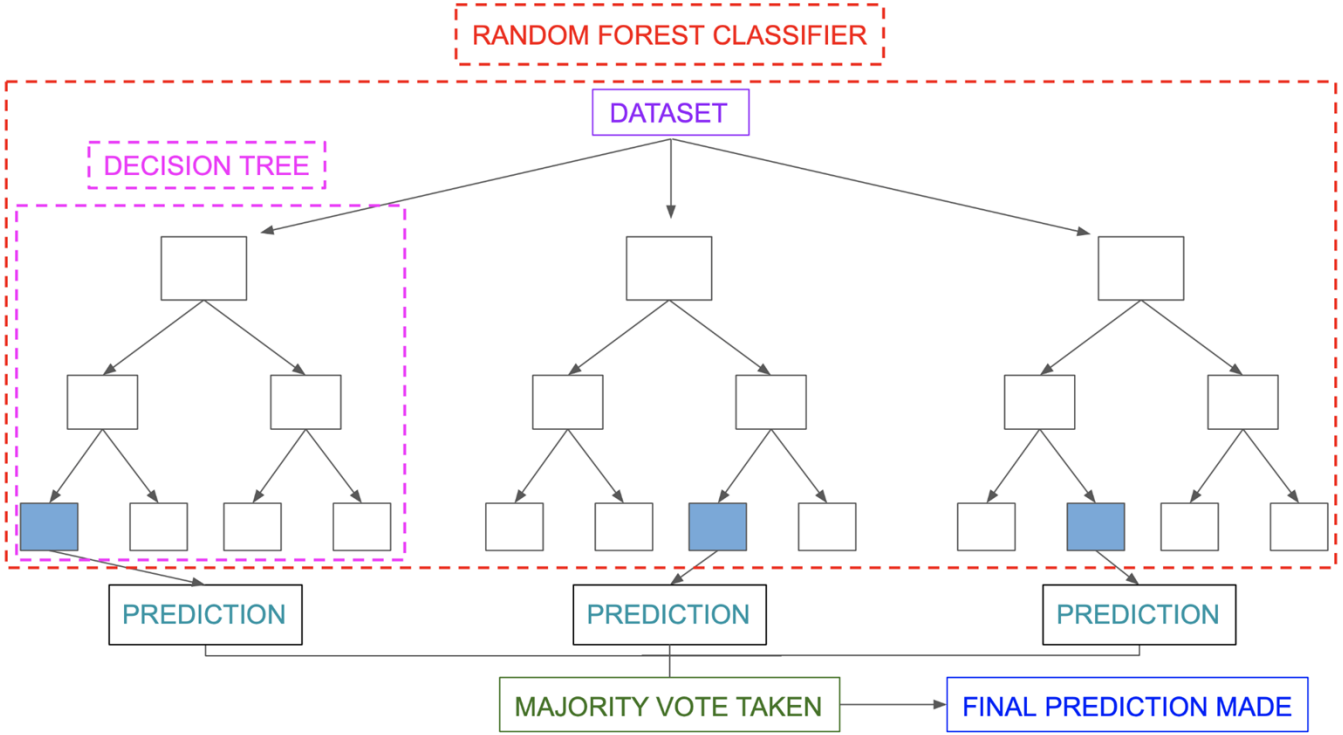
Random Forest

- **Adım 1:** Veri seti üzerinden n adet rastgele örnek kümesi seçilir
- **Adım 2:** Her veri seti için karar ağacı oluşturulur.
- **Adım 3:** Belirli bir test kümesi için her ağaçtan tahmin sonucu elde edilir.
- **Adım 4:** Test kümesindeki her örnek için en çok oylanan tahmin sonucu nihai sonucu verir..



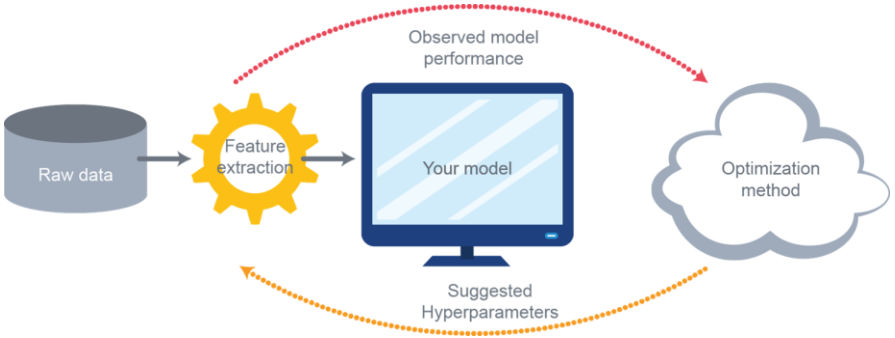
https://www.tibco.com/reference-center/what-is-a-random-forest

Random Forest



Hiperparametre Optimizasyonu
(Random Forest)

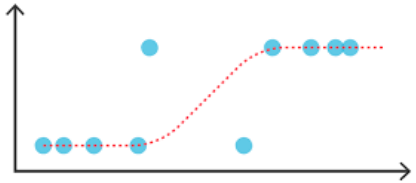
- n_estimators
- criterion
- max_depth
- min_samples_split
- max_leaf_nodes
- class_weight
- bootstrap
- ...

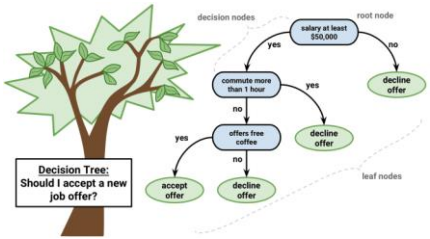


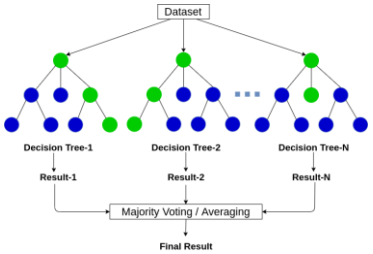
https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
https://sigopt.com/blog/common-problems-in-hyperparameter-optimization/

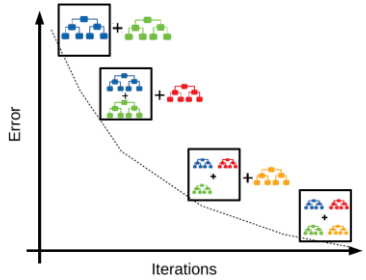
Algoritmalar

- Lojistik Regresyon
- Karar Ağacı
- Random Forest
- **Gradient Boosting Tree** **(https://www.youtube.com/watch?v=jxuNLH5dXCs)**



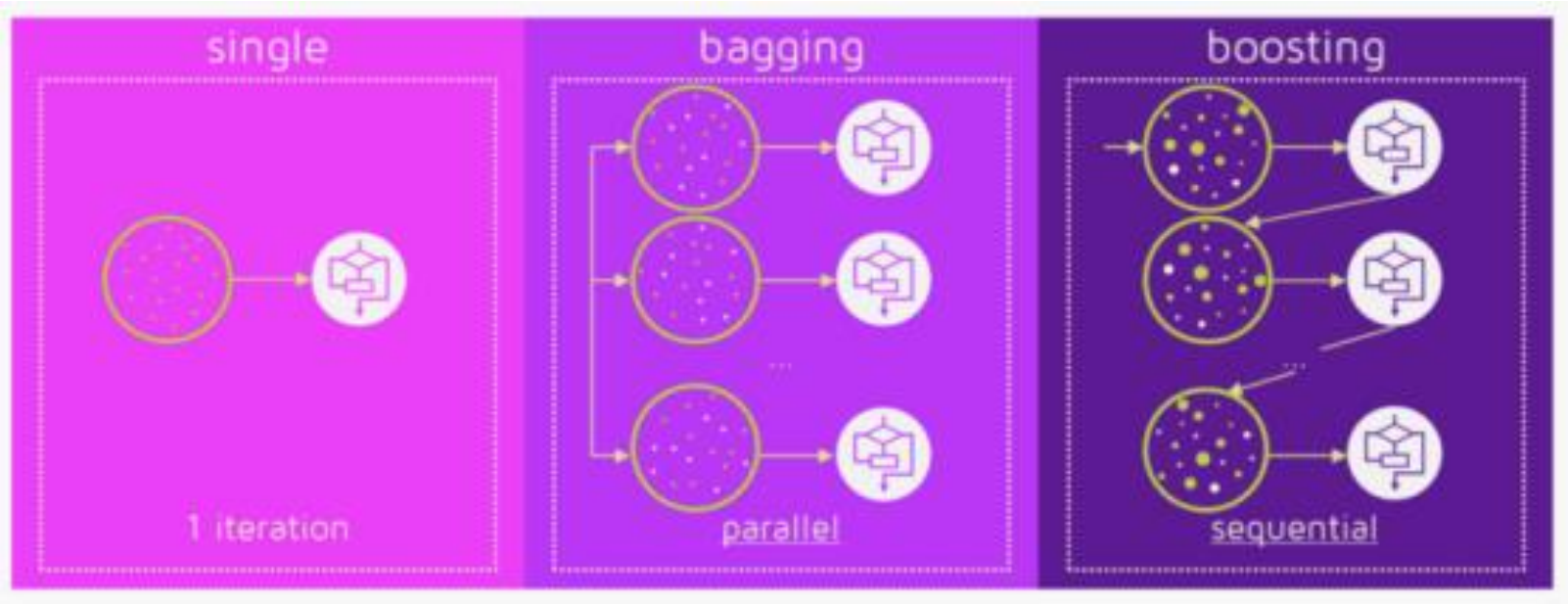






https://www.tibco.com/reference-center/what-is-logistic-regression
https://mychartguide.com/decision-tree/
https://www.tibco.com/reference-center/what-is-a-random-forest
https://morioh.com/p/e108a4521555

Algoritmalar



https://devhunteryz.wordpress.com/2018/07/11/gradyan-arttirmagradient-boosting/