

Seminer Dersi:

Konu: Random Forest Regresyon

Yücel Aytaç Akgün 20360859059

Bilgisayar Mühendisliği 3.Sınıf

21/03/2024



İçindekiler

- I. Regresyon Nedir?**
- II. Regresyon'daki Varsayımlar**
- III. Regresyon Tarihçesi**
- IV. Regresyon vs Classification**
- V. Regresyon Modelleri**
- VI. Random Forest Regresyon**
- VII. Random Forest vs Decision Tree**

İçindekiler(Devam)

VIII. Error ve Verim Skorları

- Mean Square Error
- Mean Absolute Error
- R^2 Score

Regresyon Nedir?

- **Regresyon analizi**, iki ya da daha çok nicel değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur.
- **Regresyon analizi** ile değişkenler arasındaki ilişkinin varlığı, eğer ilişki var ise bunun gücü hakkında bilgi edinilebilir.

Regresyon Nedir?(Devam...)

- İki değişken varsa, **tahminin temelini oluşturan değişken bağımsız değişkendir**. Değeri tahmin edilecek **değişken bağımlı değişken olarak bilinir**.

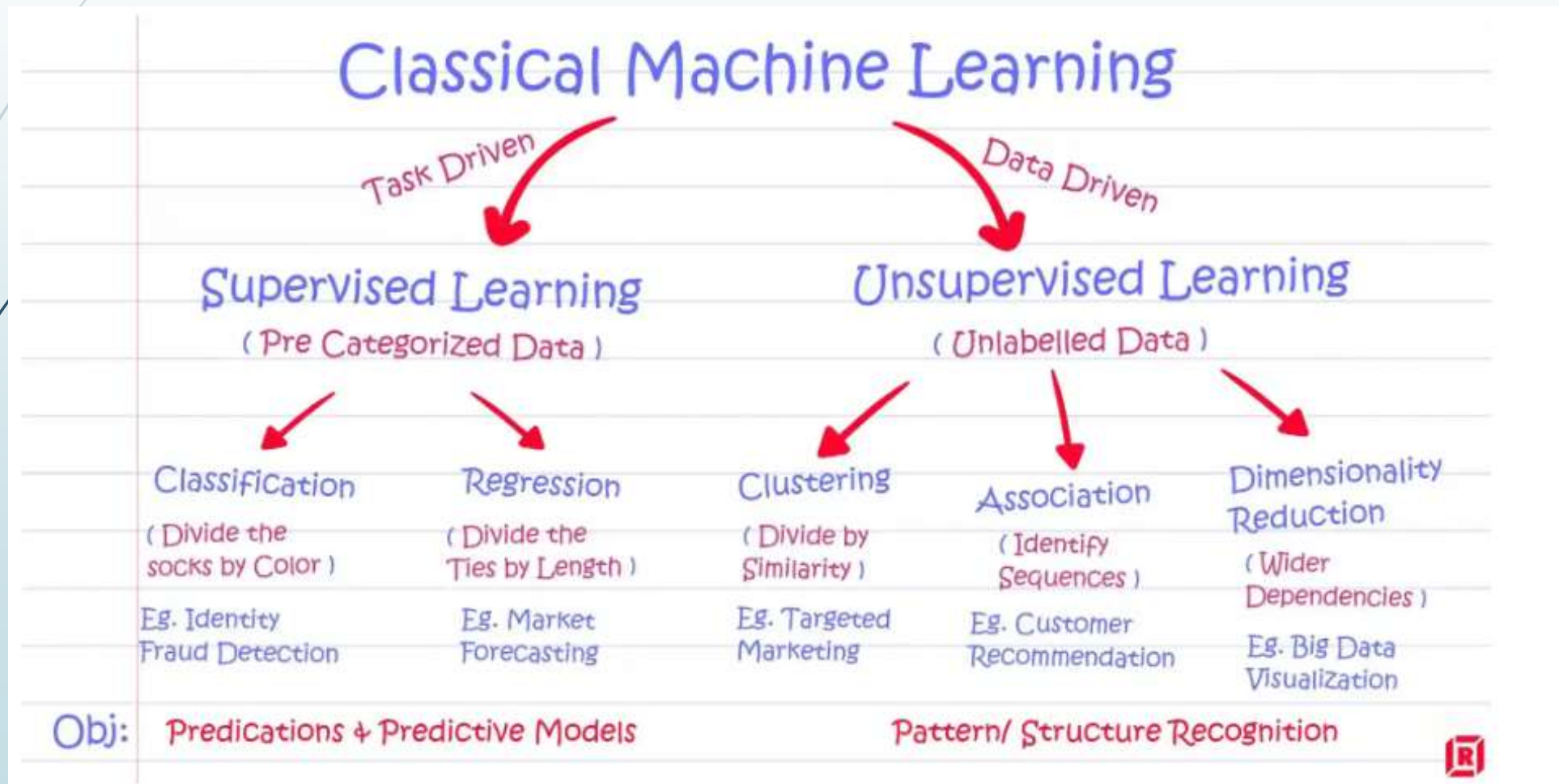
Regresyon'daki Varsayımlar

- Bağımlı değişken içinde hata bulunmaktadır.
- Bu hatalar **Sistemik hata** veya rassal değişken olduğu söylenir.
- Bağımsız değişken hatasızdır.

Regresyon Tarihçesi

- Regresyon yönteminin ilk şekli **en küçük kareler** prensibidir ve ilk olarak **Adrien Marie Legendre** tarafından 1805 yılında ortaya atılmıştır.
- Hemen takiben 1809 yılında **Gauss** aynı yöntemi açıklamıştır.

Regresyon vs Classification



DEVAM...

	<i>Supervised Learning</i>	<i>Unsupervised Learning</i>
<i>Discrete</i>	classification or categorization	clustering
<i>Continuous</i>	regression	dimensionality reduction



Regresyon Modelleri

- ❑ Simple Linear Regression
- ❑ Multiple Linear Regression
- ❑ Polynomial Regression
- ❑ Support Vector for Regression (SVR)
- ❑ Decision Tree Regression
- ❑ Random Forest Regression

Random Forest Regresyon

- ❑ Karar ağaçlarının(Decision Tree)'nin en büyük problemlerinden biri aşırı öğrenme-veriyi ezberlemesidir (overfitting).
- ❑ Random Forrest Regresyon ise, 10'larca 100'lerce karar ağacı oluşturuyor ve her bir karar ağacı bireysel olarak tahminde bulunuyor.
- ❑ Ve bu olayların sonunda tahminlerin ortalaması alınıyor.



Random Forest vs Decision Tree

- **Random Forest** Regresyon **Karar Ağaçlarına göre (Decision Tree)** daha iyi tahmin yapar.
- Fakat **Random Forest'in** negatif tarafları ise çok daha karmaşıktır ayrıca görselleştirmesi daha zordur. Karar ağaçlarının görselleştirilmesi daha basittir.

Error ve Verim Skorları

$$\text{MSE} = \overset{\text{Mean}}{\boxed{\frac{1}{n} \sum_{i=1}^n}} \overset{\text{Error}}{\boxed{(Y_i - \hat{Y}_i)}} \overset{\text{Squared}}{\boxed{^2}}$$

Error ve Verim Skorları

Mean Absolute Error (MAE):

The diagram illustrates the Mean Absolute Error (MAE) formula with the following components and annotations:

- Divide by the total number of data points:** Points to the $\frac{1}{n}$ term in the formula.
- Sum of:** Points to the Σ symbol.
- Actual output value:** Points to the y term inside the absolute value.
- Predicted output value:** Points to the \hat{y} term inside the absolute value.
- The absolute value of the residual:** Points to the entire absolute value expression $|y - \hat{y}|$.

$$MAE = \frac{1}{n} \sum |y - \hat{y}|$$

Error ve Verim Skorları

R Squared(R^2) Score:

Formula

$$R^2 = 1 - \frac{RSS}{TSS}$$

R^2 = coefficient of determination

RSS = sum of squares of residuals

TSS = total sum of squares

$$RSS = \sum (y_i - \hat{y}_i)^2$$

Where: y_i is the actual value and, \hat{y}_i is the predicted value.

$$TSS = \sum (y_i - \bar{y})^2$$

Where: y_i is the actual value and \bar{y} is the mean value of the variable/feature



KAYNAKÇA

- <https://bulutistan.com/blog/regresyon>
- https://tr.wikipedia.org/wiki/Regresyon_analizi
- <https://www.labellerr.com/blog/supervised-vs-unsupervised-learning-whats-the-difference/>
- <https://dev.to/afozbek/supervised-learning-vs-unsupervised-learning-4b65>
- <https://medium.com/data-science-tr/makine-%C3%B6%C4%9Frenmesi-dersleri-5a-random-forest-regresyon-2a91715a8b66>
- <https://suboptimal.wiki/explanation/mse>
- <https://medium.com/@polanitzer/the-minimum-mean-absolute-error-mae-challenge-928dc081f031>
- <https://www.freecodecamp.org/news/evaluation-metrics-for-regression-problems-machine-learning>



Sorularınız?





Beni Dinlediğiniz için
Teşekkür Ederim