



Mehmet Fatih AKCA

Follow

Feb 22, 2021 · 4 min read



Save



## Lojistik Regresyon Nedir? Nasıl Çalışır?



Open in app ↗

Sign up

Sign In

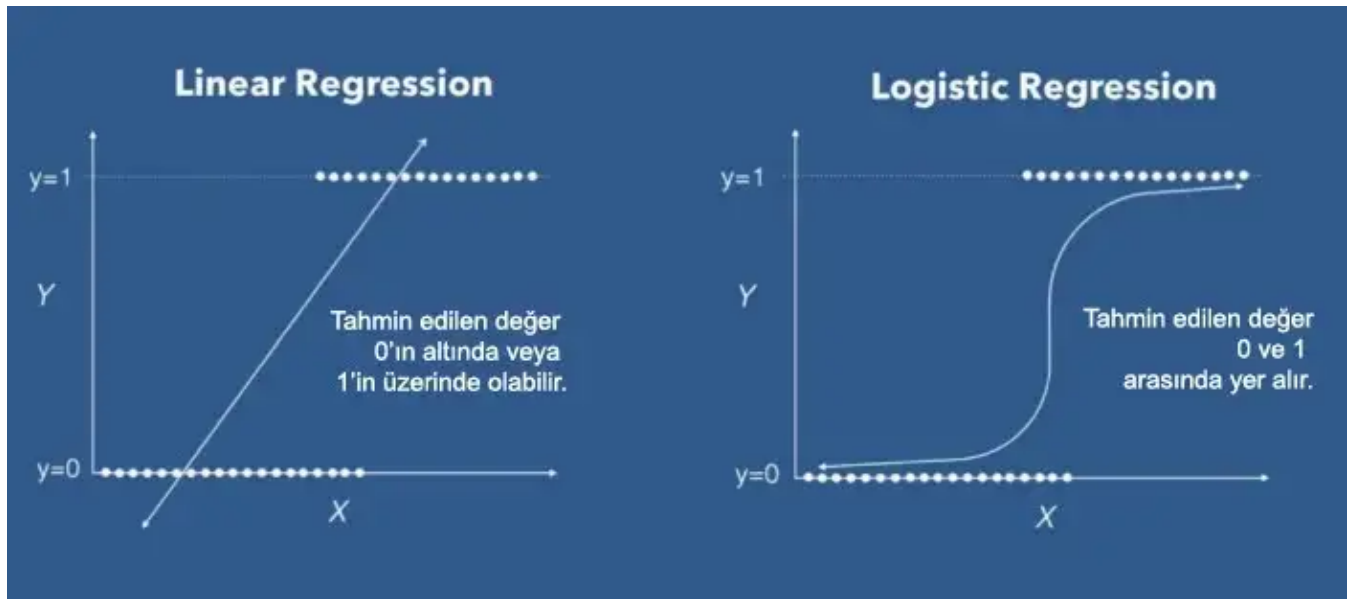


$$y = 1 + e^{-x}$$

Kaynak: <https://sefiks.com/2017/01/21/sigmoid-function-as-an-activation-function/>

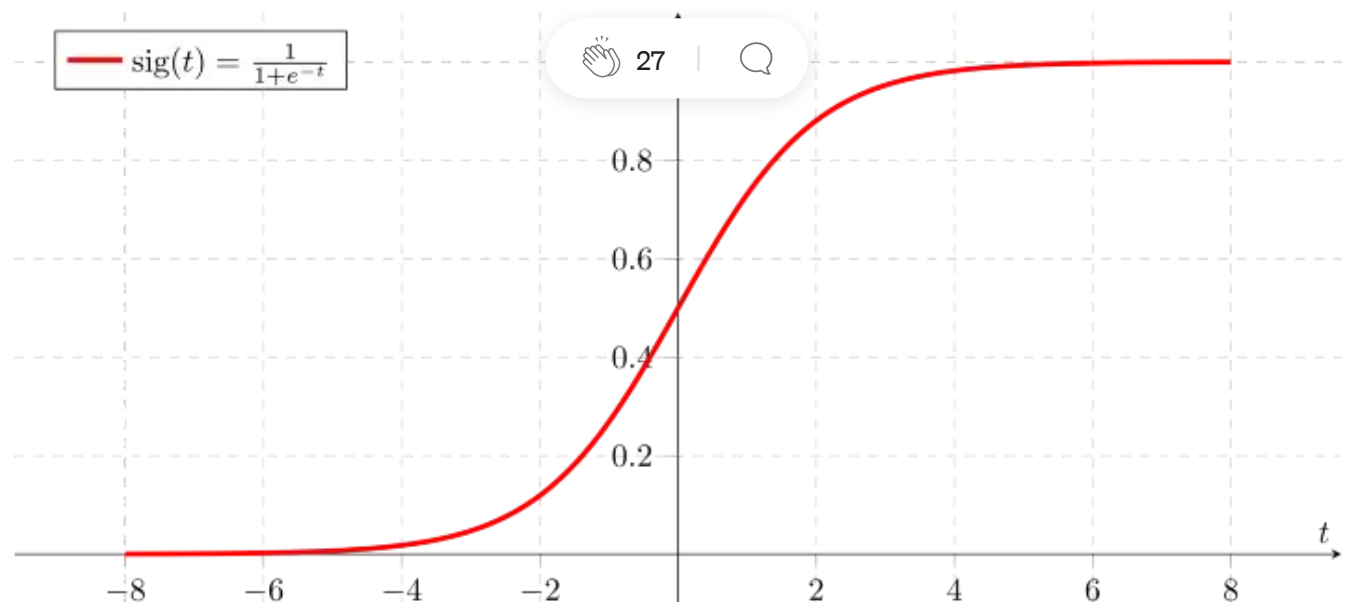
Lojistik regresyon, isminde “**regresyon**” geçmesine rağmen bir **sınıflandırma** algoritmasıdır. Yani görseldeki hayvanın kedi mi, köpek mi olduğu veya verilmiş olan bilgilerin bir erkeğe mi yoksa bir kadına mı ait olduğunu tahmin etme gibi iki sınıflı sınıflandırma problemlerinde sıkça kullanılır.

Lojistik regresyonun, lineer regresyon ile arasındaki en büyük farkı iki sınıflı birbirinden ayıracak çizgiyi nasıl uyguladığıdır (fit). Lineer regresyon, optimum çizgiyi çizmek için “*En Küçük Kareler Yöntemi*” (Least Squares) kullanırken, lojistik regresyon “*Maksimum Olabilirlik*” (Maximum Likelihood) kullanır.



Kaynak: <https://medium.com/@nesayan.saha/machine-learning-basics-logistic-regression-classification-254a8342d1b7>

Lojistik regresyon, sınıflandırma yapmak için **Sigmoid** (Lojistik) Fonksiyonu kullanır. Sigmoid fonksiyonu “S” şeklinde bir eğridir.



## 1. Sigmoid Fonksiyonu nedir?

Sigmoid fonksiyonu basitçe, verilerimizi 0 ve 1 arasına sıkıştırmak için kullanılan fonksiyondur. Bu fonksiyon sayesinde sınıflandırma yapabiliriz. Derin Öğrenme içerisinde aktivasyon fonksiyonları altında da sıkça kullanılır.

**Formülü:**

$$g(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(\theta^T x)}}$$

Sigmoid Fonksiyonu Formülü

**Bir örnek yapalım:**

$$\theta = [1, 2, 3] \quad x = [1, 2, 2]$$

İşlemleri yapmamız için  $\theta$  (katsayı/ağırlık) ve  $x$  (girdi değerleri) değerlerini almamız gerekiyor.

$$\theta^T x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} * [1 \ 2 \ 2] = [ (1*1) + (2*2) (3*2) ] = 11$$

$\theta^T x$  değerini buluyorum.

$$\frac{1}{1 + e^{-(\theta^T x)}} = \frac{1}{1 + e^{-11}} = 0,999983...$$

Yani 0,999983 olasılıkla sonuç 1 olur. Bu değer, Lojistik Regresyonun'un eşik değeri (threshold) olan 0.5'in üzerinde olduğu için sınıflandırma sonucu 1 olur.

$$h(x) = \begin{cases} > 0.5, \text{ eğer } \theta^T x > 0 \longrightarrow \text{Sınıfı 1 olarak tahmin eder.} \\ < 0.5, \text{ eğer } \theta^T x < 0 \longrightarrow \text{Sınıfı 0 olarak tahmin eder.} \end{cases}$$

$\theta^T x > 0$  ise  $h(x) > 0.5$ 'tir ve tahmin edilen sınıfı 1 olur. Aksine  $\theta^T x < 0$  ise  $h(x) < 0.5$ 'tir ve tahmin edilen sınıfı 0 olur.

Rastgele verilmiş ağırlıklar üzerinden bir tahmin yaptık ama doğru (optimum) parametreleri nasıl buluyoruz? Hemen anlatayım.

## 2. Doğru parametreleri nasıl buluyoruz?

Öncelikle tekil bir sonuç için Maliyet (Cost) Fonksiyonuna bakalım.

$$\text{Cost}(h_{\theta}(x), y) = \begin{cases} -\log(h_{\theta}(x)) & \text{eğer } y=1 \\ -\log(1 - h_{\theta}(x)) & \text{eğer } y=0 \end{cases}$$

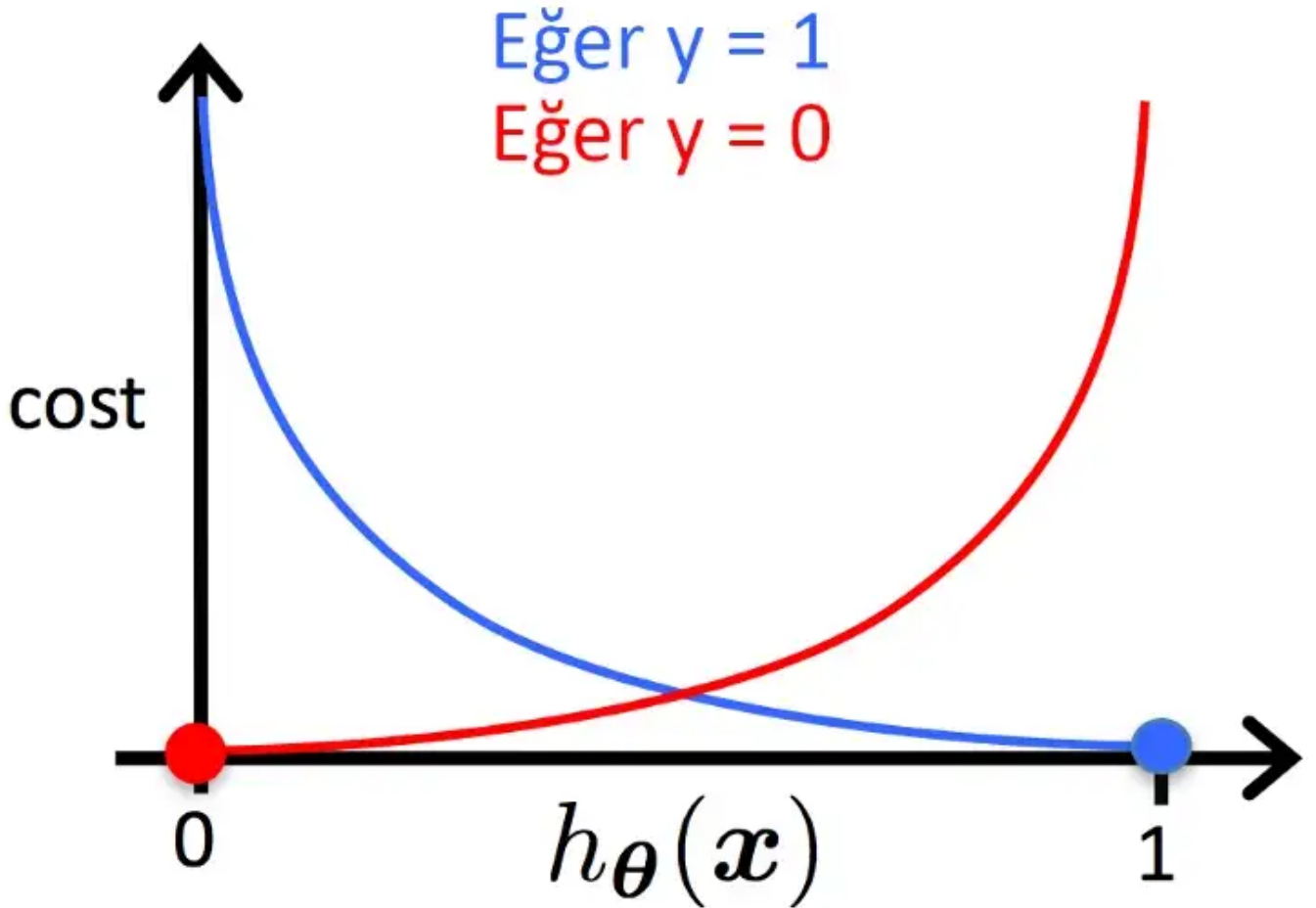
Lojistik Regresyon için Maliyet (Cost) Fonksiyonu

Yani gerçek  $y$  değeri 1 ise üstteki, 0 ise alttaki formül kullanılıyor. Bu formülü tek satırda yazmak gerekirse:

$$\text{Cost}(h_{\theta}(x), y) = -y \log(h_{\theta}(x)) - (1-y) \log(1 - h_{\theta}(x))$$

Bu nasıl oluyor dediğiniz duyar gibiyim. Daha yakından bakalım. Eğer  $y = 1$  ise ikinci kısımda, parantez içerisindeki  $1-y$ 'nin sonucu 0 olacak ve ikinci kısım 0 ile çarpıldığı için 0 sonucunu alacak ve elimizde sadece birinci kısım kalacak. Tam tersine, eğer  $y = 0$  ise, birinci kısımda, eşittirin sağındaki  $y$ , birinci kısmı 0'a eşitleyecek ve elimizde sadece ikinci kısım kalacak.

Maliyet Fonksiyonuna bir grafik ile bakalım.



X ekseninde “Tahmin Edilen Değer”, y ekseninde “Maliyet” var. Bu grafikte mavi renkli eğri  $-\log(h_{\theta}(x))$ , kırmızı renkli eğri ise  $-\log(1 - h_{\theta}(x))$ 'in eğrisidir. Kaynak:

<https://datascience.stackexchange.com/questions/40982/logistic-regression-cost-function>

Eğer  $y = 1$ , tahmin edilen değerde 1 ise cost (y eksen) 0 olur ama  $y = 1$  iken tahmin edilen değer 0 ise cost sonsuza gider. Diğer taraftan  $y = 0$ , tahmin edilen değerde 0

ise cost 0 olur ama  $y = 0$  iken tahmin edilen değer 1 ise cost sonsuza gider. Bu şekilde yanlış yaptıklarını cezalandırabilir.

Bu formül bize tekil bir değer için sonuç verir. Bütün veri seti içerisinde değerlendirme yapmak için bu formülü biraz değiştirmemiz gerekecek.

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_{\theta}(x^{(i)}), y^{(i)})$$

$m$  = örnek sayısı

Bu kadar teori yeter, şimdi bunları uygulayalım.

```
1  # Gerekli kütüphanelerin içe aktarılması
2  import numpy as np
3  from sklearn.linear_model import LogisticRegression
4  import matplotlib.pyplot as plt
5  import seaborn as sns
6
7
8
9  # X ve y değerlerinin oluşturulması
10 x = np.arange(100).reshape(-1, 1)
11 y = np.array([0]*50 + [1]*50)
12
13
14
15
16 # Verisetinin train ve test olarak ayrılması
17 from sklearn.model_selection import train_test_split
18
19 X_train, X_test, y_train, y_test = train_test_split(x,y, test_size =.3)
20
21
22
23 # Modelin eğitilmesi
24 model = LogisticRegression().fit(X_train,y_train)
25
26
27 y_pred = model.predict(X_test)
28
29
30
31 # Accuracy score değerinin hesaplanması
32 from sklearn.metrics import accuracy_score
33
34 print("Accucarcy score: ", accuracy_score(y_test,y_pred))
35
36
37 # Karmaşıklık matrisinin ısı haritası olarak çizilmesi
38 from sklearn.metrics import confusion_matrix
39
40 sns.heatmap(confusion_matrix(y_test,y_pred),annot=True)
41 plt.show()
```

Accuracy score: 0.96666...

0	15	1
1	0	14
	0	1

Karmaşıklık Matrisi

Biraz da avantaj/dezavantajından bahsedelim.

**Avantajları:**

1. Lojistik regresyonun uygulanması, yorumlanması kolaydır.
2. Veri seti doğrusal olarak ayrılabiliriyorsa oldukça iyi performans gösterir.
3. Overfitting'e daha az meyillidir ama büyük veri setlerinde overfit olabilir.

**Dezavantajları:**

1. Gözlem sayısı özellik sayısından azsa, Lojistik Regresyon kullanılmamalıdır, aksi takdirde overfit olabilir.
2. Lojistik regresyonun ayırım yapabilmesi için veri setinin doğrusal olarak ayrılabiliriyor olması lazım.

Lojistik Regresyon bu kadardı. Okuduğunuz için teşekkür ederim.

Esen Kalın.



## Kaynaklar:

<b>Machine Learning</b> Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade... <a href="http://www.coursera.org">www.coursera.org</a>	
<b>StatQuest with Josh Starmer</b> Statistics, Machine Learning and Data Science can sometimes seem like very scary topics, but since each technique is... <a href="http://www.youtube.com">www.youtube.com</a>	

<https://www.datacamp.com/>

Logistic Regression

Lojistik Regresyon

Makine Ogrenmesi

Yapay Zeka

Siniflandirma

[About](#) [Help](#) [Terms](#) [Privacy](#)

Get the Medium app

