

REGRESYON NEDİR?

Regresyon, bir bağımlı değişkenin bir veya daha fazla bağımsız değişken tarafından açıklanması veya tahmin edilmesi amacıyla kullanılan istatistiksel bir yöntemidir.

REGRESYON ANALİZİ NEDİR?

Regresyon analizi, iki ya da daha çok nicel değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur. Eğer tek bir değişken kullanılarak analiz yapılıyorsa buna tek değişkenli regresyon, birden çok değişken kullanılıyorsa çok değişkenli regresyon analizi olarak isimlendirilir.

1. **Amaç:** Regresyon analizinin temel amacı, bir bağımlı değişkenin bir veya daha fazla bağımsız değişken tarafından açıklanması veya tahmin edilmesidir. Örneğin, bir kişinin gelirini, eğitim düzeyini ve yaşını kullanarak ev sahibi olma olasılığını tahmin etmek gibi.
2. **Doğrusal ve Doğrusal Olmayan Regresyon:** Regresyon analizi, doğrusal regresyon (lineer regresyon) ve doğrusal olmayan regresyon olarak iki ana kategoriye ayrılabilir. Doğrusal regresyon, bağımlı değişken ile bağımsız değişken arasındaki ilişkinin doğrusal olduğunu varsayar. Doğrusal olmayan regresyon ise, bu ilişkinin doğrusal olmadığı durumlar için kullanılır.
3. **Modelleme Süreci:** Regresyon analizi genellikle şu adımları içerir: veri toplama, model seçimi, model uygulama, modelin değerlendirilmesi ve sonuçların yorumlanması. Bu süreçte, uygun bir model seçmek için veri analizi ve model uyumu değerlendirilir.
4. **Tahmin ve Açıklama:** Regresyon analizi, bir değişkenin değerini tahmin etmenin yanı sıra, değişkenler arasındaki ilişkileri de açıklar. Bu nedenle, tahmin yapma ve neden-sonuç ilişkilerini anlama açısından kullanışlıdır.
5. **Uygulama Alanları:** Regresyon analizi çeşitli uygulama alanlarında kullanılır, örneğin ekonomi, finans, pazarlama, sosyal bilimler, sağlık bilimleri ve mühendislik gibi.

1. Tek Değişkenli Lineer Regresyon Örneği:

Tek değişkenli lineer regresyon, bir bağımlı değişkenin (genellikle y olarak gösterilir) tek bir bağımsız değişken tarafından (genellikle x olarak gösterilir) açıklanmaya çalışıldığı durumlarda kullanılır.

Örneğin, bir evin fiyatını belirlemek istediğinizi düşünün. Ev fiyatı (y), evin büyüklüğüne (x) bağlı olabilir. Bu durumda, evin büyüklüğü (x) tek değişkenli lineer regresyon analizi için bağımsız değişken olur.

Örneğin, aşağıdaki veri seti, evlerin büyüklüğünü (x) ve fiyatlarını (y) içeriyor olabilir:

EvBüyüklüğü(x):[100,150,200,250,300]Fiyat (y):[200000,250000,300000,350000,400000]
Ev Büyüklüğü (x)Fiyat (y)
:[100,150,200,250,300]:[200000,250000,300000,350000,400000]

Bu veri setine tek değişkenli lineer regresyon uygulandığında, ev büyüklüğü ile fiyat arasındaki ilişkiyi açıklamak için bir doğru elde edilir.

2. Çok Değişkenli Lineer Regresyon Örneği:

Çok değişkenli lineer regresyon, bir bağımlı değişkenin birden fazla bağımsız değişken tarafından açıklanmaya çalışıldığı durumlarda kullanılır.

Örneğin, bir şirketin günlük satışlarını tahmin etmek istediğinizi düşünün. Günlük satışlar (y), reklam harcamaları (x1), mağaza büyüklüğü (x2) ve mevsim (x3) gibi birden fazla faktöre bağlı olabilir.

Bu durumda, reklam harcamaları (x1), mağaza büyüklüğü (x2) ve mevsim (x3) çok değişkenli lineer regresyon analizi için bağımsız değişkenler olur.

Örneğin, aşağıdaki gibi bir veri seti olabilir:

Reklam Harcamaları (x1):[1000,1500,2000,2500,3000]Mağaza Büyüklüğü (x2):[500,600,700,800,900]Mevsim (x3):[1,2,3,4,1]Günlük Satışlar (y):[20000,25000,30000,35000,40000]
Reklam Harcamaları (x1)Mağaza Büyüklüğü (x2)Mevsim (x3)Günlük Satışlar (y)
:[1000,1500,2000,2500,3000]:[500,600,700,800,900]:[1,2,3,4,1]:[20000,25000,30000,35000,40000]
Bu veri setine çok değişkenli lineer regresyon uygulandığında, reklam harcamaları, mağaza büyüklüğü ve mevsimin günlük satışlar üzerindeki etkilerini açıklamak için bir model elde edilir.

```
import numpy as np

import matplotlib.pyplot as plt

from sklearn.linear_model import LinearRegression


# Örnek veri setini oluşturma
X = np.array([1, 2, 3, 4, 5]).reshape(-1, 1) # Bağımsız değişken (x)
y = np.array([2, 3, 4, 5, 6]) # Bağımlı değişken (y)


# Lineer regresyon modelini oluşturma ve eğitme
model = LinearRegression()

model.fit(X, y)


# Eğim (slope) ve kesişim (intercept) katsayılarını alınması
slope = model.coef_[0]
intercept = model.intercept_

print("Eğim (Slope):", slope)
print("Kesişim (Intercept):", intercept)


# Veri seti ve regresyon doğrusunu görselleştirme
plt.scatter(X, y, color='blue', label='Veri Noktaları') # Veri noktalarını çizme
plt.plot(X, model.predict(X), color='red', label='Regresyon Doğrusu') # Regresyon
doğrusunu çizme

plt.xlabel('X')
plt.ylabel('Y')
plt.title('Lineer Regresyon Örneği')
plt.legend()
plt.show()
```

Bu kod örneğinde, bir lineer regresyon modeli, örnek bir veri seti üzerinde eğitilir. Veri seti, X ve y olarak adlandırılan bağımsız ve bağımlı değişkenleri içerir. Model, veri setine uygun bir regresyon doğrusu oluşturmak için eğitilir. Son olarak, veri seti ve regresyon doğrusu matplotlib kütüphanesi kullanılarak görselleştirilir.