grafik, grafik tasarım, metin, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu 

**DERS:** YAPAY ÖĞRENMENİN TEMELLERİ

**AD:** Sude Beyza Çifci Hafize Cemile Öğüt

**NO:** B210109064                  B210109058

**Facebook İçerik Etkileşimi İle Tahmin Modeli**

Veri seti incelendiğinde 18 özellik ve 1 etiket değerimden oluşan 19 sütun vardır. Type sütunu dışında bütün değerlerim sayısal. Like share gibi bazı özelliklerimin sütununda boş değerler var. Adım adım veri setime uygun model oluşturma aşamaları gerçekleştiriyoruz.

**1 Giriş**

**Lineer Regresyon Nedir?**

Bağımsız bir değişken ile bağımlı değişken arasında lineer bir ilişki olduğunda bağımsız değişken yardımı ile bağımsız değişkeni tahmin etmektir.

Bağımlı değişken: Tahmin edilmek istenen değişkendir.

Bağımsız değişken: Girdilerdir.

**Çoklu Lineer Regresyon Nedir?**

Çoklu regresyonda birden fazla bağımsız değişken ( x1, x2, ………, x n )  ile bir bağımlı değişken ( y ) arasındaki ilişki incelenmektedir.

**2 Metodoloji**

**Gradyan İniş:**

Gradyan iniş,1 bir fonksiyonun minimumunu bulmak için kullanılan yinelemeli bir optimizasyon algoritmasıdır. Doğrusal regresyon bağlamında, maliyet fonksiyonunu (ortalama kare hata) en aza indirmek ve verilerimiz için en uygun doğruyu bulmak için gradyan inişini kullanırız.

diyagram, taslak, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. Ağırlıkları Başlat
2. Tahmin için model yapısı oluştur
3. Maliyet Hesabı yap
4. Ağırlıkları hataya göre güncelle
5. Doğruya yakın sonuca ulaşana kadar 2 den 4 e kadar olan adımları tekrarla

metin, el yazısı, kağıt, doküman, belge içeren bir resim

Açıklama otomatik olarak oluşturuldu**Matematiksel Model**

**1-Parametreleri Başlatma:**

* Ağırlıkları (w) rastgele başlatma: w=rastgele değer ata
* Bias'ı (b) sıfırla: b=0

**2- İleri Yayılım :**

* İleri yayılım, giriş verilerini kullanarak tahmin değerlerini hesaplar.
* Formül: z=w⋅X+b
* z: Tahmin değerleri
* w: Ağırlıklar
* X: Giriş verileri
* b: Bias (kesim noktası)

**3-Geriye Yayılım (Backward Propagation):**

* Geriye yayılım, maliyet fonksiyonunun gradyanını hesaplar.
* Formüller:
* Dz=1/m (z−y)
* *w*=dz⋅*XT*
* db=∑dz

***4- Maliyet Fonksiyonu (Cost Function):***

* *Formül: J=1/2m ∑(z−y)2*
* *J: Maliyet*
* *m: Veri noktalarının sayısı*
* *z: Tahmin değerleri*
* *y: Gerçek değerler*

**5-Gradient İniş Güncelleme:**

* Ağırlıkları ve bias'ı günceller.
* Formüller:
* w=w−learning rate  × dw
* b=b−learning rate  ×  db

**6-Hata Metrikleri:**

* Ortalama Mutlak Hata (Mean Absolute Error - MAE):
* Formül: MAE=1/m ∑∣z−y∣
* Toplam Kare Hata (Sum of Squared Error - SSE):
* Formül: SSE=∑(z−y)2

**7-Eğitim:**

* Belirtilen epoch sayısı boyunca tekrarla:
* Eğitim verisi üzerinde ileri yayılım, maliyet hesaplama ve geriye yayılımı gerçekleştir.
* Her 150 epoch'ta eğitim maliyetini kaydet.
* Test verisi üzerinde ileri yayılım ve maliyet hesaplama.
* Her 150 epoch'ta test maliyetini kaydet.
* Sonuçları Görselleştirme:
* Eğitim maliyetini grafiğe çiz.
* Her 150 epoch'ta gerçek ve tahmin edilen değerlerin grafiğini çiz.

**3)Bulgular ve Tartışma**

   Bağımsız Değişken X :'Category','Page total likes', 'Post Month', 'Post Hour', 'Post Weekday', 'Paid'

    Bağımlı Değişken Y: ' Total Interactions'

Facebook veri setimizdeki özelliklerimizden yukarıdaki seçtiğimiz özellikler bizim modelimizin X özellikleridir.

Yukarıdaki veriler kullanarak modelimizi eğittik.

Amaç: Total Interactions değerini test verilerinde doğruya yakın tahmin etmek.

metin, ekran görüntüsü, diyagram, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Eğitim Maliyeti - İterasyon Grafiği:**

Yukarıdaki grafikte görüldüğü üzere iterasyon sayımım artmasına paralel olarak maliyetim düşmüştür.

Buna dayanarak modelimizin öğrenme sürecinde iyi bir iş çıkardığını düşünüyoruz.

metin, ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Eğtim Verileri İçin Gerçek ve Tahmin Değerler Grafiği:**

 Eğitim setindeki veri noktalarında tahmin değerlerimin gerçek değerlerime oldukça yakın olduğunu gözlemledik.

Buna dayanarak modelimizin eğitim veri setinde başarılı olduğu çıkarımını yaptık.

metin, ekran görüntüsü, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Test Veri Seti Gerçek ve Tahmin Değerler Grafiği:**

Test veri setindeki veri noktalarına bakıldığında, tahmin edilen değerlerin genellikle gerçek değerlere yakın olduğunu gözlemledik. Ancak, eğitim setinde elde edilen performansa kıyasla, modelin test setindeki performansının biraz daha düşük olduğunu gördük.

Bunun sebebi ise modelin belki de eğitim setine aşırı uyum sağlamasıdır. Ancak  daha iyi genelleme yapabilmesi için girdi özelliklerinin güncellenmesini düşünüyoruz. Test veri sayımızı artırmak da bir çözüm olabilir. Öğrenme katsayısında değişikliklerde yapılabilir. Hataları minimize etmek için ise normalizasyona başvurabiliriz.

Girdi özelliklerini değiştirip noramilizasyon uyguladığımda elde ettiğim sonuçlar:

metin, ekran görüntüsü, ekran, görüntüleme, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

**ekran görüntüsü, metin, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**ekran görüntüsü, metin, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**4 Sonuç**

Yukarıda bahsedilen durumlar değerlendirerek aldığımızda farklı çıktılar ile karşılaştık. Tamamı kodumuzun olduğu dosyada bulunmaktadır. Öğrenme katsayısındaki değişiklikler, epoch sayısı ve özellik olarak alacağım sütunlar sonucu büyük oranda etkilemektedir. Sonuç olarak, modelin eğitim setinde iyi performans sergilediği test setinde de yeterli başarıyı sağladığı sonucuna vardık. Fakat Eğitim setine göre sağladığı başarıya oranladığımızda biraz daha iyileştirmeye gittik. Modelimizin özelliklerini değiştirdiğimizde ise biraz daha iyi bir performans sağladığını tespit ettik. Normalizasyon işlemi sonucunda modelim daha iyi bir performans sergiledi.

**Kaynakça**

<https://medium.com/international-school-of-ai-data-science/gradient-descent-multi-linear-regression-ml-from-scratch-3-8773421be2ce>

<https://stackoverflow.com/questions/65539013/how-to-plot-a-graph-of-actual-vs-predict-values-in>

<https://youtu.be/ALpO0iTnR40?si=jHDXJCZOGXavIkT5>

metin, el yazısı, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

[GitHub - MrGranddy/Yapay-Zeka-Kulubu-Teorik-ML: Yapay Zeka Kulübü 2023 Teorik Machine Learning Eğitim Grubu için bir repo](https://github.com/MrGranddy/Yapay-Zeka-Kulubu-Teorik-ML)

<https://github.com/learn-co-students/multivariable-calculus-gradient-descent-ds-apply-000/blob/master/gradientdescent.png>

[Building Linear Regression Model from Scratch (No Libraries) | by Jordan | Medium](https://medium.com/@jordans2299/building-linear-regression-model-from-scratch-no-libraries-1d75782aa617)

[Machine-Learning-Algorithms/Multiple Linear Regression/Multiple Linear Regression from Scratch.ipynb at main · nishant-sg/Machine-Learning-Algorithms · GitHub](https://github.com/nishant-sg/Machine-Learning-Algorithms/blob/main/Multiple%20Linear%20Regression/Multiple%20Linear%20Regression%20from%20Scratch.ipynb)

Beyza Çifci & Hafize Öğüt Github Proje Linki:

 https://github.com/hafizeogut/Linear\_Regression\_Algortima