

İncelenen Problemin Başlığı: Instagram İçerik Metriklerinin Toplam Erişim (Etki)
Üzerindeki Etkisinin Makine Öğrenmesi ile Tahmini **Ders Adı:** Yapay Zeka ve Makine
Öğrenimine Giriş (YZMÖ Giriş) **Okul:** Kocaeli Üniversitesi (KOÜ)

Öğrenci Adı-Soyadı: Talha Özgür **Öğrenci Numarası:** 250121009

1. Giriş ve Problem Tanımı

Çözülmek istenen problem:Sosyal medyanın önemli bir parçası olan “Instagram”ın etkileşimi; beğeni, yorum, kaydetme, paylaşma, hashtag kullanımı ve keşfet trafigi gibi özelliklerin etkileşim üzerinde ne kadar ağırlığının ne ölçüde olduğunu tahmin edilmesi.

Çalışmanın amacı: Bu çalışmanın temel amacı, gerçek hayat verileri içeren bir veri seti üzerinden Instagram içerik metriklerini analiz ederek, gelecekte paylaşılacak içeriklerin potansiyel erişim gücünü öngörebilecek bir **Denetimli Öğrenme (Regresyon)** modeli geliştirmektir.Bu model sayesinde hangi etkileşim türünün ağırlığının daha önemli olduğunu matematiksel şekilde tahmin edicektir.

2.Veri seti ve Yöntem

Veri setinin tanıtımı:Bu çalışmada sosyal medya içeriklerine ait performans metriklerini içeren "Instagram Etkileşim Veri Seti" kullanılmıştır.Bu veri seti gönderilerin aldığı beğeni, yorum, kaydetme, paylaşma, profil ziyareti gibi etkileşim sayılarını ve bu gönderilerin ulaştığı toplam kişi sayısını (Etki) içermektedir.Bu veri setini “Instagram.csv” şeklinde yeniden isimlendirip içerisindeki metrikleri Python ve Pandas kullanarak hedefim için şekillendirilmiştir.

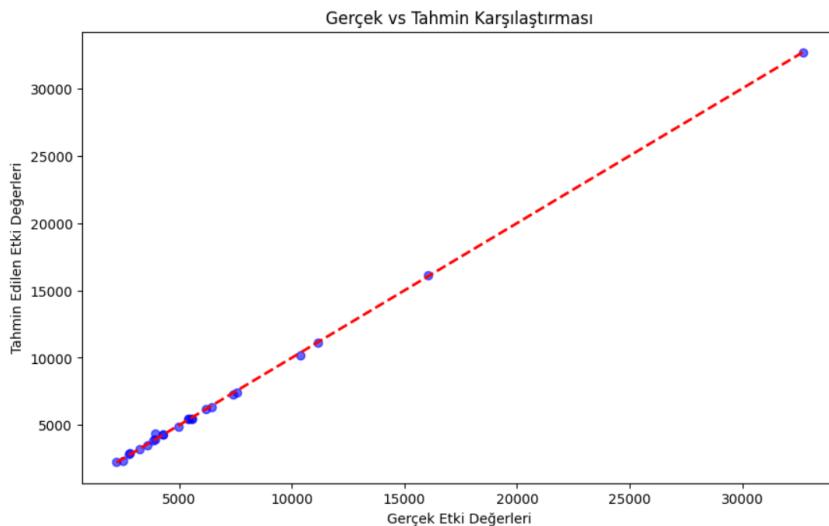
Kullanılan Algoritma ve Yöntem:Problemim tipi sayısal veriler içerir.Bu yüzden Denetimli öğrenme yöntemlerinden olan Lineer Regresyon algoritmasını kullandım.Lineer regresyonu kullanırken modelimi %80 Eğitim %20 Test şeklinde tercih ettim.Böylelikle modelimin daha önce hiç görmediği değerler ile eğitilmiştir.

3.Deneysel Sonuçlar

Geliştirdğim modelde 2 tane performan metriği kullandım

Performans Metrikleri:

- R2 Skoru (Belirleme Katsayısı):** Modelin verideki değişkenliği açıklama oranını gösterir. Test seti üzerinde yapılan ölçümde R2 **Skoru:** (`0.99959404890355`) olarak hesaplanmıştır. Bu değer, modelin erişim sayılarını tahmin etmede %'lik oranında başarılı olduğunu gösteriyor.
- MAE (Ortalama Mutlak Hata):** Tahmin edilen değerler ile gerçek değerler arasındaki farkın ortalamasını ifade eder. Hesaplanan **MAE Değeri:** (`83.83232873523849`) şeklindedir. Bu sonuç, modelin bir gönderinin erişimini tahmin ederken ortalama +/- kişilik bir sapma ile çalıştığını gösterir.



. Sonuç ve Değerlendirme

Bu çalışmada, Instagram gönderilerinin etkileşim metrikleri kullanılarak toplam erişim sayılarını tahmin eden bir makine öğrenmesi modeli geliştirilmiştir. Elde ettiğim sonuçlar şunlardır:

Model katsayıları incelendiğinde `1.48` katsayısı ile "Yorum" değişkenidir.

Sonuç olarak içerik üreticilerinin sadece beğeni sayılarına değil algoritma tarafından daha değerli bulunan "Yorum" ve "Paylaşım" gibi metrikleri artırmaya yönelik stratejiler geliştirmesi gerektiği, bu çalışma ile veri bilimsel olarak desteklenmiştir

Bağlantılar ve Kaynak:

Veri seti kaynağı: Kaggle <https://www.kaggle.com/datasets/tahirmohd/my-datasat>

Bağlantı: <https://github.com/TalhaOzgur>