

MAKİNE ÖĞRENİMİ

LASSO REGRESYON İLE SATIŞ TAHMİNİ

1. Veri Yükleme ve Ön İşleme

- Veri Yükleme: train, feature ve test veri setlerini birleştirerek kullandık.
- **Store ve Dept Değişkenlerini Düzenleme:** "Store" ve "Dept" değişkenlerini one-hot encoding ile dönüştürdük.
- **Tarih Dönüşümü:** Tarih sütununu datetime formatına çevirdik.
- **Hafta ve Çeyrek Ekleme:** Tarih bilgisine dayalı olarak "Week" ve "Quarter" sütunlarını ekledik.

2. Eksik Verilerin Doldurulması

- **Boolean Dönüşümü:** Boolean sütunlarını sayısal değerlere çevirdik.
- **One-Hot Encoding:** "Week" ve "Quarter" sütunlarını one-hot encoding ile dönüştürdük.
- **Eksik Verilerin Doldurulması:** Markdown sütunlarındaki eksik değerleri 0 ile doldurduk, diğer eksik değerleri ortalama ile doldurduk.

LASSO REGRESYON İLE SATIŞ TAHMİNİ

3. Ölçeklendirme

- Min-Max Ölçeklendirme: Tüm değişkenleri 0-1 aralığına ölçeklendirdik, ancak "Weekly_Sales" sütununu hariç tuttuk.

4. Model Eğitimi

- **Veri Bölme:** Veriyi eğitim ve test setlerine ayırdık.
- **Lasso Regresyon:** L1 regularizasyonu ile Lasso regresyon modelini eğittik. Lasso, modelin karmaşıklığını azaltarak aşırı uyumlamayı önlemek ve özellik seçimi yapmak için kullanıldı.

5. Model Değerlendirme

- Tahmin Yapma: Test setinde tahminler yaptık.
- Performans Ölçümleri: RMSE (Kök Ortalama Kare Hatası) ve R-squared (R^2) değerlerini hesapladık.
- Sonuçların Gösterimi: İlk 10 tahmin ve gerçek değerleri karşılaştırdık.

Özellikler:

Size, IsHoliday, Temperature, Fuel_Price, MarkDown1, MarkDown2, MarkDown3, MarkDown4, MarkDown5, CPI, Unemployment,

One-hot encoding ile eklenenler:

Type_B, Type_C, Week-of-month_1, Week-of-month_2, Week-of-month_3, Week-of-month_4, Quarter_1, Quarter_2, Quarter_3, Quarter_4.

Lasso ile Satış Tahmini Sonuçları

```
Root Mean Squared Error (RMSE): 8881.6835957517  
R-squared: 0.6659621895290404
```