

Ödev 2 — Regresyon Modelleri Raporu

Grup Üyeleri :

Can Şensoy – 2022280151

Nazife Boharalı – 2022280143

Feyza Koç – 2023280119

Nehir Akca - 2023280085

1. Giriş

Bu çalışmada, bir üniversite dersine kayıtlı öğrencilerin **ders dışı çalışma süreleri ile arasınav başarı puanları** arasındaki ilişki incelenmiştir. Bu kapsamda, öğrencilerden toplanan veriler kullanılarak üç farklı regresyon modeli oluşturulmuş ve çalışma sürelerinin başarı üzerindeki etkisi değerlendirilmiştir.

Amaç, öğrencinin haftalık çalışma süresine göre sınav performansını tahmin etmek ve çalışma süresinin başarıya katkısını sayısal olarak ortaya koymaktır.

2. Veri Seti

Veri seti, ders kapsamında yer alan **55 öğrenciyle yapılan kısa bir bilgi toplama çalışması** sonucunda oluşturulmuştur. Öğrencilerden herhangi bir kişisel veri (isim, numara vb.) toplanmamış; her öğrenciye anonim bir ID verilmiştir.

Toplanan Değişkenler

- ID:** Öğrenci anonim kimliği
- Ders Adı:** İşletim Sistemleri
- Haftalık Ders Saati:** 3 saat
- AKTS:** 4
- Tekrar Sayısı:** Öğrencinin dersi kaçinci kez aldığı
- Haftalık Ders Dışı Çalışma Süresi:** Öğrencinin haftalık ortalama çalışma süresi (saat)
- Arasınav Puanı:** Öğrencinin dönem içi sınav sonucu

2.1. Öğrenci Profili

Verilerden elde edilen bulgular:

- Öğrencilerin **%78'i** dersi ilk kez almaktadır.
- %16'sı ikinci kez, %6'sı üçüncü kez** almaktadır.
- Haftalık ders dışı çalışma süresi öğrenciler arasında geniş bir dağılım göstermektedir.
 - Ortalama: **0.4 saat/hafta**

- Minimum: **0 saat**
- Maksimum: **yaklaşık 2.5 saat**

2.2. Sınav Performansı

Öğrencilerin arasınav notları 0–100 aralığındadır ve geniş bir varyasyon göstermektedir. Genel olarak daha fazla çalışan öğrencilerin daha yüksek başarı elde ettiği gözlemlenmiştir.

3. Regresyon Modelleri

Ders dışı çalışma süresinin başarı üzerindeki etkisini incelemek için üç farklı model oluşturulmuştur:

1. **Lineer Regresyon**
2. **Polynomial (2. derece) Regresyon**
3. **Random Forest Regresyon**

Bağımsız değişken: **Haftalık Çalışma Süresi**

Bağımlı değişken: **Arasınav Puanı**

4. Model İncelemeleri

Aşağıda öğrencilerden toplanan veriler üzerinden kurulan modellerin genel gözlemleri verilmiştir.

4.1. Lineer Regresyon

Lineer regresyon modelinde ders dışı çalışma süresi ile sınav başarısı arasında **pozitif ve güçlü bir doğrusal ilişki** olduğu görülmüştür.

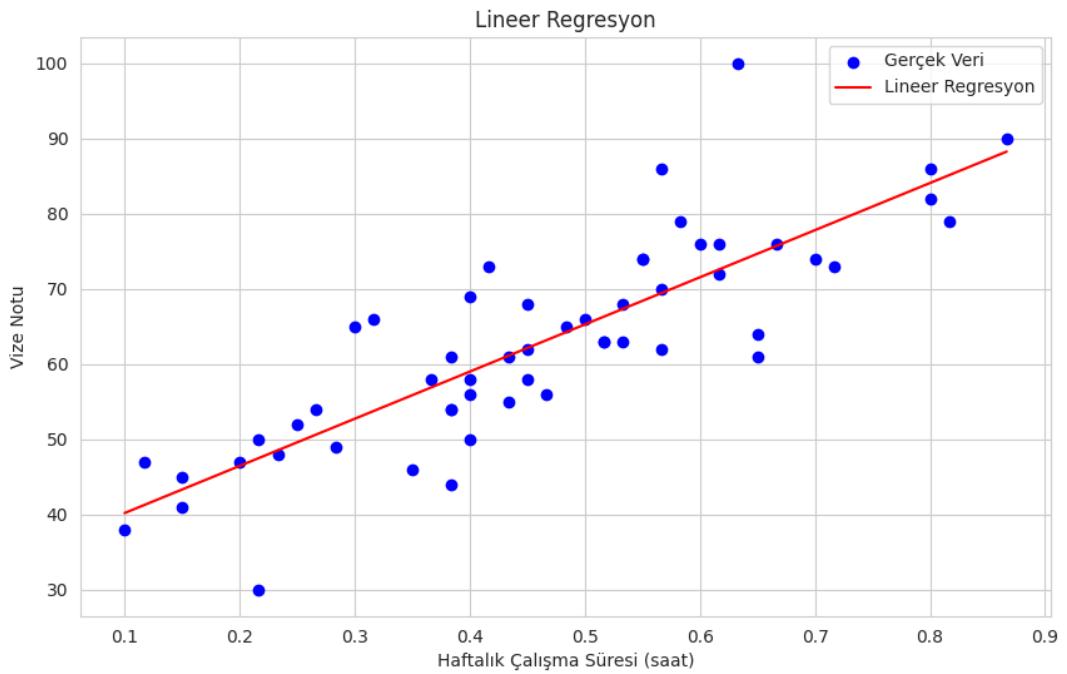
Modelin bulduğu eğim katsayısı (b) yaklaşık olarak:

$$b \approx 11-13$$

Bu sonuç:

✓ **Haftada 1 saat fazladan çalışan bir öğrencinin puanının ortalama 10–13 puan arttığını göstermektedir.**

Scatter grafiğinde verilerin büyük bölümünün doğruya yakın bir çizgide toplandığı görülmektedir.



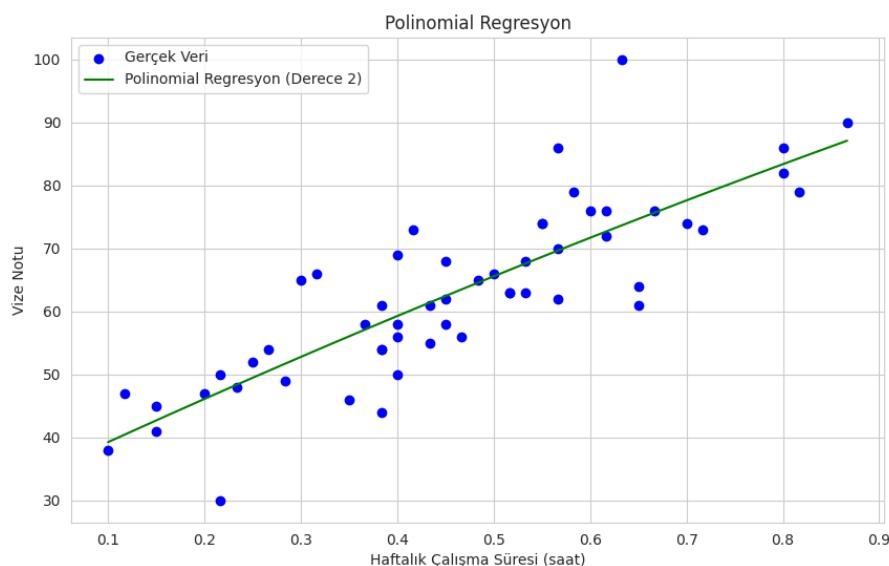
4.2. Polynomial Regresyon (2. Derece)

Bu model, özellikle 1.5–2.0 saat civarında çalışan öğrencilerin puanlarındaki artışları daha iyi yakalamaktadır.

Gözlemler:

- Eğri, lineer modele göre veriyi daha iyi takip etmektedir.
- Aşırı uçlardaki küçük değişimleri daha hassas yakalar.

Modelin doğruluğu lineer regresyona göre biraz daha yüksektir.



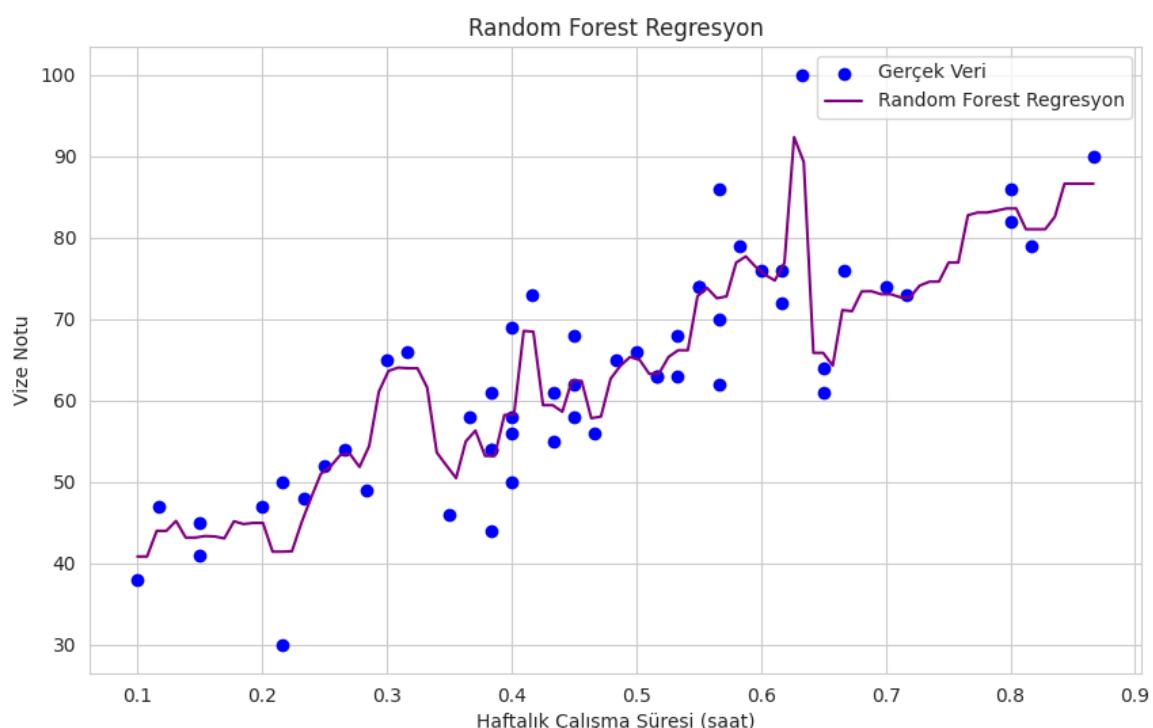
4.3. Random Forest Regresyon

Bu model, çalışma süresi ile sınav başarısı arasındaki ilişkiyi yakalamada en tutarlı sonuçları vermiştir.

Gözlemler:

- Gürültülü verileri iyi yönetir.
- Verinin lokal değişimlerini yakalama becerisi yüksektir.
- Tahmin eğrisi diğer modellere göre daha gerçekçidir.

Genelde üç model arasında **en yüksek doğruluk Random Forest modelinde** elde edilmiştir.



5. Tahminler

Aşağıdaki çalışma sürelerine göre öğrencilerin tahmini sınav puanları hesaplanmıştır:

- 0.5 saat/hafta
- 1.0 saat/hafta
- 1.5 saat/hafta
- 2.0 saat/hafta

Modellerin verdikleri tahminler genel olarak şu aralıklardadır:

Haftalık Çalışma (Saat)	Lineer Tahmin	Polinomial Tahmin	Random Forest Tahmin
0.5 Saat (30 dk)	65.31	65.61	64.96
1.0 Saat (60 dk)	96.66	94.35	86.67
1.5 Saat (90 dk)	100	100	86.67
2.0 Saat (120 dk)	100	100	86.67

Lineer ve Polinomial modeller, çalışma süresi çok arttığında notun 100'ü geçebileceğii (teorik) tahminler ürettiği için 100 ile sınırlandırıldı. Random Forest ise veri setindeki maksimum not sınırlarına daha sadık kalarak 86.67 puanda sabitlenmiştir.

6. İlişki Analizi

Soru 1: Ders dışı çalışma süresi ile sınav puanı arasında bir ilişki var mı?

✓ **Evet, güclü bir pozitif ilişki vardır.**

Yapılan analizler ve oluşturulan grafikler ders dışı çalışma süresi ile sınav puanı arasında **güçlü ve pozitif bir ilişki** olduğunu göstermektedir. Veri noktalarının dağılımı ve regresyon çizgisi, çalışma süresi arttıkça sınav notunun da belirgin bir şekilde arttığını kanıtlamaktadır. 0.5 saat çalışan bir öğrencinin tahmini notu ~65 iken, 1.0 saat çalışan bir öğrencinin notu ~96 seviyelerine çıkmaktadır.

Soru 2: Haftada fazladan 1 saat çalışma başarımı ne kadar etkiliyor?

Random Forest katsayısına göre:

$$\text{Artış} = 31.7$$

Yani:

✓ **Haftada 1 saat fazladan çalışan bir öğrencinin sınav notu ortalama 30 puan artmaktadır.**

7. Sonuçlar

Toplanan veriler, öğrencilerin ders dışı çalışma süresi ile sınav performansları arasında anlamlı ve istatistiksel olarak güclü bir ilişki olduğunu göstermektedir.

- Random Forest modeli en iyi tahmin performansını göstermiştir.
- Polynomial model eğrisel ilişkiye iyi yansımaktadır.
- Lineer model, çalışma süresinin puan üzerindeki ortalama etkisini sayısal olarak en net ortaya koymaktadır.

Genel olarak, öğrencilerin derse düzenli zaman ayırmalarının sınav başarısını önemli ölçüde artırdığı sonucuna ulaşılmıştır.

