

DERS İZLENESİ (SYLLABUS): VERİ BİLİMİ VE UYGULAMALARI

Dersin Adı: Veri Bilimi ve Uygulamaları

Öğretim Elemanı: Dr. Ömer Gökdaş (Geomatik Müh. PhD)

1. Dersin Tanımı

Bu ders; ham veriyi işleyerek anlamlı bilgiye ve katma değere dönüştürme sanatı olan Veri Bilimi'ne (Data Science) uygulamalı bir giriş sağlar. Öğrenciler; karmaşık matematiksel formüllerle boğuşmak yerine, Python ekosistemindeki güçlü kütüphaneleri (Pandas, Scikit-learn) kullanarak gerçek hayat verilerini analiz etmeyi, temizlemeyi ve geleceğe yönelik temel tahminler (Makine Öğrenmesi) yapmayı öğreneceklerdir.

2. Öğrenme Çıktıları

Bu dersi başarıyla tamamlayan öğrenciler:

- Python ile (NumPy/Pandas) büyük veri setlerini yükler, filtreler ve manipüle eder.
- Gerçek dünyadaki "kirli" verileri (Eksik, Hatalı, Tekrarlı) tespit eder ve temizler.
- Keşifçi Veri Analizi (EDA) ile verinin hikayesini ortaya çıkarır.
- Matplotlib ve Seaborn kullanarak veriyi profesyonel grafiklerle sunar.
- Temel Makine Öğrenmesi modellerini (Regresyon, Sınıflandırma) Scikit-learn aracılığıyla kurar.
- Bir veri bilimi projesini uçtan uca (Veri -> Analiz -> Model -> Karar) yönetir.

3. Haftalık Ders Planı

Hafta	Konu Başlığı	İçerik ve Uygulama
1	Veri Bilimine Giriş	Veri Bilimci ne iş yapar? Yapay Zeka vs. Veri Bilimi. Google Colab ile tanışma (Kurulumuz Python).
2	Sayısal İşlemler: NumPy	Listeler neden yetersiz? NumPy Dizileri (Arrays). Vektörizasyon: Döngü kullanmadan hızla işlem yapma.
3	Veri Manipülasyonu - I (Pandas)	"Excel'in Python hali": DataFrame ve Series yapıları. CSV/Excel dosyalarını okuma ve özetleme.
4	Veri Manipülasyonu - II (Filtreleme)	Veri madenciliği: <code>loc</code> ve <code>iloc</code> ile veriyi dilimleme. Koşullu sorgular ("Satışları 1000'den büyük olanlar").
5	Veri Temizleme (Data Cleaning)	Eksik verileri (NaN) doldurma stratejileri. Aykırı değerleri ve tekrarlayan satırları temizleme.
6	Veri Birleştirme	Farklı tabloları birleştirme (Merge, Join). Pivot tablolar ile raporlama.

Hafta	Konu Başlığı	İçerik ve Uygulama
7	Keşifçi Veri Analizi (EDA)	Veriyi anlamak: Değişkenler arası ilişkiler. Korelasyon Matrisi: Hangi özellikler birbirini etkiliyor?
8	ARA SINAV (VİZE)	Kod okuma, kütüphane mantığı, veri manipülasyonu fonksiyonlarının kağıt üzerinde uygulanması.
9	Veri Görselleştirme - I	Matplotlib: Temel grafikler (Çizgi, Sütun). Bir mühendis veriyi nasıl sunmalı?
10	Veri Görselleştirme - II	Seaborn: İstatistiksel grafikler. Isı Haritaları (Heatmap) ve Boxplot ile veri dağılımını görme.
11	Makine Öğrenmesine Giriş	Bilgisayar nasıl öğrenir? Gözetimli vs. Gözetimsiz Öğrenme. Train/Test ayrımı nedir?
12	Geleceği Tahmin: Regresyon	Scikit-learn ile model kurma. Ev fiyatı tahmini gibi senaryolar üzerinden Doğrusal Regresyon.
13	Sınıflandırma (Classification)	"Spam mi değil mi?" Karar Ağaçları ve K-NN algoritmalarının mantığı ve kodlanması.
14	Dönem Sonu Projesi	Öğrencilerin seçtikleri (Kaggle vb.) bir veri seti üzerinde; Temizleme, Görselleştirme ve Modelleme adımlarını içeren sunumu.

4. Yazılım ve Araçlar

- **Platform:** Google Colab (Önerilen) veya Jupyter Notebook.
- **Kütüphaneler:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn.

5. Sınav Politikası (Cheat Sheet)

- **Format:** Sınavlar ezberden uzak, problem çözme odaklıdır.
- **Kopya Kağıdı:** Öğrenciler sınavlara, kod sözdizimlerini (syntax) içeren 1 adet A4 kağıdı (arkalı-önlü) ile girebilirler.

6. Kaynaklar (WEB)

1. **Veri Setleri:** Projeler için [kaggle.com/datasets](https://www.kaggle.com/datasets) kullanılacaktır.