

Ara sınav projesi

Veri Madenciliğinde İleri Konular – Yüksek Lisans – 2025-2026 – Güz

Teslim şekli

- Ödevinizi yalnızca LMS sistemine **tek bir .zip dosyası** olarak yükleyin.
- Zip dosyası adı: **öğrencinumarası_ad_soyad.zip**
- Zip dosyasının içinde şu dosyalar bulunmalıdır:
 - Jupyter Notebook proje dosyanız (.ipynb uzantılı), yorumlar, Python kodu ve sonuçlar dahil. Dosya adı: **öğrencinumarası_ad_soyad.ipynb**
 - Aynı notebook dosyasının HTML çıktısı. Dosya adı: **öğrencinumarası_ad_soyad.html**

Veri seti

kaggle.com adresinden **“1000 drugs and side effects”** adlı CSV dosyasını indirin.
Bağlantı: <https://www.kaggle.com/datasets/palakjain9/1000-drugs-and-side-effects>

Sorular

1. Tüm sayısal sütunlar için temel istatistikleri hesaplayın: **mean, std, min, %25, median, %75 ve max.**
2. Kategorik sütunlardaki verileri **her sütun için ayrı ayrı** olacak şekilde **pie chart** (pasta grafiği) ile görselleştirin.
3. Condition sütunu hastalık sebebinde içermektedir. Bu sebeplere karşılık gelen **tedavi süresi (Treatment_Duration_days)** dağılımını tek bir **boxplot** üzerinde görselleştirin.
4. Drug_Name ve Side_Effects sütunları için **2 boyutlu histogram** (veya uygun bir 2D frekans görselleştirmesi) çizdirin ve **sonuçları yorumlayın.**
5. Bir önceki sorudaki (4. madde) analizde kullandığınız veriler için, seaborn kütüphanesinden **clustermap** fonksiyonunu kullanarak **hiyerarşik kümeleme** yapın. Elde ettiğiniz grafiği yorumlayın ve **yorumlarınızı markdown hücrelerinde** yazın.
6. **Yaş (Age)** ve **Improvement_Score** değişkenleri arasındaki **korelasyonu** hesaplayın ve sonucu yorumlayın.

7. Erkekler ve kadınlar için **hastalık (Condition)** oranlarını normalize edilmiş bir şekilde hesaplayın. (Örneğin her cinsiyet için toplam 1 olacak şekilde olasılık dağılımı çıkarabilirsiniz.)
8. Hangi ilaçlar için **baş dönmesi (Dizziness)** yan etkisi raporlanmıştır? Bu ilaçları **azalan sıklığa göre** sıralayın.
9. **Metoprolol** ilacı için, diğer yan etkilere göre **baş dönmesi (Dizziness)** yan etkisinin **olasılığını** hesaplayın (Metoprolol için raporlanan tüm yan etkiler içinde Dizziness'in oranı).

Tüm soruların puanı eşittir.

Başarılar

Dr. Sajjad NEMATZADEH