

Hospital Readmission Prediction (30-Day Readmission Risk)

Makine öğrenmesi kullanarak **taburcu edilen hastaların 30 gün içinde tekrar hastaneye yatış yapma olasılığını tahmin eden** bir yapay zeka projesidir.

Bu sistem, doktorlara objektif bir **erken uyarı risk skoru** sunarak hem hasta güvenliğini artırmayı hem de hastane maliyetlerini düşürmeyi hedefler.

Proje Amacı

Hastanelerde bazı hastalar taburcu olduktan kısa süre sonra yeniden yatış yapmak zorunda kalır. Bu durum:

- Servis ve yoğun bakımlarda kapasite baskısı oluşturur
- Tedavi maliyetlerini artırır
- Hastanın sağlık durumunu olumsuz etkiler
- Hastaneler için kritik kalite metriği olan *readmission rate* değerini yükseltir

Bu proje, geçmiş verilere dayalı bir **makine öğrenimi modeli ile tekrar yatış riskini önceden tahmin etmeyi** amaçlar.

Problem Tanımı

Doktorların tek tek tüm hastalar için risk değerlendirmesi yapması mümkün değildir çünkü:

- Hastaların tıbbi geçmişi çok karmaşıktır
- Kronik hastalıklar farklı seviyededir
- Evde tedavi süreci kontrol edilemez
- Sosyal/çevresel faktörler değişkendir

Bu yüzden veri temelli bir model, doktorlar için güçlü bir karar destek sistemi sağlar.

Başarı Kriterleri

Model başarı değerlendirmesinde kullanılan metrikler:

- **ROC-AUC ≥ 0.70**
- **PR-AUC** (dengesiz veri setleri için ideal)
- Yanlış negatiflerin azaltılması (yüksek riskli hastaları kaçırılmamak)
- SMOTE ile veri dengelenmesi
- SHAP ile model açıklanabilirliği

Kullanılan Veri

Kaynak: UCI Machine Learning Repository – Diabetic Readmission Dataset

Boyut: ~100.000 hasta kaydı

Öne çıkan değişkenler:

- Yaş grubu
 - Cinsiyet
 - Önceki yatış sayısı
 - Acil servis ziyaret sayısı
 - Laboratuvar test sayısı
 - İlaç değişimi
 - Sigorta ve taburcu şekli
 - **Target:** readmitted (Yes/No)
-

Teknik Yaklaşım

1 Veri Hazırlama & EDA

- Eksik veri temizleme
- Aykırı değer işleme (IQR & winsorizing)
- Korelasyon analizi
- Kategorik değişken dönüştürme
- Feature engineering (ilaç sayısı vb.)

2 Dengesiz Veri Çözümü

Dataset dengesizdir:

- **%11 readmitted (Yes)**
- **%89 readmitted (No)**

Çözüm: **SMOTE Oversampling**

3 Modelleme

Kullanılan model:

✓ **XGBoost Classifier**

En iyi sonuçlar:

- **ROC-AUC: 0.74**
- **PR-AUC: 0.32**

SHAP Açıklanabilirlik

Modelin hangi özelliğe nasıl etki ettiğini açıklamak için:

- SHAP Summary Plot
- Feature Importance
- Lokal açıklama mantığı

Bu sayede doktor ‘neden?’ sorusuna yanıt bulabilir.

API (FastAPI Servisi)

Model, dış sistemler tarafından kullanılabilmesi için **FastAPI ile servis haline getirilmiştir.**

Çalıştırma:

uvicorn app:app --reload

Swagger arayüzü:

 <http://127.0.0.1:8000/docs>

Streamlit Arayüzü (Kullanıcı Paneli)

Hastane personeli ve doktorlar için kullanıcı dostu bir arayüz geliştirilmiştir.

Çalıştırma:

streamlit run ui.py

Arayüz üzerinden:

- Yaş
 - Laboratuvar testleri
 - Acil ziyaret sayısı
 - İlaç sayısı
 - Yatış geçmişi
- gibi bilgiler girilerek **30 günlük tekrar yatış riski** hesaplanır.
-

Proje Dosya Yapısı

project/

|

|— data/

|

|— diabetic_data.csv

```
|  
|   └── models/  
|       └── readmission_xgb.pkl  
|  
|  
└── app.py  
└── main.py  
└── ui.py  
└── shap_summary.png  
└── README.md
```

Kurulum

1. Sanal ortam oluştur

```
python -m venv .venv
```

2. Ortamı aktif et

```
.venv\Scripts\activate
```

3. Gerekli kütüphaneleri yükle

```
pip install -r requirements.txt
```

Karşılaşılan Zorluklar & Çözümler

Zorluk	Çözüm
Eksik veriler yükseltti	Drop & medyan doldurma
Veri dengesizdi	SMOTE
Çok kategorik sütun vardı	One-hot / target encoding
Açıklanabilirlik gerekiyordu	SHAP uygulanması
API & UI entegrasyonu	FastAPI + Streamlit

Sonuç

Bu proje, sağlık sektöründe **hayati öneme sahip bir erken uyarı sistemi** sunar.
Doktorların taburcu işlemleri sırasında riskli hastaları hızlıca belirlemesine yardımcı olur.

Sonuç olarak:

- **Hastane maliyetleri azalır**
 - **Hasta güvenliği artar**
 - **Doktorlara karar destek mekanizması sağlanır**
-



Geliştirici

Busenur Durak
Yönetim Bilişim Sistemleri – İzmir Bakırçay Üniversitesi
AI & Data Analytics