

T.C. İSTANBUL TOPKAPI ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ YAZILIM
MÜHENDİSLİĞİ BÖLÜMÜ
FİNAL PROJESİ RAPORU

DERS ADI: Programlama Dilleri

DÖNEM: Güz 2025

PROJE ADI: Kişisel Hipertansiyon Takip ve Uyarı
Sistemi

Adı Soyadı: Şevval ARSLAN

Öğrenci Numarası: 22040301030

ÖZET

Bu projede, bireylerin hipertansiyon (HTN) ve diyabet (DM) risklerini değerlendirmek ve kişiselleştirilmiş sağlık önerileri sunmak amacıyla masaüstü tabanlı bir Karar Destek Sistemi geliştirilmiştir. Python programlama dili kullanılarak geliştirilen yazılım, Scikit-learn kütüphanesi ile Lojistik Regresyon algoritmasını temel alan bir makine öğrenmesi modeli kullanmaktadır.

Çalışmada hazır bir veri seti üzerinde; eksik veri tamamlama (imputation), ölçeklendirme (StandardScaler) ve sınıf dengesizliği giderme (Class Weighting) gibi veri ön işleme adımları uygulanmıştır. Kullanıcı dostu arayüz Tkinter ile tasarlanmış, sonuçlar Matplotlib grafik kütüphanesi ile görselleştirilmiş ve analiz çıktıları ReportLab ile PDF formatında raporlanabilir hale getirilmiştir. Sistem, veri tabanı gereksinimi olmadan CSV dosya yapısı ile çalışarak taşınabilir bir çözüm sunmaktadır.

İÇİNDEKİLER

- 1. GİRİŞ VE PROJENİN AMACI**
- 2. MATERİYAL VE VERİ SETİ**
- 3. YÖNTEM: VERİ ÖN İŞLEME VE
MODELLEME**
- 4. YAZILIM MİMARİSİ VE KULLANILAN
TEKNOLOJİLER**
- 5. SENARYO BAZLI RİSK ANALİZİ**
- 6. BULGULAR VE SİSTEMİN İŞLEYİŞİ**
- 7. SONUÇ VE DEĞERLENDİRME**
- 8. EKLER VE KAYNAKÇA**

1. GİRİŞ VE PROJENİN AMACI

1.1. Projenin Tanımı

Kişisel Hipertansiyon Takip ve Uyarı Sistemi: kullanıcının günlük tansiyon, kan şekeri, tuz tüketimi ve yaşı bilgileri gibi verilerini analiz ederek; kişiselleştirilmiş hipertansiyon riskini ve buna eşlik edebilecek diyabet eğilimini tespit eden, erken uyarılar sunarak kullanıcıya sağlık durumu hakkında bütüncül geri bildirim vermeyi amaçlayan bir programdır.

1.2. Projenin Amacı

Hipertansiyon ve diyabet, dünya genelinde en yaygın görülen ve erken teşhis edilmediğinde ciddi komplikasyonlara yol açan kronik hastalıklardır. Bu projenin temel amacı, kullanıcıların klinik verilerini (tansiyon, şeker, kolesterol vb.) ve yaşam tarzı alışkanlıklarını analiz ederek, olası risk durumlarını simüle eden ve farkındalık yaratan bir yazılım geliştirmektir.

Sistem sadece mevcut durumu analiz etmekle kalmayıp, "Hipertansiyon var ama Diyabet riski nedir?" gibi senaryo bazlı çapraz sorgulamalar yaparak kullanıcıya, veri madenciliği teknikleri ile desteklenmiş etkileşimli bir sağlık simülasyonu sunmaktadır.

2. MATERİYAL VE VERİ SETİ

Projede taşınabilirliği artırmak adına CSV (Comma Separated Values) dosya formatı tercih edilmiştir.

2.1. Veri Kaynağı ve İçeriği

Modelin eğitimi için Kaggle platformundan temin edilen hypertension_data.csv dosyası kullanılmıştır. Modelin eğitimi için kullanılan ham veri seti, bireylerin sosyo-ekonomik durumları ve stres seviyeleri dahil olmak üzere geniş kapsamlı veriler içermektedir. Ancak projenin odak noktası **tıbbi ve biyolojik risk analizi** olduğu için modele gürültü oluşturabilecek veya doğrudan tıbbi tanı ile ilişkisi zayıf olan sütunlar (Ülke, Eğitim Seviyesi, İş Durumu vb.) veri ön işleme aşamasında elenmiştir. Model eğitiminde aktif olarak kullanılan öznitelikler şunlardır:

- Klinik Veriler:** Sistolik ve Diyastolik Tansiyon, Glikoz, Vücut Kitle İndeksi (BMI), Toplam Kolesterol, LDL, HDL, Triglicerit, Nabız, Mevcut Hastalık Tanısı (Hipertansiyon Hastası mı, Diyabet (Şeker) Hastası mı?)
- Yaşam Tarzı:** Tuz Tüketimi, Alkol, Sigara Kullanımı, Uyku Süresi, Fiziksel Aktivite Düzeyi.
- Demografik:** Yaş, Cinsiyet ve Aile Öyküsü.

3. YÖNTEM: VERİ ÖN İŞLEME VE MODELLEME

Ham verinin makine öğrenmesi modeline uygun hale getirilmesi için aşağıdaki adımlar uygulanmıştır:

- Eksik Verilerin Giderilmesi (Imputation):** Veri setindeki boş değerler SimpleImputer kullanılarak sütun ortalamaları ile doldurulmuştur.
- Ölçeklendirme (Scaling):** Lojistik Regresyon modelinin doğru çalışabilmesi için StandardScaler ile tüm veriler standartlaştırılmıştır ($\text{Mean}=0$, $\text{Std}=1$). Bu sayede Kolesterol (200 mg/dL) gibi büyük sayıların, Tuz (5 gr) gibi küçük sayıları ezmesi engellenmiştir.
- Algoritma Seçimi:** Model sonuçlarının yorumlanabilir olması ve hangi klinik parametrenin (katsayı/coefficient) riski ne yönde etkilediğinin kullanıcıya şeffafça sunulabilmesi hedeflenmiştir. Bu doğrultuda, modelin **açıklanabilirliğini (Explainability)** sağlamak ve neden-sonuç ilişkisini net bir şekilde ortaya koymamak adına **Lojistik Regresyon (Logistic Regression)** algoritması tercih edilmiştir.
- Sınıf Dengesizliği:** Veri setindeki hasta ve sağlıklı birey sayılarındaki dengesizliği yönetmek için class_weight='balanced' parametresi kullanılmıştır.

4. YAZILIM MİMARİSİ VE KULLANILAN TEKNOLOJİLER

Proje, yazılım mühendisliğinin temel taşı olan **Sorumlulukların Ayrılığı (Separation of Concerns)** prensibine uygun olarak tasarlanmıştır. "Programlama Dilleri" dersi kapsamında aşağıdaki mimari yaklaşımlar ve teknolojiler kullanılmıştır:

4.1. Programlama Prensipleri ve Mimari

- Nesne Yönelimli Programlama (OOP):** Kod yapısı, LoginPenceresi, HipertansiyonApp, VeriIsleyici ve GrafikCizici gibi sınıflara (Classes) bölünmüştür. Bu sayede kodun yönetilebilirliği, okunabilirliği ve tekrar kullanılabilirliği (Reusability) artırılmıştır.
- Modülerlik (Modularity):** Proje, tek bir dosya yığını yerine mantıksal katmanlara ayrılmıştır. Veri işleme (modules/veri_isleme.py) ve grafik çizim (modules/grafikler.py) işlemleri ana dosyadan (main.py) izole edilerek modüler bir yapı kurulmuştur.
- Hata Yönetimi (Exception Handling):** Dosya okuma hataları, eksik kütüphane sorunları ve kullanıcıdan gelen hatalı veri girişleri (örn: sayı yerine harf girilmesi) try-except blokları ile yakalanarak programın çökmesi engellenmiş ve sistemin dayanıklılığı (Robustness) sağlanmıştır.

- **Olay Güdümlü Programlama (Event-Driven):** Tkinter arayüzünde buton tıklamaları ve veri girişleri, kullanıcı etkileşimine dayalı fonksiyonları (callback functions) tetikleyen bir yapıda kurgulanmıştır.
- **Çapraz Platform Uyumluluğu (Cross-Platform Compatibility):** Yazılım, işletim sistemini (Windows/MacOS/Linux) otomatik algılayarak font ayarlarını (Arial veya DejaVu Sans) ve dosya yollarını dinamik olarak yapılandırmaktadır. Bu sayede platform kütüphanesi kullanılarak taşınamabilirlik bir üst düzeye çıkarılmıştır.

4.2. Kullanılan Teknolojiler (Tech Stack)

Proje geliştirme sürecinde tercih edilen temel kütüphaneler ve kullanım amaçları şöyledir:

- **Python 3.13.5:** En güncel kararlı sürümlerden biri olması, performans iyileştirmeleri ve Nesne Yönelimli Programlama (OOP) desteği nedeniyle ana geliştirme dili olarak kullanılmıştır.
- **Tkinter:** Harici kurulum gerektirmeyen hafif yapısı nedeniyle **Kullanıcı Arayüzü (GUI)** geliştirmesinde tercih edilmiştir.
- **Scikit-learn & Pandas:** Veri ön işleme, temizleme ve **Lojistik Regresyon** modelinin eğitimi (backend) için kullanılmıştır.
- **Matplotlib:** Risk analizlerinin ve trend grafiklerinin çizdirilerek arayüze entegre edilmesi (**Veri Görselleştirme**) amacıyla kullanılmıştır.
- **ReportLab:** Analiz sonuçlarının dinamik olarak **PDF raporuna** dönüştürülmesi için kullanılmıştır.

5. SENARYO BAZLI RİSK ANALİZİ

Sistem, kullanıcının beyanına göre 4 farklı senaryoyu otomatik algılar ve buna göre analiz yapar:

1. **Tam Sağlıklı Birey:** Hem Tansiyon hem Diyabet riski hesaplanır.
2. **Sadece Hipertansiyon Hastası:** Mevcut hastalığın Diyabeti tetikleme riski analiz edilir.
3. **Sadece Diyabet Hastası:** Mevcut hastalığın Tansiyonu tetikleme riski analiz edilir.
4. **İki Hastalığa da Sahip Birey:** Risk hesabı yerine "Sağlık Yönetimi Dashboard'u devreye girer ve değerlerin ideal aralıklara uzaklığını gösterilir.

6. BULGULAR VE SİSTEMİN İŞLEYİŞİ

Geliştirilen "Kişisel Hipertansiyon Takip ve Uyarı Sistemi'nin fonksiyonel testleri tamamlanmış; veri girişi, model tahmini, görselleştirme ve raporlama modüllerinin entegre bir

şekilde çalıştığı gözlemlenmiştir. Sistemin çalışma süreci ve elde edilen bulgular aşağıda üç aşamada detaylandırılmıştır:

6.1. Veri Giriş Süreci ve Validasyon

Sistemin ilk aşamasında, kullanıcıdan klinik (Tansiyon, Şeker, Kolesterol vb.) ve yaşam tarzı (Tuz, Sigara, Uyku vb.) verileri alınmaktadır. Testler sırasında, sayısal alanlara (örn: yaş, tansiyon) harf veya geçersiz karakter girilmesi durumunda sistemin try-except blokları sayesinde hatayı yakaladığı ve kullanıcıyı uyardığı tespit edilmiştir. Ayrıca, tuz tüketimi gibi kritik verilerin farklı birimlerden (çay kaşığı, gram vb.) girilse dahi arka planda otomatik olarak gram cinsine dönüştürülerek modele iletiliği doğrulanmıştır.

6.2. Risk Analizi ve Senaryo Tespiti

"Analizi Başlat" komutıyla birlikte sistemin, kullanıcının beyan ettiği hastalık durumuna göre dört farklı senaryodan (Tam Sağlıklı, Sadece HTN, Sadece DM veya İkişi Birlikte) uygun olanı otomatik algıladığı gözlemlenmiştir.

- Model, sağlıklı görünen bir bireyde dahi yüksek tuz tüketimi ve hareketsizlik gibi gizli değişkenleri analiz ederek **%0 ile %100 arasında** bir risk skoru üretmiştir.
- Elde edilen bulgular, yüksek BMI ve 140 mmHg üzeri sistolik kan basıncının risk skorunu logaritmik olarak artırdığını göstermiştir.

6.3. Görselleştirme ve Çıktıların Raporlanması

Analiz sonuçları, kullanıcıya sadece ham veri olarak değil, karar vermeyi kolaylaştıracı grafikler halinde sunulmuştur:

- Konum Analizi:** Kullanıcının verileri, veri setindeki benzer profillerle (benzer yaşı ve cinsiyet grubu) karşılaştırılmış ve kullanıcının "hasta" kümesine mi yoksa "sağlıklı" kümesine mi daha yakın olduğu saçılım (scatter) grafiği üzerinde işaretlenmiştir.
- Etki Analizi:** Kullanıcının riskini artırın en büyük faktörün (örn: Tuz tüketimi veya Uyku düzensizliği) hangisi olduğu, pozitif ve negatif yönlü çubuk grafiklerle (bar chart) net bir şekilde ortaya konulmuştur.
- Raporlama:** Son aşamada sistem, tüm bu analizleri ve kişiye özel oluşturulan "İyileştirme Önerilerini" (örn: "Tuz tüketimini 2 gram azaltmalısınız") içeren PDF formatında, çıktısı alınabilir resmi bir rapor oluşturmayı başarmıştır.

7. SONUÇ VE DEĞERLENDİRME

Geliştirilen proje, teorik makine öğrenmesi algoritmalarının, son kullanıcıya hitap eden pratik bir yazılıma dönüştürülmüş halidir. Sistem, sadece matematiksel bir olasılık vermekle

kalmayıp, kullanıcının "Tuz tüketimim riski ne kadar artırıyor?" sorusuna görsel yanıt veren **Etki Analizi** grafiklerini de içermektedir. Sistem klinik bir tanı aracı değil, bir **Karar Destek Sistemi** olarak tasarlanmıştır.

8. EKLER VE KAYNAKÇA

EKLER:

- **Ek-1: Proje Kaynak Kodları**
(main.py, modules/veri_isleme.py, modules/ grafikler.py)
 - **Ek-2: Kullanılan Veri Seti**
(data/ hypertension_data.csv)
 - **EK-3: Kurulum ve Çalıştırma Kılavuzu**
Projenin hedef bilgisayarda sorunsuz çalışabilmesi için aşağıdaki yazılım ve kütüphane gereksinimlerinin sağlanması gerekmektedir.
 - 1. Sistem Gereksinimleri** Yazılım, **Python 3.13.5** sürümü üzerinde geliştirilmiştir ve test edilmiştir. Çalıştırma ortamında Python'un bu sürümünün veya güncel bir 3.x sürümünün yüklü olması önerilmektedir.
 - 2. Gerekli Kütüphaneler** Projenin çalışması için standart Python kütüphanesine ek olarak aşağıdaki harici paketlerin sisteme yüklü olması gerekmektedir:
pandas: Veri analizi ve manipülasyonu için.
numpy: Sayısal hesaplamalar için.
scikit-learn: Makine öğrenmesi modelleri ve veri ön işleme araçları için.
matplotlib: Grafik ve veri görselleştirme işlemleri için.
reportlab: PDF formatında çıktı üretimi için.

(Not: Grafik arayüz için kullanılan **Tkinter** kütüphanesi, Python kurulumu ile birlikte standart olarak geldiği için harici bir yüklemeye gerektirmez.)

 - 3. Kurulum ve Başlatma** Gerekli kütüphaneler, Python'un paket yöneticisi olan **pip** aracı kullanılarak yüklenmelidir. Kurulum tamamlandıktan sonra proje klasörü içerisinde yer alan **main.py** dosyası çalıştırılarak uygulama başlatılabilir. Veri setinin (hypertension_data.csv) proje ana dizinindeki "data" klasörü içinde veya main.py ile aynı dizinde bulunduğuundan emin olunmalıdır.
- ÖNEMLİ:** main.py dosyası çalıştırıldığında, hypertension_data.csv dosyasının, proje dizini içindeki **data/** klasöründe bulunduğu teyit edilmelidir.

KAYNAKÇA:

- **Panday, A.** (2024). *Hypertension Risk Prediction Dataset*. Kaggle. Erişim Tarihi: Aralık 2025, <https://www.kaggle.com/datasets/ankushpanday1/hypertension-risk-prediction-dataset>
- **Sowers, J. R., Epstein, M., & Frohlich, E. D.** (2001). Diabetes, Hypertension, and Cardiovascular Disease: An Update. *Hypertension*, 37(4), 1053–1059. <https://doi.org/10.1161/01.HYP.37.4.1053>
- **Yamazaki, D., Hitomi, H., & Nishiyama, A.** (2018). Hypertension with diabetes mellitus complications. *Hypertension Research*, 41(3), 147–156. <https://doi.org/10.1038/s41440-017-0008-y>
- **UK Prospective Diabetes Study Group.** (1998). Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. *BMJ*, 317(7160), 703–713. <https://doi.org/10.1136/bmj.317.7160.703>
- **Akhmedova, S. M.** (2024). Specific Characteristics of Clinical and Laboratory Changes in the Course of Arterial Hypertension Against the Background of Diabetes. *International Journal of Health Systems and Medical Sciences*, 3(1), 127–132.