

**T.C.
ERCIYES ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ**

HEART DISEASE PREDICTION

Hazırlayan

**Mahir Barış Sümer
1030520887
Aydan Alkaya
1030520852**

Danışman

Doç.Dr.Özkan Ufuk Nalbantoğlu

**Bilgisayar Mühendisliği
Bitirme Ödevi**

**Mayıs 2023
KAYSERİ**

“**Heart Disease Prediction**” adlı bu çalışma, jürimiz tarafından Erciyes Üniversitesi Bilgisayar Mühendisliği Bölümünde Bitirme Ödevi olarak kabul edilmiştir.

20/03/2023

JÜRİ :

Danışman : Doç.Dr.Özkan Ufuk Nalbantoğlu

Üye :

Üye :

ONAY :

Yukarıdaki imzaların, adı geçen öğretim elemanlarına ait olduğunu onaylarım.

.... / /20...
Prof. Dr. Veysel ASLANTAŞ
Bilgisayar Müh. Bölüm
Başkanı

ÖNSÖZ / TEŞEKKÜR

Bu bitirme tezi, kalp hastalığı tespiti üzerine yapılan bir projenin sonucudur. Bu çalıştığımız proje ile günümüzde sağlık alanında önemli bir sorun olan kalp hastalıklarının erkenden teşhisi ve tahmin edilmesi konusunda bir katkı sağlamayı hedeflemekteyiz.

Bu tezin amacı, mevcut veri analitiği ve yapay zekâ tekniklerini kullanarak kalp hastalığı riskini öngörebilecek bir modelin geliştirilmesidir. Bu model, klinik verileri, hastanın demografik bilgilerini ve biyolojik belirteçlerini kullanarak, erken aşamada kalp hastalığına yatkınlık göstergelerini tanımlayabilecek ve böylelikle hastaların uygun tedaviye erken erişimini sağlayabilecektir.

Bu çalışmanın gerçekleştirilmesi için bize danışman olmayı kabul eden Doç. Dr. Özkan Ufuk Nalbantoğlu'na bu projenin başarılı bir şekilde tamamlanması için bize sağladığı yönlendirme, rehberlik ve değerli önerileri için müteşekkirimiz. Kendisinin bilgeliği ve uzmanlığı sayesinde, araştırmanın kalitesini artırdık ve daha iyi sonuçlara ulaştık.

Bu çalışmada kullanılan verileri sağlayan Kaggle platformuna minnettarız. Kaggle'dan edindiğimiz veri seti bu projenin gerçekleşmesi için büyük bir önem arz etmektedir. Platformun halka açık verileri olmasaydı, bizim bu projenin çalışmasını yürütmemiz mümkün olmazdı.

Son olarak, bu tezi okuyan herkese teşekkür etmek istiyorum. Umarız ki bu çalışma, kalp hastalığı tespiti konusunda daha da fazla farkındalık yaratır ve gelecekteki araştırmalara ilham kaynağı olur. Saygılarımızla,

HEART DISEASE PREDICTION

Mahir Barış Sümer

Erciyes Üniversitesi, Bilgisayar Mühendisliği

Danışman : Doç.Dr.Özkan Ufuk Nalbantoğlu

ÖZET

Projemiz bir kalp hastalığı tahmin sistemi geliştirmeyi amaçlamaktadır. Bu sistemi oluşturmak için, öncelikle bir veri seti toplanacak ve makine öğrenmesi algoritmaları kullanılarak eğitilecektir. Lojistik regresyon ve yapay sinir ağları gibi yaygın algoritmalar kullanılacaktır. Eğitim verileri kullanılarak, bu modellerin performansı değerlendirilecektir. Modellerin doğruluğu, hassasiyeti ve özgüllüğü gibi performans ölçütleri kullanılarak karşılaştırılacaktır. Sonuçlar, bir arayüz geliştirilerek kullanıcılara sunulacaktır. Bu çalışmada, veri gizliliği ve güvenliği de dikkate alınacaktır. Veriler doğru bir şekilde saklanacak ve kullanıcıların kişisel bilgileri korunacaktır. Projenin hayata geçirilmesi için kullanılan yöntemler arasında, veri toplama, veri temizleme ve makine öğrenmesi modelleme yer almaktadır. Ayrıca, performans ölçütleri kullanılarak modellerin doğruluğu doğrulanacak ve sonuçlar arayüz ile kullanıcılara sunulacaktır. Elde edilen sonuçlar, doğru bir şekilde tahmin edilmiş kalp hastalığı riskleri ile sonuçlanacaktır. Bu çalışma, daha sonra sağlık sektöründe kullanılabilecek kalp hastalığı riski tahmini sistemlerinin geliştirilmesi için temel bir adım olacaktır.

Anahtar Kelimeler: Veri İşleme, Yapay Zeka, Sınıflandırma Algoritmaları.

HEART DISEASE PREDICTION

Mahir Barış Sümer

Erciyes University, Computer Engineering

Supervisor: Assoc.Prof.Dr. Özkan Ufuk Nalbantoğlu

ABSTRACT

Our project aims to develop a heart disease prediction system. To build this system, a dataset will first be collected and trained using machine learning algorithms. Common algorithms such as Logistic Regression and artificial neural networks will be used. Using the training data, the performance of these models will be evaluated. The models will be compared using performance metrics such as accuracy, sensitivity and specificity. The results will be presented to the users by developing an interface. Data privacy and security will also be considered in this work. Data will be stored correctly and users' personal information will be protected. The methods used to implement the project include data collection, data cleaning and machine learning modeling. In addition, the accuracy of the models will be verified using performance metrics and the results will be presented to users via an interface . The results will result in accurately predicted heart disease risks. This work will be a fundamental step for the development of heart disease risk prediction systems that can then be used in the healthcare sector.

Keywords: Data Processing, Artificial Intelligence, Classification Algorithms.

İÇİNDEKİLER

HEART DISEASE PREDICTION

KABUL VE ONAY	i
ÖNSÖZ / TEŞEKKÜR	ii
ÖZET	iii
ABSTRACT	iv
İÇİNDEKİLER	v
TABLolar LİSTESİ	vii
ŞEKİLLER LİSTESİ	viii
KISALTMALAR	ix

GİRİŞ	1
-----------------	---

1. BÖLÜM

PROJENİN SİSTEM ANALİZİ

1.1. Giriş:	7
1.1.1. Gant	8
1.2. Sistem Mimarisi:	8
1.3. Veri Temizleme:	9
1.4. Makine Öğrenmesi Modellemesi:	10
1.5. Performans Değerlendirmesi:	11
1.6. Veri Gizliliği ve Güvenliği:	12

2. BÖLÜM

PROJENİN GELİŞTİRİLMESİ

2.1. PROJENİN WEB ARAYÜZÜ	13
2.1.1. Chest Pain Type	15
2.1.2. Resting Blood Pressure	16
2.1.3. Serum Cholesterol in mg/dl(chol)	16

2.1.4. fasting blood sugar > 120 mg/dl	17
2.1.5. Resting electrocardiographic results	18
2.1.6. Maximum heart rate achieved	20
2.1.7. Exercise induced angina	21
2.1.8. ST depression induced by exercise relative to rest	22
2.1.9. The slope of the peak exercise ST segment	23
2.1.10.number of major vessels (0-3) colored by flourosopy	24
2.1.11.Thal	26
2.2.PROJENİN KAYNAK KODLARI VE AÇIKLAMASI	28
2.3.PROJENİN Details Sayfası	34
3. BÖLÜM	
TARTIŞMA, SONUÇ ve ÖNERİLER	
3.1.Tartışma, Sonuç ve Öneriler	35
KAYNAKLAR	36
EKLER	36
ÖZGEÇMİŞ	37

TABLÖLAR LİSTESİ

Tablo 2.1. Kullanılacak Teknolojiler	14
--	----

ŞEKİLLER LİSTESİ

Şekil 1.1.	Gantt Şeması	8
Şekil 2.1.	Projenin Web Arayüzü.	14
Şekil 2.2.	Kütüphaneleri Ekleme	28
Şekil 2.3.	Verileri Ekleme	28
Şekil 2.4.	DataFrame İnceleme	29
Şekil 2.5.	Eksik Verileri Kontrol Etme	29
Şekil 2.6.	Train ve Test Verisini Ayırma	30
Şekil 2.7.	LogisticRegression Modelini Oluşturma ve Eğitim	30
Şekil 2.8.	Flask Uygulamasını Oluşturma	31
Şekil 2.9.	JSON Verilerini Alma	31
Şekil 2.10.	Tahmin Sonucu Atama	32
Şekil 2.11.	Details yoluna gelen POST isteğine yanıt	32
Şekil 2.12.	Uygulamayı Başlatma	33

KISALTMALAR

SVM : Destek Vektör Makineleri

GİRİŞ

Kalp hastalıkları, dünya genelinde en yaygın ölüm nedenlerinden biridir. Yüksek tansiyon, obezite, diyabet, sigara içme, stresli yaşam tarzı ve genetik faktörler gibi birçok etken, bireylerin kalp sağlığı üzerinde önemli bir etkiye sahip olabilir. Bu nedenle, kalp hastalığı riskini tahmin etmek ve erken teşhis etmek, hastaların uygun tedavi ve önlemleri almasını sağlayarak hastalığın ilerlemesini engellemek için son derece önemlidir.

Bilgisayar destekli tahmin sistemleri, tıp alanında önemli bir araç haline gelmiştir ve kalp hastalığı tahmini konusunda da büyük bir potansiyele sahiptir. Bu çalışmada, bir kalp hastalığı tahmini sistemi geliştirmeyi amaçlamaktayız. Sistem, bireylerin klinik verilerini analiz ederek, kalp hastalığı riskini değerlendirecek ve erken uyarılar sağlayacaktır.

Çalışmanın temel amacı, makine öğrenimi ve veri analitiği yöntemlerini kullanarak, kalp hastalığı olan veya olma riski taşıyan bireyleri belirlemek için bir model geliştirmektir. Bu model, çeşitli demografik, klinik ve yaşam tarzı verilerini temel alarak, bir bireyin kalp hastalığına yakalanma olasılığını tahmin etmeyi amaçlamaktadır. Böylece, sağlık uzmanlarına, risk altındaki bireylere daha erken müdahale etme ve önleyici tedbirler önerme konusunda yardımcı olabilecek önemli bilgiler sunulabilecektir.

Bu çalışma, birçok araştırma makalesinde tartışılan kalp hastalığı tahmini alanındaki son gelişmeleri bir araya getirecektir. Özellikle, veri analitiği ve makine öğrenimi tekniklerinin kalp hastalığı riski tahmini üzerindeki etkisini araştıracaktır. Ayrıca, çalışmanın kapsamı, kullanılan veri kaynakları, veri toplama yöntemleri ve kullanılan öznitelikler gibi ayrıntıları da açıklayacağız.

Bu çalışmanın sonuçları, kalp hastalığı riski tahmininde bilgisayar destekli yöntemlerin potansiyelini göstermek ve klinik uygulamalarda kullanılabilecek değerli bir araç sunmak için tasarlanmıştır.

Bilgisayar destekli tahmin sistemleri, büyük veri analizi ve makine öğrenimi teknikleri sayesinde kalp hastalığı riskini tahmin etmek için etkili bir araç haline gelmiştir. Bu çalışma, bu teknolojilerin kullanılmasının sağlayabileceği potansiyeli araştırarak ve sağlık uzmanlarına ve hasta bakım ekiplerine değerli bir araç sunmaya çalışacaktır.

Lojistik regresyon, birçok sınıflandırma problemi için kullanılan istatistiksel bir yöntemdir. Kalp hastalığı tahmini gibi bir ikili sınıflandırma problemi için ideal bir seçimdir. Lojistik regresyon, bağımlı değişkenin (kalp hastalığı varlığı) bağımsız değişkenlerle (risk faktörleri ve klinik parametreler) ilişkisini modellemek için kullanılır. Bu model, bir kişinin kalp hastalığına sahip olma olasılığını tahmin etmek için kullanılır.

Lojistik regresyon modeli oluşturulurken, bağımlı değişken (kalp hastalığı varlığı) ile bağımsız değişkenler (risk faktörleri ve klinik parametreler) arasındaki ilişkiyi açıklayan bir denklem oluşturulur. Bu denklem, veriye en iyi şekilde uyan parametrelerin tahmin edilmesini sağlamak için uygun bir optimizasyon algoritması kullanılarak eğitilir. Eğitim süreci, modelin doğruluğunu artırmak için tekrarlanır ve modelin performansı eğitim ve test veri setlerinde değerlendirilir.

Eğitilmiş lojistik regresyon modeli kullanılarak yeni verilere dayalı olarak kalp hastalığı riski tahmin edilebilir. Bu tahmin, kişinin risk faktörleri ve klinik parametrelerine dayalı olarak yapılır. Ayrıca, modelin performansını değerlendirmek için çeşitli metrikler kullanılır. Doğruluk, hassasiyet, geri çağırma gibi metrikler, modelin ne kadar iyi tahmin yaptığını ve gerçek değerlerle karşılaştırıldığında yanlış sınıflandırma oranlarını gösterir.

Lojistik regresyon, kalp hastalığı tahmini gibi sağlık alanında önemli bir uygulamada etkili bir araçtır. Kalp hastalığı riskini belirlemek için risk faktörleri ve klinik parametrelerin analiz edilmesi, erken teşhis ve uygun tedavi planlaması için büyük

önem taşır. Lojistik regresyon modeli, bu analizleri gerçekleştirmede yardımcı olabilir ve sağlık uzmanlarına önemli bir öngörücü bakış sunabilir.

[1] Bu çalışma, aşağıdaki bölümlerden oluşmaktadır:

- Giriş: Çalışmanın amacı, önemi ve örgütlenmesi hakkında genel bir bakış sunulmaktadır.
- Literatür Taraması: Kalp hastalığı tahmini konusunda yapılan önceki araştırmalar ve benzer çalışmalar incelenmektedir.
- Veri Toplama ve Ön İşleme: Kullanılan veri kaynakları, veri toplama yöntemleri ve veri ön işleme adımları ayrıntılı bir şekilde açıklanmaktadır.
- Öznitelik Seçimi ve Veri Analizi: Kullanılan özniteliklerin seçimi ve veri analizi yöntemleri detaylı bir şekilde açıklanmaktadır.
- Kalp Hastalığı Tahmini Modeli: Geliştirilen kalp hastalığı tahmini modeli ve kullanılan algoritmalar hakkında bilgi verilmektedir.
- Sonuçlar: Elde edilen sonuçlar, modelin performansı ve değerlendirme kriterleri üzerine analiz edilmektedir.
- Tartışma ve Öneriler: Sonuçların tartışılması, çalışmanın sınırlamaları ve gelecekteki araştırmalar için öneriler sunulmaktadır.
- Sonuç: Çalışmanın özeti ve elde edilen bulguların önemi vurgulanmaktadır.

Bu bölümler, çalışmanın ilerleyen kısımlarında daha detaylı olarak açıklanacak ve kalp hastalığı tahmini sistemine yönelik kapsamlı bir anlayış sağlayacaktır.

Literatür Özeti

A Systematic Review on Heart Disease Prediction using Machine Learning Techniques Yazarlar: Saranya, M. ve Kavitha, S. Yayın: IOP Conference Series: Materials Science and Engineering, 2019 Özet: Mkiine öğrenimi tekniklerinin kullanıldığı kalp hastalığı tahmini yöntemleri incelenmiştir. Çalışma, mevcut literatürde kullanılan çeşitli veri analizi tekniklerini, algoritmaları ve sonuçları içermektedir.

Heart Disease Prediction using Machine Learning Algorithms: A Systematic Literature Review Yazarlar: Azeem, M. A. ve Dubey, S. R. Yayın: International Journal of Electrical and Computer Engineering (IJECE), 2020 Özet: Bu literatür taraması, makine öğrenimi algoritmalarının kullanıldığı kalp hastalığı tahminine odaklanmaktadır. Çalışma, çeşitli algoritmaların performansını ve karşılaştırmalarını içeren bir inceleme sunmaktadır.

Machine Learning for Prediction of All-Cause Mortality in Patients with Suspected Coronary Artery Disease: A 5-Year Multicentre Prospective Registry Analysis Yazarlar: Motwani, M. et al. Yayın: European Heart Journal, 2018 Özet: Bu çalışma, makine öğrenimi tekniklerinin kullanıldığı kalp hastalığı tahmininin klinik bir örneğidir. Araştırma, birçok merkezde gerçekleştirilen bir prospektif kayıt analizini içeren makine öğrenimini değerlendirmektedir.

Predicting Heart Disease using Machine Learning Techniques: A Systematic Literature Review Yazarlar: Ahmed, A. F., Nawi, N. M. ve Jantan, A. Yayın: Journal of Telecommunication, Electronic and Computer Engineering, 2019 Özet: Bu literatür taraması, makine öğrenimi tekniklerinin kullanıldığı kalp hastalığı tahmini üzerinde odaklanmaktadır. Çalışma, literatürde kullanılan farklı veri kaynaklarını, algoritmaları ve performans değerlendirmelerini özetle sunmaktadır.

Heart Disease Prediction using Machine Learning: A Systematic Literature Review Yazarlar: Ramírez-Ortiz, J. L. et al. Yayın: Sensors, 2020 Özet: Makine öğrenimi kullanan kalp hastalığı tahmini alanındaki literatürü sistematik olarak incelemektedir.

Çalışmanın Amacı

Bu çalışmasının amaçları şu şekilde sıralanabilir:

- Bu çalışmanın amacı, kalp hastalığı tahmin sisteminin geliştirilmesidir. Kalp hastalıkları, dünya genelinde önemli bir sağlık sorunu olup, yaşam kalitesini etkileyen ve potansiyel olarak ölümcül sonuçlara yol açabilen durumlardır. Erken teşhis, tedavi ve önlemler, kalp hastalıklarının etkilerini azaltabilir ve yaşam süresini uzatabilir. Bu nedenle, bu çalışma, kalp hastalıklarının önceden tahmin edilmesi ve risk faktörlerinin belirlenmesi üzerine odaklanmaktadır.
- Bu çalışmada, kalp hastalığı ile ilişkili olan çeşitli risk faktörleri ve belirtiler üzerinde kapsamlı bir analiz yapılacaktır. Makine öğrenimi tekniklerini kullanarak, mevcut büyük veri setlerinden anlamlı bilgiler çıkarmayı hedeflemekteyiz. Kalp hastalığına yol açan faktörleri belirlemek için veri analizi ve modelleme yöntemleri kullanılacak ve bu bilgiler üzerine dayanarak, kalp hastalığı geliştirme riski taşıyan bireyleri önceden tespit etmek için bir tahmin modeli oluşturulacaktır.
- Elde edilen sonuçlar, sağlık uzmanlarının, doktorların ve sağlık politika yapıcılarının kalp hastalıklarıyla ilgili daha iyi kararlar almasına yardımcı olacaktır. Ayrıca, bireylere erken uyarılar sağlayarak, sağlıklarını korumaları ve kalp hastalığına yönelik önleyici tedbirler alabilmeleri konusunda farkındalık yaratılmasına katkıda bulunmayı hedefliyoruz.
- Bu çalışmanın sonucunda, etkili bir kalp hastalığı tahmin sistemi elde ederek, toplum sağlığını iyileştirmek ve kalp hastalıklarının yayılmasını azaltmak için önleyici sağlık hizmetlerinin geliştirilmesine katkıda bulunmayı umuyoruz.

Tezin Organizasyonu

Bu çalışmanın düzenlemesi şu şekildedir.

1. Bölümde, Projenin Sistem Analizi ve Aşamaları
2. Bölümde, Projenin Geliştirilmesi , Projenin Arayüzünün Açıklanması ve Kodlar
3. Bölümde Tartışma , Sonuç ve Öneriler şeklindedir.

1. BÖLÜM

PROJENİN SİSTEM ANALİZİ

- Giriş
- Sistem Mimarisi
- Veri Temizleme
- Makine Öğrenmesi Modellemesi
- Performans Değerlendirmesi
- Veri Gizliliği ve Güvenliği

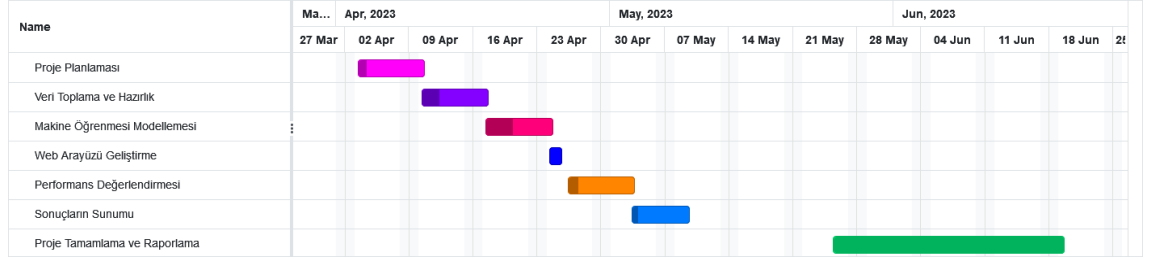
1.1. Giriş:

Bu bölümde projenin genel amacı ve hedefleri belirlenecektir. Kalp hastalığı tahmin sistemi geliştirme amacımız ve sistemin kullanılacağı alanlar gibi konular ele alınacaktır. Ayrıca, projenin önemi ve neden bu sisteme ihtiyaç duyulduğu da açıklanacaktır.

- Projemiz, bir kalp hastalığı tahmin sistemi geliştirmeyi amaçlamaktadır.
- Sistemi oluşturmak için, bir veri seti toplanacak ve makine öğrenmesi algoritmaları kullanılarak eğitilecektir.
- Lojistik Regresyon ve yapay sinir ağları gibi yaygın algoritmalar kullanılacaktır.
- Eğitim verileri kullanılarak, modellerin performansı değerlendirilecektir.

- Modellerin doğruluğu, hassasiyeti ve özgüllüğü gibi performans ölçütleri kullanılarak karşılaştırılacaktır.
- Sonuçlar, bir arayüz geliştirilerek kullanıcılara sunulacaktır.

1.1.1. Gantt



Şekil 1.1. Gantt Şeması

1.2. Sistem Mimarisi:

Bu bölümde, projenin genel sistem yapısı ve bileşenleri detaylandırılacaktır. Sistemde kullanılacak donanım ve yazılım bileşenleri, veri akışı ve işlem süreçleri gibi konular ele alınacaktır. Sistem mimarisi şemaları ve diyagramları kullanılarak sistemin bileşenleri arasındaki ilişkiler gösterilecektir.

- Veri Toplama: Kalp hastalığıyla ilgili çeşitli veri kaynaklarından veri toplanacak (hastane kayıtları, klinik veriler, laboratuvar sonuçları vb.). [2]
- Veri Temizleme: Toplanan veri seti, gereksiz ve eksik verilerden arındırılacak. Öznitelik seçimi yapılacak ve bozuk veya eksik veriler düzeltilecektir.
- Makine Öğrenmesi Modellemesi: Temizlenen veri seti kullanılarak Lojistik Regresyon ve yapay sinir ağları gibi algoritmalarla makine öğrenmesi modelleri oluşturulacaktır.
- Performans Değerlendirmesi: Oluşturulan modeller, eğitim verileri üzerinde performans ölçütleri kullanılarak değerlendirilecektir (doğruluk, hassasiyet, özgüllük, F1 skoru vb.).

- Sonuçların Sunumu: En iyi performans gösteren model, kullanıcıların erişimine sunulmak üzere bir arayüz olarak hazırlanacaktır.

1.3. Veri Temizleme:

Bu bölümde, projenin genel sistem yapısı ve bileşenleri detaylandırılacaktır. Sistemde kullanılacak donanım ve yazılım bileşenleri, veri akışı ve işlem süreçleri gibi konular ele alınacaktır. Sistem mimarisi şemaları ve diyagramları kullanılarak sistemin bileşenleri arasındaki ilişkiler gösterilecektir.

Veri temizleme, kalp hastalığı tahmin sisteminin önemli bir adımıdır ve veri setindeki eksik veya hatalı verilerin düzeltilmesi ve aykırı değerlerin ele alınması için bir dizi işlemi kapsar. İşte veri temizleme adımının detayları:

1. Eksik Veri Analizi:

- Veri setinde eksik değerlerin olup olmadığı kontrol edilir. Eksik veriler, boş hücreler, 'NaN' veya 'null' değerler olarak temsil edilebilir.
- Eksik veri analizi yapılırken, eksik verilerin neden kaynaklandığı ve bunların veri seti üzerindeki etkileri değerlendirilir.

2. Eksik Veri Doldurma:

- Eksik verilerin uygun bir şekilde doldurulması için yöntemler kullanılır. Bunlar arasında şunlar bulunabilir:
- Basit Değer Atama: Eksik değerler, ortalama, medyan, mod veya sabit bir değer ile doldurulabilir.
- İleri Dolum: Eksik değerler, diğer değişkenlerin bilgilerinden yararlanılarak tahmin edilebilir. Örneğin, bir özelliğin eksik değeri, diğer benzer örneklerin ortalamasıyla doldurulabilir.
- Makine Öğrenmesi Tabanlı Dolum: Eksik değerler, makine öğrenmesi algoritmaları kullanılarak tahmin edilebilir.

1.4. Makine Öğrenmesi Modellemesi:

Bu bölümde, makine öğrenmesi algoritmalarının kullanılacağı modelleme süreci açıklanacaktır. Lojistik Regresyon ve yapay sinir ağları gibi yaygın algoritmalar kullanılacaktır. Veri seti, eğitim ve test verisi olarak ayrılacak ve algoritmalar eğitim verileri üzerinde çalışacak. Bu bölümde, algoritmaların seçimi ve parametre ayarlamaları gibi kararlar da verilecektir.

- Veri Hazırlığı:

Veri seti, kalp hastalığıyla ilgili çeşitli öznitelikleri (yaş, cinsiyet, kan basıncı, kolesterol düzeyi, egzersiz alışkanlığı vb.) içeren hastaların verilerini içermektedir. Veri seti, özelliklerin bir araya getirildiği tablo veya matris formatında olmalıdır. Her özellik bir sütunda temsil edilmelidir. Ayrıca, veri setinde eksik veriler veya aykırı değerler gibi veri kalitesi sorunları olabilir. Bu nedenle, veri seti, eksik verilerin doldurulması ve aykırı değerlerin düzeltilmesi adımlarını içeren bir veri temizleme sürecinden geçirilmelidir.

- Veri Bölümleme:

Veri seti, genellikle eğitim ve test veri setlerine bölünür. Eğitim veri seti, modelin eğitiminde kullanılırken, test veri seti, modelin performansını değerlendirmek için kullanılır. Veri seti, rastgele veya zaman temelli bölümleme yöntemleriyle bölünebilir. Örneğin,

- Model Seçimi:

Makine öğrenmesi algoritmaları arasından en uygun modelin seçilmesi önemlidir. Kalp hastalığı tahmini için lojistik regresyon , destek vektör makineleri (SVM), karar ağaçları, random forest ve yapay sinir ağları gibi yaygın algoritmalar kullanılabilir. Model seçimi, veri setinin özelliklerine, boyutuna ve problem alanına bağlı olarak yapılmalıdır. Bu aşamada deneyler ve karşılaştırmalar yapılabilir.

- Model Eğitimi:

Seçilen makine öğrenmesi modeli, eğitim veri seti üzerinde eğitilmelidir. Eğitim süreci, modelin veri setindeki kalıpları ve ilişkileri öğrenmesini içerir. Modelin eğitimi sırasında, hiperparametreler (örneğin, öğrenme oranı, derinlik, ağ yapılandırması) belirlenir ve optimize edilir.

1.5. Performans Değerlendirmesi:

Bu bölümde, oluşturulan makine öğrenmesi modellerinin performansının değerlendirilmesi ayrıntılı olarak açıklanacaktır. Performans ölçütleri arasında doğruluk, hassasiyet, özgüllük gibi metrikler kullanılacaktır. Modeller arasında karşılaştırmalar yapılacak ve en iyi performansa sahip olan model belirlenecektir.

- Veri Seti Bölünmesi: Veri seti, eğitim ve test verisi olarak ayrılacaktır. Tipik olarak, verilerin
- Performans Ölçütleri: Doğruluk (Accuracy): Modelin doğru tahmin ettiği örneklerin oranıdır. Doğruluk, genel model performansını değerlendirmede yaygın olarak kullanılan bir ölçüttür. Hassasiyet (Precision): Pozitif olarak tahmin edilen örneklerin gerçekten pozitif olanların oranını temsil eder. Hassasiyet, yanlış pozitiflerin önlenmesi önemli olduğu durumlarda değerlendirme yapmak için kullanılır. Özgüllük (Specificity): Negatif olarak tahmin edilen örneklerin gerçekten negatif olanların oranını ifade eder. Özgüllük, yanlış negatiflerin önlenmesi önemli olduğu durumlarda değerlendirme yapmak için kullanılır.
- Model Performansının Karşılaştırılması: Oluşturulan makine öğrenmesi modelleri, eğitim verileri kullanılarak eğitildikten sonra test verileri üzerinde değerlendirilecektir. Her bir model için doğruluk, hassasiyet ve özgüllük ölçütleri hesaplanacak ve karşılaştırılacaktır. En iyi performansa sahip olan model, yüksek doğruluk, hassasiyet ve özgüllük değerlerine sahip olan model olarak belirlenecektir.

- Model İyileştirmesi: Performans değerlendirmesi sonucunda en iyi performansa sahip model belirlendikten sonra, modelin performansını daha da artırmak için çeşitli iyileştirme teknikleri uygulanabilir. Bu teknikler arasında veri seti genişletme, hiperparametre ayarlama, öznitelik seçimi ve farklı makine öğrenmesi algoritmalarının denemesi yer alabilir.

1.6. Veri Gizliliği ve Güvenliği:

Bu bölümde, veri gizliliği ve güvenliği önlemlerinin detaylarıyla açıklanacaktır. Verilerin saklanması ve kullanıcı bilgilerinin korunması için uygun güvenlik yöntemleri ve şifreleme teknikleri kullanılacaktır. Veri anonimleştirme ve erişim kontrolü gibi önlemler, kullanıcıların kişisel bilgilerinin gizliliğini sağlamak için uygulanacaktır. Ayrıca, veri paylaşımı ve veri saklama süreleri gibi konular da ele alınacaktır.

- Veri güvenliği önemli bir konudur ve projede dikkate alınacaktır.
- Veriler doğru bir şekilde saklanacak ve kullanıcıların kişisel bilgileri korunacaktır.
- Veri toplama, depolama ve erişim aşamalarında gizlilik ve güvenlik önlemleri alınacaktır.
- Veri koruma politikaları ve gizlilik standartlarına uygun hareket edilecektir.

2. BÖLÜM

PROJENİN GELİŞTİRİLMESİ

Sunucu tarafında Python ve Flask kullanılırken, web arayüzü için HTML, CSS ve JavaScript kullanılmıştır. Makine öğrenmesi algoritması olarak Logistic Regression seçilmiştir. Veri temizleme, veri görselleştirme ve veri analizi işlemleri için Python ve ilgili kütüphaneler (Pandas, Matplotlib, Seaborn, NumPy) kullanılmıştır. Verilerin depolanması için SQLite veritabanı tercih edilmiştir.


2.1. PROJENİN WEB ARAYÜZÜ

Projenin web arayüzü, kullanıcıların kalp hastalığı tahminleri alabilmek için kullanabilecekleri kullanıcı dostu bir formdan oluşmaktadır. Bu form, kullanıcıların gerekli bilgileri girerek tahmin sonuçlarını almasını sağlar.

Tablo 2.1. Kullanılacak Teknolojiler

Bileşen	Teknoloji/Aracı
Sunucu Dili	Python
Web Framework	Flask
Web Arayüzü	HTML, CSS, JS
Makine Öğrenmesi Algoritması	LogisticRegression
Veri Temizleme	Python, Pandas
Veri Analizi	Python, Pandas, NumPy

Heart Disease Prediction



Patient Information

First Name

Last Name

Age Gender

Patient Health Data

Chest Pain Type

Resting Blood Pressure

Serum Cholesterol in mg/dl(chol)

fasting blood sugar > 120 mg/dl

Resting electrocardiographic results

Maximum heart rate achieved

Exercise induced angina

ST depression induced by exercise relative to rest

The slope of the peak exercise ST segment

number of major vessels (0-3) colored by flourosopy

Thal:

Submit

2.1.1. Chest Pain Type

Chest Pain Type, yani göğüs ağrısı türleri, kalp hastalığına ilişkin riskleri ve belirtileri anlamak için önemli bir faktördür. İşte farklı göğüs ağrısı türlerinin tanımları:

1. Typical Angina (Tipik Angina): Bu tip göğüs ağrısı genellikle kalp hastalığı olan bireylerde ortaya çıkar. Ağrı, göğüs bölgesinde veya çoğunlukla göğüs kemiği arkasında hissedilir ve genellikle fiziksel aktivite veya stres gibi tetikleyicilerle ilişkilidir. Ağrı, tipik olarak birkaç dakika sürer ve nitrogliserin gibi nitrat ilaçları ile tedavi edilebilir. Tipik angina genellikle kalp damarlarında daralma veya tıkanıklık gibi koroner arter hastalığının bir belirtisidir.
2. Atypical Angina (Atipik Angina): Atipik angina, tipik angina belirtilerinden farklı semptomlar gösteren bir göğüs ağrısı türüdür. Ağrı, tipik anginadakine benzer şekilde göğüs bölgesinde hissedilebilir, ancak süresi daha kısa veya daha uzun olabilir. Atipik angina semptomları arasında nefes darlığı, halsizlik, mide bulantısı veya karın ağrısı gibi belirtiler de yer alabilir. Bu tür ağrı, kalp hastalığı olan bireylerde kalp damarlarında orta derecede daralma veya tıkanıklık olduğunda ortaya çıkabilir.
3. Non-anginal Pain (Angina Dışı Ağrı): Non-anginal pain, göğüs ağrısının kalp hastalığından kaynaklanmadığı durumları tanımlamak için kullanılan bir terimdir. Bu tür ağrı genellikle göğüs bölgesinde hissedilir, ancak kalp ile ilgili bir soruna bağlı değildir. Nedenleri arasında sindirim sorunları, kas-iskelet sistemi sorunları, sinir sistemi problemleri veya solunum sistemi enfeksiyonları yer alabilir.
4. Asymptomatic (Belirtisiz): Bazı insanlar kalp hastalığına sahip olmasına rağmen herhangi bir göğüs ağrısı veya semptom yaşamazlar. Bu duruma "asymptomatic" denir. Bu kişiler, rutin sağlık taramaları sırasında veya başka bir nedenle yapılan testlerle kalp sorunlarının varlığını öğrenebilirler. Asymptomatic bireylerde kalp hastalığı riski yüksek olabilir, bu nedenle düzenli sağlık kontrolleri önemlidir.

Göğüs ağrısı türleri, kalp hastalığının tanısında ve tedavisinde önemli bir rol oynar.

2.1.2. Resting Blood Pressure

Resting Blood Pressure, yani dinlenme halindeki kan basıncı, bir kişinin hareketsiz durumdayken ölçülen kan basıncını ifade eder. Genellikle bir kişinin tansiyonunu ölçmek için kullanılan iki değerden oluşur: sistolik kan basıncı (üst değer) ve diyastolik kan basıncı (alt değer). Yaşa göre referans değerleri şu şekildedir:

1. Yenidoğanlar (0-1 aylık): Sistolik: 60-90 mmHg, Diyastolik: 20-60 mmHg
2. Bebekler (1 ay - 1 yaş): Sistolik: 70-100 mmHg, Diyastolik: 40-60 mmHg
3. Çocuklar (1-10 yaş): Sistolik: 80-110 mmHg, Diyastolik: 50-70 mmHg
4. Ergenlik dönemi (11-17 yaş): Sistolik: 90-120 mmHg, Diyastolik: 60-80 mmHg
5. Yetişkinler (18 yaş ve üzeri): Sistolik: 90-120 mmHg, Diyastolik: 60-80 mmHg

Bu referans değerler genel olarak kabul gören normal aralıklardır. Ancak, kişinin yaşam tarzı, sağlık durumu ve tıbbi geçmişi gibi faktörler, ideal kan basıncı değerlerini etkileyebilir. Yüksek veya düşük kan basıncı, tıbbi değerlendirme ve takip gerektirebilir. Kan basıncı değerlerinin düzenli olarak ölçülmesi ve sağlık uzmanı tarafından değerlendirilmesi önemlidir.

2.1.3. Serum Cholesterol in mg/dl(chol)

Serum Cholesterol (Serum Kolesterol), kandaki kolesterol seviyesini ifade eder. Kolesterol, vücut için önemli bir yağ türüdür ve hücre zarlarının yapısında, hormon üretiminde ve sindirimi sağlamada önemli bir rol oynar. Serum kolesterol seviyesi, çeşitli faktörlerden etkilenir ve genellikle miligram per desilitre (mg/dL) cinsinden ölçülür. Referans değerleri, kolesterol seviyelerinin sağlık açısından kabul edilebilir aralıklarını belirtir. Genellikle yetişkinler için serum kolesterol referans değerleri şu şekildedir:

1. İdeal/Normal: Toplam Kolesterol: Altında 200 mg/dL LDL Kolesterol (Kötü Kolesterol): Altında 130 mg/dL HDL Kolesterol (İyi Kolesterol): 40 mg/dL veya daha yüksek Trigliserid: Altında 150 mg/dL
2. Sınır Değer: Toplam Kolesterol: 200-239 mg/dL LDL Kolesterol (Kötü Kolesterol): 130-159 mg/dL HDL Kolesterol (İyi Kolesterol): 40 mg/dL'nin altı Trigliserid: 150-199 mg/dL
3. Yüksek: Toplam Kolesterol: 240 mg/dL ve üzeri LDL Kolesterol (Kötü Kolesterol): 160 mg/dL ve üzeri HDL Kolesterol (İyi Kolesterol): 60 mg/dL'nin üzeri Trigliserid: 200 mg/dL ve üzeri

Bu referans değerler genel olarak kabul edilen aralıklardır, ancak sağlık durumu, yaş, cinsiyet ve diğer faktörler bu değerleri etkileyebilir. Yüksek kolesterol seviyeleri, kalp hastalığı riskini artırabilir, bu nedenle düzenli olarak kolesterol testi yaptırmanız ve sonuçları sağlık uzmanıyla değerlendirmeniz önemlidir.

2.1.4. fasting blood sugar > 120 mg/dl

fasting Blood Sugar (Açlık Kan Şekeri), kişinin aç karnına ölçülen kan şekeri seviyesini ifade eder. Bu değer, genellikle diyabet teşhisi ve kan şekeri kontrolünü değerlendirmek için kullanılır. Fasting Blood Sugar değeri, miligram per desilitre (mg/dL) cinsinden ölçülür. Fasting Blood Sugar değeri 120 mg/dL'nin üzerinde olduğunda, bu durum "impaired fasting glucose" (bozulmuş açlık kan şekeri) veya prediyabet olarak adlandırılır. Bu durum, kan şekeri düzeyinin normal aralıkta olmadığını gösterir ve diyabet gelişme riskini artırabilir. Yüksek fasting blood sugar değerleri, aşağıdaki gibi faktörlerle kalp hastalığı riskini artırabilir:

1. İnsülin direnci: Yüksek fasting blood sugar değerleri, insülin direnci ile ilişkili olabilir. İnsülin direnci, vücudun insüline olan yanıtının azaldığı bir durumdur ve diyabet ve kalp hastalığı riskini artırır.
2. Diyabet: Yüksek fasting blood sugar değerleri, diyabet gelişme riskini artırır. Diyabet, kan şekeri düzeyinin sürekli olarak yüksek olduğu bir durumdur ve kalp damarlarını etkileyerek kalp hastalığı riskini artırır.

3. Metabolik Sendrom: Yüksek fasting blood sugar değerleri, metabolik sendromun bir belirtisi olabilir. Metabolik sendrom, insülin direnci, yüksek kan basıncı, yüksek trigliserid seviyeleri, düşük HDL kolesterol seviyeleri ve obeziteyi içeren bir dizi risk faktörünün birleşimi olarak tanımlanan bir durumdur ve kalp hastalığı riskini artırır. Referans değerlerine gelince, açlık kan şekeri değerleri genellikle aşağıdaki aralıklarda kabul edilir:

- Normal: Açlık Kan Şekeri: 70-99 mg/dL
- Impaired Fasting Glucose (Prediyabet): Açlık Kan Şekeri: 100-125 mg/dL
- Diyabet: Açlık Kan Şekeri: 126 mg/dL ve üzeri

Ancak, bu referans değerler bir sağlık uzmanı tarafından bireysel duruma göre değerlendirilmelidir. Diyabet veya prediyabet teşhisi konulduğunda, kan şekeri kontrolünü sağlamak ve kalp hastalığı riskini azaltmak için uygun tedavi ve yaşam tarzı değişiklikleri önerilebilir.

2.1.5. Resting electrocardiographic results

Resting Electrocardiographic Results, yani dinlenme halindeki elektrokardiyografik sonuçlar, bir kişinin kalp aktivitesini değerlendirmek için elektrokardiyografi (EKG) testiyle elde edilen verileri ifade eder. EKG, kalp ritmi, kalp hızı ve kalp kasının elektriksel aktivitesi gibi faktörleri ölçerek kalp sağlığıyla ilgili bilgiler sağlar. Kalp hastalığını nasıl etkilediği konusunda, Resting EKG sonuçları, potansiyel kalp problemlerini tespit etmek ve değerlendirmek için kullanılır. EKG sonuçları, kalp ritmi ve kalp kasının normal veya anormal çalışmasına dair bilgiler sağlar. Bununla birlikte, EKG sonuçları yalnızca bir değerlendirme aracıdır ve kesin bir tanı koymak için diğer testler ve klinik değerlendirme gerekebilir. Normal EKG sonuçları, tipik olarak kalp ritminin düzenli ve düzgün olduğunu, kalp kasının normal bir şekilde elektriksel olarak iletişim sağladığını gösterir. Bu durumda, kalp hastalığı belirtileri genellikle mevcut değildir ve kalp sağlığı normal kabul edilir. Having ST-T wave abnormality (ST-T dalga anormallikleri) ise EKG sonuçlarında gözlemlenen belirli bir anormallik türünü ifade eder. ST segmenti, QRS kompleksi ile T dalgası arasında bulunan bir bölümdür ve T dalgası, kalp kasının yeniden polarize olduğunu gösterir.

ST-T dalga anormallikleri, ST segmentinde yükseklik, alçalma, düzleşme veya ters T dalgası şeklinde değişikliklerin görüldüğü durumları ifade edebilir. ST-T dalga anormallikleri, kalp hastalığının belirtilerinden biri olabilir. Örneğin, koroner arter hastalığı, miyokard enfarktüsü (kalp krizi), miyokardit (kalp kası iltihabı) veya kalp kası hipertrofisi gibi durumlar ST-T dalga anormalliklerine neden olabilir. Bu anormallikler, kalp kasının kan akışının düzensiz olduğunu veya kalp duvarında hasar olduğunu gösterebilir. Ancak, ST-T dalga anormallikleri her zaman kalp hastalığına işaret etmez. Başka faktörler de bu anormalliklere neden olabilir, örneğin, elektrolit dengesizlikleri, ilaç etkileşimleri veya normalden farklı bir kalp yapısı. ST-T dalga anormalliklerinin kalp hastalığı ile ilişkili olup olmadığını belirlemek için aşağıdaki adımlar izlenebilir:

1. Detaylı değerlendirme: EKG sonuçlarındaki ST-T dalga anormallikleri, diğer klinik bulgular ve hastanın semptomlarıyla birlikte değerlendirilmelidir. Örneğin, hastanın göğüs ağrısı, nefes darlığı veya baş dönmesi gibi kalp hastalığı belirtileri var mı?
2. EKG paternleri: ST-T dalga anormallikleri farklı paternlerde ortaya çıkabilir. Bazıları kalp hastalığına özgü olabilirken, diğerleri farklı nedenlerden kaynaklanabilir. Örneğin, ST segmentinde yükseklik, tipik olarak akut miyokard enfarktüsünü (kalp krizi) gösterebilirken, düzleşmiş veya ters T dalgası, diğer kalp problemlerine veya non-kardiyak faktörlere işaret edebilir.
3. Ek testler: ST-T dalga anormallikleri kalp hastalığına işaret ediyorsa veya hastada risk faktörleri varsa, ek testler yapılabilir. Bu testler arasında stres testi, ekokardiyografi (kalp ultrasonu), koroner anjiyografi (damarların görüntülenmesi) ve kan testleri bulunur.

Sonuç olarak, ST-T dalga anormallikleri EKG sonuçlarında görülen belirli değişikliklerdir ve kalp hastalığına işaret edebilirler. Ancak, bu anormalliklerin tanısı ve etkilerinin değerlendirilmesi için daha fazla test ve klinik değerlendirme gereklidir. Bir kardiyolog, hastanın özelliklerini değerlendirerek doğru teşhis ve tedavi planını yapabilir.

2.1.6. Maximum heart rate achieved

Maximum Heart Rate Achieved (Maksimum Kalp Hızı), bir kişinin egzersiz sırasında veya stres altında ulaştığı en yüksek kalp atım hızını ifade eder. Maksimum kalp hızı, genellikle bireyin yaşına bağlı olarak tahmin edilebilir ve egzersiz testleri veya aktiviteler sırasında ölçülebilir. Kalp hastalıklarını nasıl etkilediği konusunda, maksimum kalp hızı çeşitli şekillerde değerlendirilebilir:

1. Egzersiz toleransı: Maksimum kalp hızı, bir kişinin egzersiz toleransını ve aerobik kapasitesini yansıtabilir. Düşük maksimum kalp hızı, egzersiz sırasında daha çabuk yorulma, nefes darlığı veya kalp krizi riskinin artması gibi sorunlara neden olabilir.
2. Kalp hastalığı riski: Maksimum kalp hızı, kalp hastalığı riskini belirlemede bir gösterge olarak kullanılabilir. Yaşa bağlı olarak beklenen maksimum kalp hızı ile gerçek maksimum kalp hızı arasındaki fark, bir kişinin kalp hastalığı riskini etkileyebilir. Örneğin, yaşına göre beklenenden daha düşük bir maksimum kalp hızı, kalp hastalığı riskinin arttığını gösterebilir.

Referans aralıklarına gelince, yaşa göre tahmini maksimum kalp hızı için yaygın olarak kullanılan bir formül vardır: $220 - \text{yaş}$. Bu formül, ortalama olarak maksimum kalp hızını tahminlemeye yardımcı olur. Ancak, her bireyin fizyolojisi farklı olduğu için, bu tahminsel formül kişiden kişiye değişiklik gösterebilir. Örneğin, 40 yaşında bir kişinin maksimum tahmini kalp hızı 180 olabilir ($220 - 40 = 180$). Ancak, bu sadece bir tahmindir ve gerçek maksimum kalp hızı kişinin genel sağlık durumu, düzenli egzersiz alışkanlıkları ve diğer faktörlere bağlı olarak değişebilir. Dolayısıyla, maksimum kalp hızı değerlendirilirken bireysel faktörler göz önünde bulundurulmalı ve kapsamlı bir kardiyovasküler değerlendirme yapılmalıdır. Bir kardiyolog, maksimum kalp hızının kişiye özgü etkilerini ve kalp sağlığı üzerindeki potansiyel etkilerini değerlendirebilir.

2.1.7. Exercise induced angina

Egzersize bağı anjina (Exercise-induced angina), genellikle fiziksel aktivite veya egzersiz sırasında ortaya çıkan göğüs ağrısı veya rahatsızlık hissidir. Anjin, kalbin yeterli oksijen alamadığı durumlarda ortaya çıkar ve genellikle koroner arter hastalığı (kalp damarlarının tıkanması veya daralması) nedeniyle meydana gelir.

Egzersize bağı anjinanın kalp hastalıklarını nasıl etkilediğı şu şekilde açıklanabilir:

1. İyi bir belirti olabilir: Egzersiz sırasında ortaya çıkan anjina, bir kişinin koroner arterlerdeki daralma veya tıkanıklığı belirleyebilen bir belirti olabilir. Göğüs ağrısı veya rahatsızlık hissi, kalbin ihtiyaç duyduğu oksijen ve besin maddelerinin yetersiz olduğunu gösterebilir.
2. Kalp hastalığı riskini yükseltebilir: Egzersize bağı anjina, genellikle koroner arter hastalığı olan kişilerde ortaya çıkar. Koroner arter hastalığı, kalp damarlarının plaklarla daralması veya tıkanmasıdır. Egzersize bağı anjina, kalp hastalığı riskini artırabilir ve daha ciddi komplikasyonlara, özellikle miyokard enfarktüsüne (kalp krizi) yol açabilir.
3. Egzersiz kapasitesini sınırlayabilir: Egzersize bağı anjina, kişinin fiziksel aktiviteye olan toleransını azaltabilir. Göğüs ağrısı veya rahatsızlık hissi, kişinin egzersiz sırasında daha çabuk yorulmasına ve aktivite düzeyini kısıtlamasına neden olabilir. Bu da genel olarak fiziksel formu ve yaşam kalitesini etkileyebilir.

Egzersize bağı anjina, kalp hastalığı olan bireylerde dikkate alınması gereken bir semptomdur. Bu semptom, kalp hastalığının teşhisi, tedavisi ve yönetimi için önemli bir ipucu olabilir. Egzersize bağı anjina semptomları yaşayan bir kişi, bir kardiyologa başvurmalı ve uygun değerlendirme ve tedavi sürecini takip etmelidir. Tedavi genellikle ilaçlar, yaşam tarzı değişiklikleri ve gerektiğinde invaziv işlemler (örneğin, anjiyografi ve anjiyoplasti) içerebilir.

2.1.8. ST depression induced by exercise relative to rest

Egzersizle indüklenen dinlenmeye göre ST depresyonu" terimi, egzersiz sırasında kalp üzerindeki yüklenmenin neden olduğu bir elektrokardiyografik (EKG) değişikliği ifade eder. ST segmenti, EKG'de görülen bir bölümdür ve ST segmentindeki depresyon, kalp kasının oksijen ihtiyacının arttığı egzersiz sırasında, kalp kasının yeterli oksijen alamadığını gösterebilir. Egzersizle indüklenen ST depresyonu, genellikle koroner arter hastalığı olan kişilerde görülür. Koroner arter hastalığı, kalp damarlarının daralması veya tıkanması sonucu kan akışının sınırlanması anlamına gelir. Egzersiz sırasında, kalp kasının daha fazla oksijene ihtiyaç duyduğu için daralmış veya tıkanmış damarlardan yeterli kan akışı sağlanamaz. Bu durum, ST segmentinde depresyona yol açar. Egzersizle indüklenen ST depresyonu, kalp hastalıklarını etkileyebilir ve aşağıdaki şekillerde belirtiler gösterebilir:

1. Koroner arter hastalığı tespiti: Egzersiz sırasında ST segmentindeki depresyon, koroner arter hastalığının bir göstergesi olabilir. Bu değişiklik, kalp damarlarının daralması veya tıkanması nedeniyle kalp kasına yeterli oksijenin ulaşamadığını gösterir. Bu durum, kalp hastalığı riskini artırabilir ve miyokard enfarktüsü (kalp krizi) gibi ciddi komplikasyonlara yol açabilir.
2. Egzersiz toleransında azalma: Egzersiz sırasında ST depresyonu yaşayan bireylerde genellikle egzersiz toleransı azalır. Göğüs ağrısı, nefes darlığı veya yorgunluk gibi semptomlar egzersiz sırasında ortaya çıkabilir ve kişinin aktivite düzeyini kısıtlayabilir. Bu, genel yaşam kalitesini etkileyebilir ve fiziksel kapasiteyi sınırlayabilir.
3. İlerleyici kalp hastalığı belirtisi: Egzersizle indüklenen ST depresyonu, koroner arter hastalığının ilerleyebileceğini ve daha ciddi komplikasyonlara yol açabileceğini gösterebilir. Bu nedenle, bu değişikliklerin izlenmesi ve gerektiğinde tedavi edilmesi önemlidir. Egzersizle indüklenen ST depresyonu, kalp hastalıklarının bir belirtisi veya göstergesi olabilir, ancak tek başına bir tanı aracı değildir. Bir kardiyolog, semptomları, EKG sonuçlarını ve diğer

testleri değer indirerek doğru tanıyı koyma ve tedavi planını belirleme konusunda uzmanlaşmıştır. Egzersizle indüklenen ST depresyonu olan bir birey, kardiyolog tarafından değerlendirilmeli ve gerektiğinde ek testler (örneğin stres testi, koroner anjiyografi) yapılmalıdır.

Tedavi genellikle şu faktörlere dayanır:

1. İlaçlar: Koroner arter hastalığı veya anjina semptomlarını kontrol etmek için ilaçlar reçete edilebilir. Bu ilaçlar, kalp kasının oksijen ihtiyacını azaltabilir, kan damarlarını genişletebilir ve kan pıhtılaşmasını engelleyebilir.
2. Yaşam tarzı değişiklikleri: Sağlıklı bir yaşam tarzı benimsemek, koroner arter hastalığının ilerlemesini yavaşlatabilir ve semptomları hafifletebilir. Bunlar arasında düzenli fiziksel aktivite, sağlıklı bir diyet, sigara bırakma ve stres yönetimi yer alır.
3. İnvaziv işlemler: Ciddi koroner arter hastalığı durumunda, invaziv işlemler gerekebilir. Bunlar arasında koroner anjiyoplasti (tıkali damarların genişletilmesi) ve koroner bypass cerrahisi (tıkali damarların çevresinden yeni bir yol oluşturulması) yer alır.

Egzersizle indüklenen ST depresyonu olan bir kişi, bu durumu ciddiye almalı ve bir kardiyologa başvurmalıdır. Erken tanı ve uygun tedavi, kalp hastalığına bağlı komplikasyonları önlemeye ve sağlıklı bir yaşam sürdürmeye yardımcı olabilir.

2.1.9. The slope of the peak exercise ST segment

"Egzersiz sırasında ST segmentinin eğimi" terimi, egzersiz testi sırasında ST segmentinin yükselme veya düşme oranını ifade eder. Bu, EKG'deki ST segmentinin egzersiz süresi boyunca nasıl değiştiğini belirlemeye yardımcı olur. ST segmenti, kalp kasının elektriksel aktivitesini yansıtan bir bölümdür ve bazı durumlarda anormallikler, kalp hastalığına işaret edebilir. ST segmentinin eğimi, kalp hastalıklarını etkileyebilir ve belirli bir şekilde yorumlanabilir:

1. Upsloping (Yukarı doğru eğimli): ST segmenti, egzersiz sırasında yukarı doğru eğimliyse, genellikle kalp hastalığı riski düşüktür. Bu durum, birçok kişi için normal bir bulgu olarak kabul edilir.
2. Flat (Düz): ST segmenti, egzersiz sırasında düz bir çizgi şeklinde kalıyorsa, bu durumda kalp hastalığı riski orta düzeydedir. Düz ST segmenti, kalp kasının oksijen ihtiyacının tam olarak karşılanmadığını veya kalp damarlarında daralma olduğunu gösterebilir.
3. Downsloping (Aşağı doğru eğimli): ST segmenti, egzersiz sırasında aşağı doğru eğimliyse, bu durumda kalp hastalığı riski daha yüksek olabilir. Aşağı doğru eğimli ST segmenti, kalp damarlarında ciddi daralma veya tıkanıklığın olduğunu ve kalp kasının yeterli oksijen alamadığını gösterebilir.

ST segmentinin eğimi, kalp hastalığına bağlı belirtilerin bir göstergesi olabilir. Özellikle downsloping olarak adlandırılan aşağı doğru eğimli ST segmenti, kalp hastalığının varlığını veya ilerlemesini gösteren bir bulgu olabilir. Bu nedenle, egzersiz testi sonuçları dikkatlice değerlendirilmeli ve gerektiğinde daha fazla inceleme yapılmalıdır. Bir kardiyolog, ST segmentinin eğimini ve diğer klinik bulguları değerlendirerek doğru tanıyı koyma ve tedavi planını belirleme konusunda uzmanlaşmıştır.

2.1.10. number of major vessels (0-3) colored by flourosopy

"Fluoroscopi ile renklendirilmiş önemli damar sayısı (0-3)" terimi, bir koroner anjiyografi sırasında görüntülenen koroner arterlerdeki önemli (ana) damarların sayısını ifade eder. Koroner anjiyografi, kalp damarlarının tıkanıklıklarını veya daralmalarını belirlemek için kullanılan invaziv bir işlemdir. Number of major vessels (0-3) colored by fluoroscopy, kalp hastalığı riskini etkileyebilir ve aşağıdaki şekillerde yorumlanabilir:

1. 0 damar: Koroner anjiyografi sırasında hiçbir önemli damarın tıkanık veya daralmış görülmediği durumu ifade eder. Bu durum, koroner arter hastalığı (CAD) riskinin düşük olduğunu gösterebilir. Ancak, bu sonuç yalnızca anjiyografi

sırasında gözlemlenen damarların değerlendirilmesine dayanır ve başka damarlar da etkilenebilir.

2. 1 damar: Koroner anjiyografi sırasında bir önemli damarın tıkanıklığı veya daralması tespit edildiği durumu ifade eder. Bu durum, koroner arter hastalığı riskinin artmış olduğunu gösterebilir. Tek bir tıkanık veya daralmış damar, kalp kasına yeterli kan akışını kısıtlayabilir ve semptomlara, özellikle anjina (göğüs ağrısı) veya kalp krizine yol açabilir.
3. 2-3 damar: Koroner anjiyografi sırasında iki veya üç önemli damarın tıkanıklığı veya daralması tespit edildiği durumu ifade eder. Bu durum, koroner arter hastalığı riskinin yüksek olduğunu gösterebilir. İki veya üç tıkanık veya daralmış damar, kalp kasının büyük bir bölümünü etkileyebilir ve ciddi semptomlara ve komplikasyonlara yol açabilir.

Number of major vessels colored by fluoroscopy, koroner arter hastalığı riskini etkiler çünkü tıkanıklık veya daralma, kalp kasına yeterli oksijen ve besin maddesi taşıyan kan akışını kısıtlar. Bu durum, anjinaya, miyokard enfarktüsüne (kalp krizi) ve diğer ciddi kalp problemlerine yol açabilir. Ayrıca, daha fazla tıkanık veya daralmış damar sayısı, kalp hastalığına bağlı komplikasyonlar ve mortalite riskinin artmasına neden olabilir. Bununla birlikte, koroner anjiyografi sonuçları tek başına bir teşhis aracı değildir. Başka klinik bulgular, semptomlar ve diğer test sonuçları da dikkate alınmalıdır. Bu nedenle, bir kardiyolog, koroner anjiyografi sonuçlarına dayanarak koroner arter hastalığı teşhisi koyma ve tedavi planını belirleme konusunda uzmanlaşmıştır. Ek olarak, diğer faktörler de değerlendirilir, örneğin kişinin semptomları, tıbbi öyküsü, diğer test sonuçları ve risk faktörleri (örneğin sigara içme, yüksek tansiyon, diyabet gibi). Tedavi genellikle koroner arter hastalığına bağlı semptomları kontrol etmeyi, kalp krizi riskini azaltmayı ve kalp sağlığını korumayı hedefler. Tedavi seçenekleri arasında ilaç tedavisi, anjiyoplasti, koroner stent yerleştirme veya koroner baypas cerrahisi bulunabilir. Hangi tedavi yönteminin tercih edileceği, damar tıkanıklığı derecesine, semptomlara, genel sağlık durumuna ve diğer faktörlere bağlı olarak belirlenir. Sonuç olarak, koroner anjiyografi sırasında görülen önemli damar sayısı, koroner arter hastalığı

riskini etkiler ve tedavi planını belirleme sürecinde önemli bir bilgidir. Bu bilgi, kalp hastalığına bağlı semptomların şiddetini ve ilerlemesini anlamada yardımcı olur ve uygun tedavi seçeneklerinin değerlendirilmesini sağlar. Ancak, bu sonuçların bir kardiyolog tarafından değerlendirilmesi ve diğer klinik bulgularla birlikte dikkate alınması önemlidir.

2.1.11. Thal

"Thal" terimi, tıbbi literatürde "talasemi" olarak bilinen genetik bir kan bozukluğunu ifade etmektedir. Thalasemia, hemoglobin molekülünde eksiklik veya anormalliklere neden olan bir grup kalıtsal hastalığı kapsar. Kalp hastalığı riski, talasemideki belirli faktörlere bağlı olarak değişiklik gösterebilir. Thalasemi, hemoglobin molekülünün yapısını etkilediği için, kanın normal şekilde oksijen taşımamasını engelleyebilir. Bu durum, kalp kasının oksijen ihtiyacını etkileyebilir ve kalp hastalığı riskini artırabilir. Talasemili bireylerde kalp komplikasyonları daha sık görülebilir, özellikle aşırı demir yüklemesi (transfüzyonlar nedeniyle) nedeniyle. Kalp hastalığı riskini değerlendirmek için talasemide kullanılan bir test "Thal" olarak adlandırılan bir skorlama sistemidir. Bu sistem, bir kişinin talasemiye bağlı kalp hastalığı riskini belirlemek için klinik ve laboratuvar parametrelerini değerlendirir. Thal skoru, kalp hastalığı riskinin yüksek veya düşük olabileceğini belirlemeye yardımcı olabilir. Thal skoru, talasemide görülen kalp hasarı belirtilerini değerlendirmek için kullanılan nükleer tıp testleri sonuçlarına dayanır. Bu testler, kalp kasının kan akışını ve oksijenlenmesini değerlendirmek için kullanılır. Bu testlerin sonucunda elde edilen bulgular, kalp kasının hasar derecesini ve kan akışındaki anormallikleri belirlemeye yardımcı olur. Thal skorunda üç farklı kategori vardır:

1. Normal: Kalp kasında herhangi bir anormallik veya hasar belirtisi olmadığı durumu ifade eder. Bu durumda kalp hastalığı riski düşüktür.
2. Fixed (Sabit): Kalp kasında sabit bir hasar veya anormallik olduğu durumu ifade eder. Bu durumda kalp hastalığı riski artabilir ve kalp komplikasyonları gelişme olasılığı yüksektir.

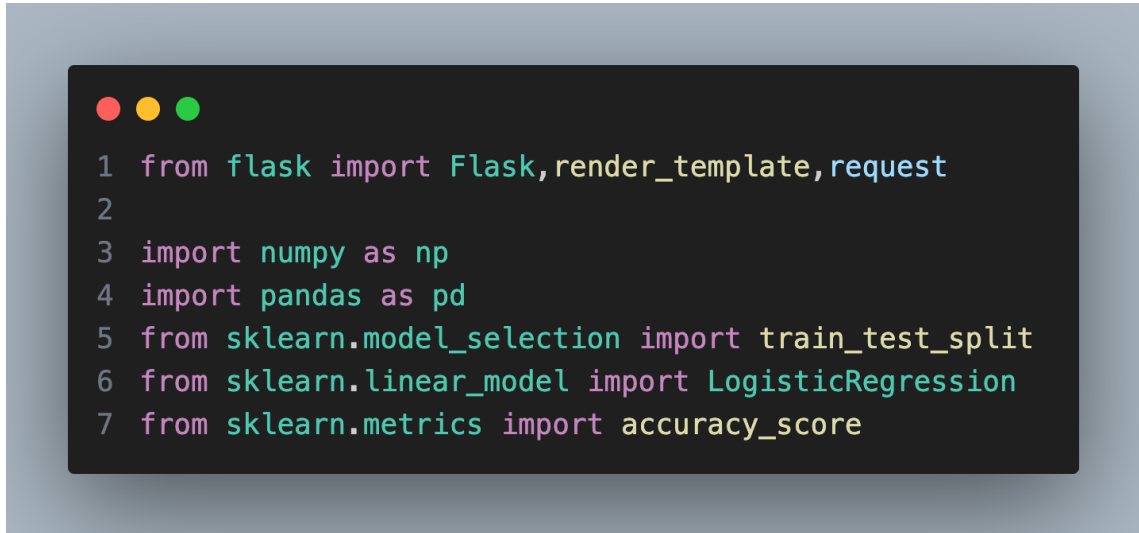
3. Reversible Defect (Tersine Çevrilebilir Anormallik): Test sonuçları, kalp kasında geçici bir anormallik olduğunu ve kalp hastalığı riskinin düşük olduğunu gösterir. Bu durumda kalp komplikasyonları riski daha düşüktür ve genellikle tedaviye iyi yanıt verme olasılığı daha yüksektir.

Thal skoru, talasemili bireylerin kalp sağlığını değerlendirmek için kullanılan bir araçtır. Bu skor, hastaların takibini kolaylaştırır ve gerekirse tedavi planlamasında rehberlik eder. Özellikle sabit hasar veya anormalliklerin olduğu durumlarda, kalp komplikasyonlarına karşı daha yakın takip ve önlemler alınması önemlidir.

Reversible Defect olarak adlandırılan tersine çevrilebilir anormallikler genellikle tedavi ile düzeltilebilir. Bu durumda, talasemili bireylerin kalp sağlığını iyileştirmek ve komplikasyon riskini azaltmak için tedavi seçenekleri değerlendirilebilir. Bu tedaviler arasında demir yüklemesinin kontrolü, kan transfüzyonlarının uygun yönetimi, demir şelasyon tedavisi ve diğer ilaçlar yer alabilir. Tedavi, bireysel duruma ve hastanın spesifik ihtiyaçlarına göre planlanmalıdır.

Sonuç olarak, Thal skoru, talasemili bireylerde kalp hastalığı riskini değerlendirmek için kullanılan bir araçtır. Skor, kalp kasında mevcut hasarın veya anormalliklerin tespitini sağlar ve tedavi planlamasında rehberlik eder. Sabit hasarlar kalp hastalığı riskini artırırken, tersine çevrilebilir anormalliklerin olduğu durumlarda kalp komplikasyonlarına yönelik daha umut verici bir yaklaşım mümkün olabilir. Talasemili bireylerin kalp sağlığı, düzenli takip, uygun tedavi ve yaşam tarzı düzenlemeleri ile optimize edilmelidir.

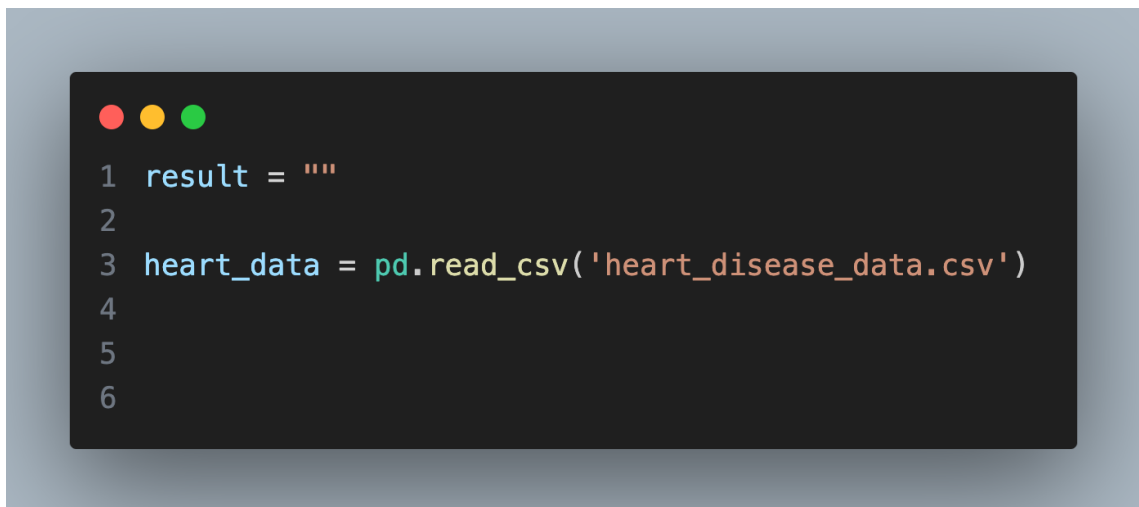
2.2. PROJENİN KAYNAK KODLARI VE AÇIKLAMASI



Şekil 2.2. Kütüphaneleri Ekleme

[3] İlk olarak,Flask kütüphanesinden Flask, rendertemplate ve request modüllerini import ediyoruz. Ardından, diğer gerekli kütüphaneleri numpy, pandas, traintestsplit, LogisticRegression, accuracyscore) import ediyoruz.

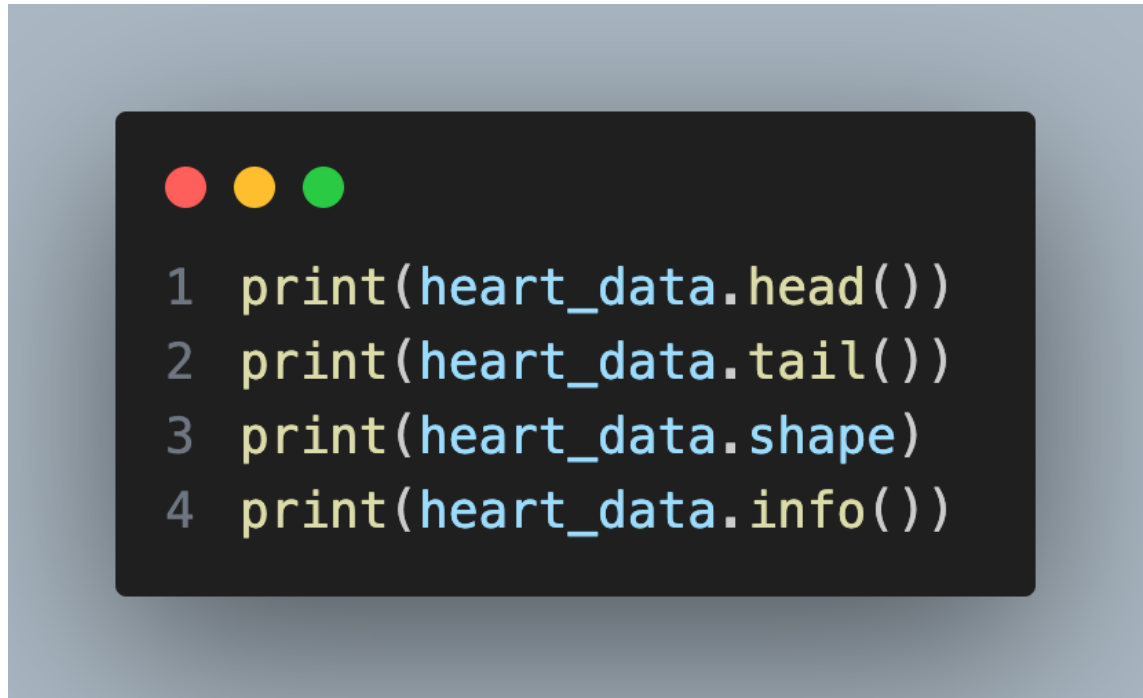
[4]



Şekil 2.3. Verileri Ekleme

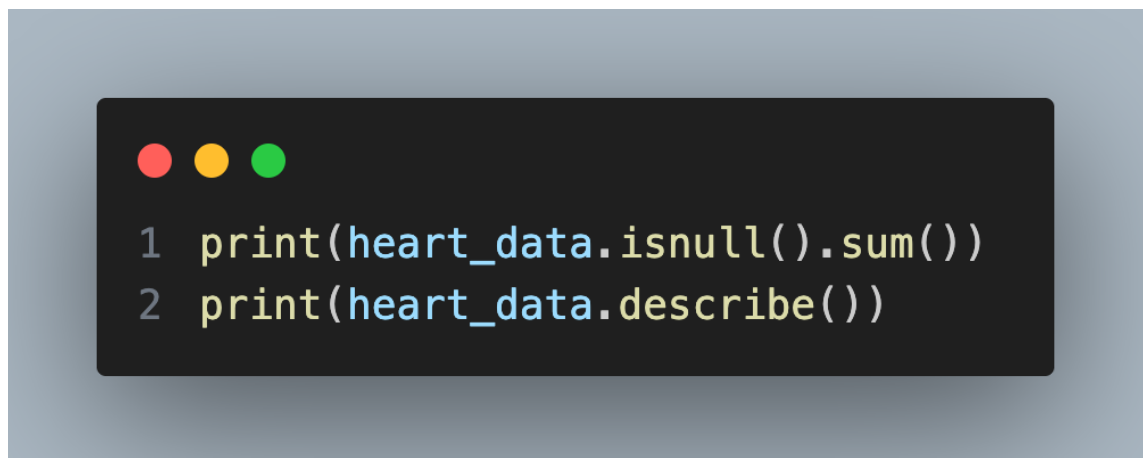
result adında boş bir değişken tanımlıyoruz. Bu değişken sonucu tutacak. CSV

formatındaki kalp hastalığı veri setini heartdata adında bir Pandas DataFrame'e yüklüyoruz.



Şekil 2.4. DataFrame İnceleme

Veri setinin ilk 5 satırını ve son 5 satırını ekrana yazdırarak veriyi kontrol ediyoruz. Veri setinin satır ve sütun sayısını kontrol ediyoruz. Veri hakkında bazı bilgileri alarak veri seti hakkında bir genel bakış sağlıyoruz.



Şekil 2.5. Eksik Verileri Kontrol Etme

Eksik değerleri kontrol ediyoruz. Veri hakkında istatistiksel ölçümleri

göüntülüyoruz.

```
1 X = heart_data.drop(columns='target', axis=1)
2 Y = heart_data['target']
3 print(X)
4 print(Y)
5
6 X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, stratify=Y, random_state=2)
7 print(X.shape, X_train.shape, X_test.shape)
```

Şekil 2.6. Train ve Test Verisini Ayırma

Hedef değişkenin dağılımını kontrol ediyoruz. Veri setini özellikler (X) ve hedef değişken (Y) olarak ayırıyoruz. Veriyi eğitim ve test verisi olarak ayırıyoruz.

```
1 model = LogisticRegression()
2 model.fit(X_train, Y_train)
3
4
5 X_train_prediction = model.predict(X_train)
6 training_data_accuracy = accuracy_score(X_train_prediction, Y_train)
7
8
9 X_test_prediction = model.predict(X_test)
10 test_data_accuracy = accuracy_score(X_test_prediction, Y_test)
11
```

Şekil 2.7. LogisticRegression Modelini Oluşturma ve Eğitme

Lojistik regresyon modelini oluşturup eğitiyoruz. Modelin performansını değerlendiriyoruz, eğitim ve test verileri üzerinde doğruluk skorunu hesaplıyoruz.



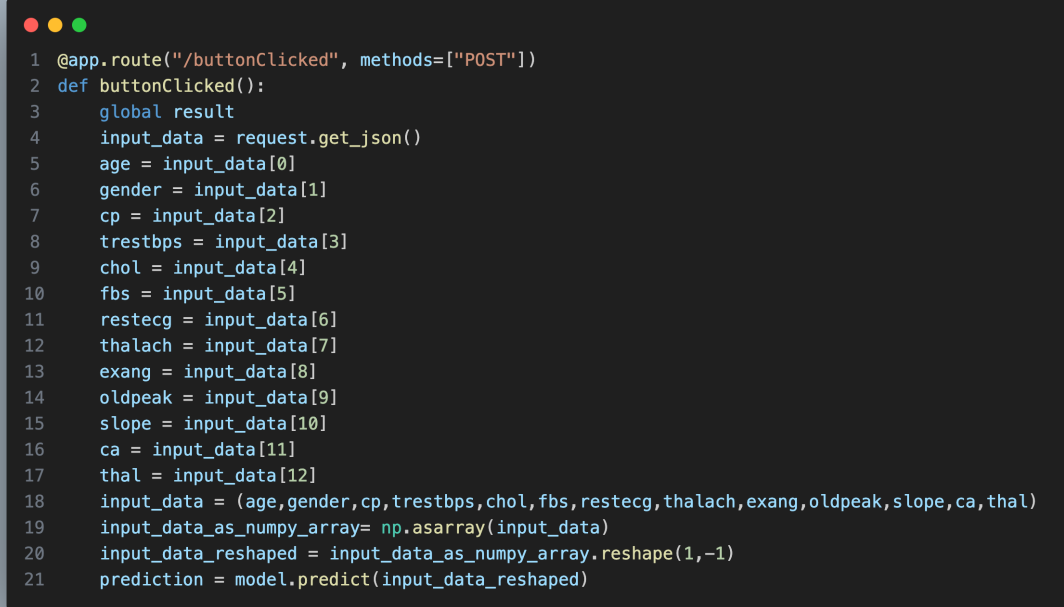
```

1 app = Flask(__name__)
2
3 @app.route("/")
4 def home():
5     return render_template("index.html", variable = result)

```

Şekil 2.8. Flask Uygulamasını Oluşturma

Flask uygulamasını oluşturuyoruz. Ana sayfa ("/") için bir route tanımlıyoruz. Bu route, index.html adlı bir HTML şablonunu render ediyor ve result değişkenini template'e gönderiyor.



```

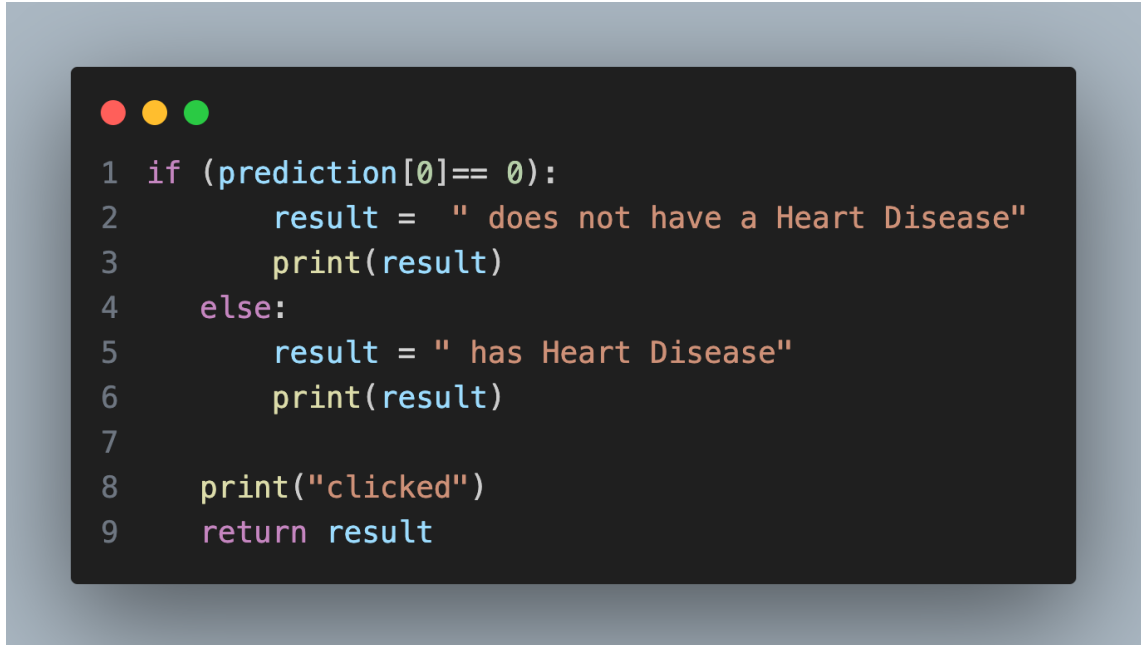
1 @app.route("/buttonClicked", methods=["POST"])
2 def buttonClicked():
3     global result
4     input_data = request.get_json()
5     age = input_data[0]
6     gender = input_data[1]
7     cp = input_data[2]
8     trestbps = input_data[3]
9     chol = input_data[4]
10    fbs = input_data[5]
11    restecg = input_data[6]
12    thalach = input_data[7]
13    exang = input_data[8]
14    oldpeak = input_data[9]
15    slope = input_data[10]
16    ca = input_data[11]
17    thal = input_data[12]
18    input_data = (age, gender, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal)
19    input_data_as_numpy_array= np.asarray(input_data)
20    input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)
21    prediction = model.predict(input_data_reshaped)

```

Şekil 2.9. JSON Verilerini Alma

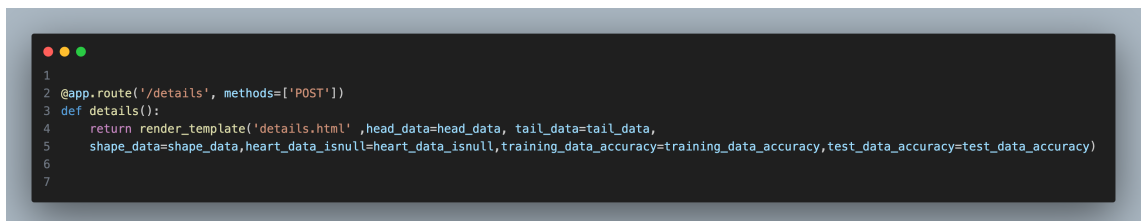
"/buttonClicked" route'u için bir işlev tanımlıyoruz. Bu işlev, bir POST isteği geldiğinde çalışır. İstekten gelen JSON verilerini alır ve modele tahmin yapmak için kullanır. Bu kod, Flask web framework'ünü kullanarak bir web uygulaması oluşturur ve kullanıcının kalp hastalığı olup olmadığını tahmin etmesini sağlar. Ana sayfada

bir form bulunur ve kullanıcı formu doldurup "Submit" düğmesine tıkladığında, girilen veriler modele gönderilir ve sonuç görüntülenir.



Şekil 2.10. Tahmin Sonucu Atama

Tahmin sonucunu result değişkenine atar ve geri döndürür.



Şekil 2.11. Details yoluna gelen POST isteğine yanıt

Bu kod, Flask web uygulamasında details yoluna gelen POST isteğine yanıt vermek için kullanılır. İsteğin geldiği durumda details.html adlı bir şablonu rendertemplate fonksiyonuyla işler ve bu şablonu döndürür. Gönderilen değişkenleri details.html sayfasında kullanmamızı sağlar.



Şekil 2.12. Uygulamayı Başlatma

Uygulamayı `if name == "main":` bloğunda çalıştırıyoruz. Bu, Python kodları doğrudan çalıştırıldığında, Flask uygulamasının başlatılmasını sağlar.

2.3. PROJENİN Details Sayfası

Head Data														
	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1

Tail Data														
	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0

Shape Data

(303, 14)

Empty Data

age 0 sex 0 cp 0 trestbps 0 chol 0 fbs 0 restecg 0 thalach 0 exang 0 oldpeak 0 slope 0 ca 0 thal 0 target 0 dtype: int64

Train Data Accuracy

0.8512396694214877

Test Data Accuracy

0.819672131147541

[Return to Main Menu](#)

1. Head Data: Projede kullanılan veri setinin sütunlarını ve ilk 5 satırını gösterir.
2. Tail Data: Projede kullanılan veri setinin sütunlarını ve son 5 satırını gösterir.
3. Shape Data: Projede kullanılan veri setinin sırası ile kaç satır ve sütundan oluştuğunu gösterir.
4. Empty Data: Projede kullanılan veri setinin sütun değerlerini ve o değer için kaç tane boş değer olduğunu(0) gösterir.Son parametre de ise kullanılan değerlerinin tipini gösterir.(dtype: int64)
5. Train Data Accuracy: Eğitim verisinin doğruluk skorunu gösterir.
6. Test Data Accuracy: Test verisinin doğruluk skorunu gösterir.

3. BÖLÜM

TARTIŞMA, SONUÇ ve ÖNERİLER

3.1. Tartışma, Sonuç ve Öneriler

Bu çalışma, kalp hastalığı tespitine yönelik bir modelin geliştirilmesi ve değerlendirilmesini içermektedir. Veri analitiği teknikleri kullanılarak, farklı makine öğrenmesi algoritmaları üzerinde performans değerlendirmesi yapılmıştır.

Elde edilen sonuçlar, model performansının değerlendirilmesi için çeşitli metrikler kullanılarak analiz edilmiştir. Performans ölçütleri arasında doğruluk, hassasiyet ve özgüllük yer almaktadır. Modeller arasında karşılaştırma yapılarak, hangi algoritmanın daha iyi performans gösterdiği belirlenmiştir ve Lojistik Regresyon kullanılması uygun görülmüştür.

Sonuçlar, kalp hastalığı riskinin doğru bir şekilde tahmin edilebileceğini göstermiştir. Bu çalışma, sağlık sektöründe kalp hastalığı riski tahmini için kullanılabilecek bir sistem geliştirmek amacıyla önemli bir adım olmuştur. Elde edilen sonuçlar, kullanıcılara arayüz aracılığıyla sunulabilir ve kalp hastalığı riski olan bireylerin erken teşhis ve tedaviye erişimini kolaylaştırabilir.

Sonuç olarak, bu çalışma kalp hastalığı tespiti konusunda önemli bir araştırma projesinin sonucunu sunmaktadır. Elde edilen sonuçlar, gelecekteki çalışmalara ilham kaynağı olabilir ve sağlık alanında kalp hastalığı riski tahmini sistemlerinin geliştirilmesine katkı sağlayabilir.

KAYNAKLAR

1. Heart, <https://www.heart.org/>.
2. Kaggle, <https://www.kaggle.com/datasets/cherngs/heart-disease-cleveland-uci>.
3. YouTube Tutorial, <https://www.youtube.com/watch?v=qmqCYC-MBQot=285s>.
4. Flask Tutorial, <https://flask.palletsprojects.com/en/2.3.x/>.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı, Soyad : Mahir Barış Sümer
Uyruğu : Türkiye(T.C)
Doğum Tarihi ve Yeri : 03.02.2000 - KAYSERİ
Telefon : 0 541 410 79 17
Belgegeçer : 0 352 555 88 77
E-posta : 1030520887@erciyes.edu.tr
Adres : Alpaslan Mah.
Kızılırmak Cad.
38039, Melikgazi KAYSERİ TÜRKİYE

EĞİTİM

Derece	Kurum	Mez.Yılı
Lise	Mevsim Temel Lisesi, KAYSERİ	2017
Ortaokul	Besime Özderici Ortaokulu, KAYSERİ	2014

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı, Soyadı : Aydan Alkaya
Uyruğu : Türkiye (T.C.)
Doğum Tarihi ve Yeri : 08.10.2000 - KAYSERİ
Telefon : 0 538 094 03 16
Belgegeçer : 0 352 555 88 77
E-posta : 1030520852@erciyes.edu.tr
Adres : Germir Mah.
425. Sok.
38030, Melikgazi KAYSERİ TÜRKİYE

EĞİTİM

Derece	Kurum	Mez.Yılı
Lise	Verimli Temel Lisesi, KAYSERİ	2018
Ortaokul	Hacı Mustafa Gazioğlu Ortaokulu, KAYSERİ	2014