**KARDİYOVASKÜLER HASTALIKLARIN TEŞHİSİ İÇİN YAPAY ZEKÂ SINIFLANDIRICI MODELİ**

**Projenin Amacı:**  
Projem, kardiyovasküler hastalıkların erken teşhisini yapabilecek bir yapay zekâ modeli (yapay sinir ağı) geliştirmeyi amaçlamaktadır. Kardiyovasküler hastalıklar, dünya çapında ölüm nedenlerinin başında gelmektedir ve yıllık yaklaşık 17.9 milyon kişinin ölümüne yol açmaktadır. Bu hastalıklar, erken teşhis ve uygun yönetimle önlenebilir veya etkileri azaltılabilir. Oluşturulacak makine öğrenmesi modeli, bireylerin hastalık risklerini tahmin ederek erken müdahaleye olanak sağlayacaktır. Ayrıca proje içerisinde hem öğrenme amaçlı hem de amacımıza ve veri setimize en uygun yapay sinir ağı ayarlarını bulmak için eğitim birkaç farklı şekilde gerçekleştirilip, karşılaştırmalar yapılmaktadır.

**Sınıflandırma Amacı:**  
Veri seti, bireylerin mevcut sağlık durumları ve risk faktörleri doğrultusunda kalp hastalığına sahip olup olmadıklarını (HeartDisease sınıfı: 1 veya 0) sınıflandırmak için kullanılacaktır.

**Veri Seti:**  
Heart Failure Prediction Dataset, kardiyovasküler hastalık tahmini için hazırlanmış ve çeşitli bağımsız veri setlerinin birleştirilmesiyle oluşturulmuş bir veri setidir. Toplamda 5 farklı veri seti kullanılarak, 11 ortak özellik üzerinde 918 örnek içeren en büyük kalp hastalığı veri setlerinden biri oluşturulmuştur. Veri seti meşhur veri bilimi zinciri Kaggle platformunda sunulmuştur ve UCI Makine Öğrenimi Deposu'ndaki farklı veri setlerinden derlenmiştir.

**Özellikler:**  
Veri seti, hastalara ait aşağıdaki özelliklerden oluşmaktadır:

1. **Age:** Hastanın yaşı (yıl).
2. **Sex:** Hastanın cinsiyeti (M: Erkek, F: Kadın).
3. **ChestPainType:** Göğüs ağrısı türü (TA: Tipik Angina, ATA: Atipik Angina, NAP: Anginal Olmayan Ağrı, ASY: Asemptomatik).
4. **RestingBP:** Dinlenme kan basıncı (mm Hg).
5. **Cholesterol:** Serum kolesterol düzeyi (mg/dl).
6. **FastingBS:** Açlık kan şekeri (1: >120 mg/dl, 0: ≤120 mg/dl).
7. **RestingECG:** Dinlenme elektrokardiyogram sonuçları (Normal, ST, LVH).
8. **MaxHR:** Elde edilen maksimum kalp hızı (60-202 arası sayısal değer).
9. **ExerciseAngina:** Egzersiz kaynaklı anjina (Y: Evet, N: Hayır).
10. **Oldpeak:** ST segment depresyonu (sayısal değer).
11. **ST\_Slope:** Egzersiz ST segmenti eğimi (Up: Yükselen, Flat: Düz, Down: Düşen).

**Hedef Sınıf Bilgisi:**

* **HeartDisease:** Kalp hastalığın olup olmamasi (1: Kalp hastalığı, 0: Normal).

**Örnek Sayısı:**  
Toplam 918 örnek bulunmaktadır. Veri seti, aşağıdaki veri kaynaklarından derlenmiştir:

* **Cleveland:** 303 gözlem
* **Hungarian:** 294 gözlem
* **Switzerland:** 123 gözlem
* **Long Beach VA:** 200 gözlem
* **Stalog (Heart) Data Set:** 270 gözlem

Bu gözlemlerin 272'si tekrarlanan veriler olduğundan çıkarılmıştır.

**Uygulama hakkında acıklama:**

**1.Eğitim setini aynı zamanda test verisi olarak kullanarak eğitim**

1. Bulunan en uygun parametre değerleri :

**Öğrenme Hızı:** 0.001

**Epoch Sayısı:** 500 epoch sayısı için %99.02 doğruluk elde edilmiştir.

**Batch Size:** 16

**Aktivasyon Fonksiyonları:** Giris katmani : Relu, Gizli katman : Relu, Cikis katmani : Sigmoid

**Optimizasyon Algoritması:** Adam optimizasyon algoritması kullanılıyor. Adam (Adaptive Moment Estimation) optimizatörü, derin öğrenmede sık kullanılan ve etkili bir optimizasyon yöntemidir.

1. En uygun ağ topolojisi :

**Giriş Katmanı :** 16 nöron var.

**Gizli Katman :** 1 gizli katman, 16 nöron var.

**Çıkış Katmanı :** Modelin sınıflandırma yaptığı ve tek bir hedef degişkenimiz olduğu için çıktı katmanı tek bir nörondan oluşur.

1. Başarı sonuçları :

50 epoch : 91,18 % dogruluk, 24.11 % kayip

100 epoch : 95,21% dogruluk, 15.82% kayip

500 epoch : 99.13%, 2.58% kayip

1. Konfüzyon matrisi

[[403 7]

[ 1 507]]

Classification Report:

precision recall f1-score support

0 1.00 0.98 0.99 410

* 1. 0.99 1.00 0.99 508

**2. 5 fold cross validation kullanarak eğitim**

1. Bulunan en uygun parametre değerleri :

**Öğrenme Hızı:** 0.001

**Epoch Sayısı:** 100 epoch sayısı için %86,27 ortalama doğruluk elde edilmiştir.

**Batch Size:** 16

**Aktivasyon Fonksiyonları:** Giris katmani : Relu, Gizli katman : Relu, Cikis katmani : Sigmoid

**Optimizasyon Algoritması:** Adam optimizasyon algoritması kullanılıyor. Adam (Adaptive Moment Estimation) optimizatörü, derin öğrenmede sık kullanılan ve etkili bir optimizasyon yöntemidir.

1. En uygun ağ topolojisi :

**Giriş Katmanı :** 16 nöron var.

**Gizli Katman :** 1 gizli katman, 16 nöron var.

**Çıkış Katmanı :** Modelin sınıflandırma yaptığı ve tek bir hedef degişkenimiz olduğu için çıktı katmanı tek bir nörondan oluşur.

1. Başarı sonuçları (ortalamalar):

50 epoch : 85,51% dogruluk, 35,42% kayip

100 epoch : 86,27% dogruluk, 38,03% kayip

200 epoch : 85,29% dogruluk, 49,05% kayip

1. Konfüzyon matrisleri :

Notebook’ta yer almaktadır.

**3. 10 fold cross validation kullanarak**

1. Bulunan en uygun parametre değerleri :

**Öğrenme Hızı:** 0.001

**Epoch Sayısı:** 30 epoch sayısı için %86,60 doğruluk elde edilmiştir.

**Batch Size:** 16

**Aktivasyon Fonksiyonları:** Giris katmani : Relu, Gizli katman : Relu, Cikis katmani : Sigmoid

**Optimizasyon Algoritması:** Adam optimizasyon algoritması kullanılıyor. Adam (Adaptive Moment Estimation) optimizatörü, derin öğrenmede sık kullanılan ve etkili bir optimizasyon yöntemidir.

1. En uygun ağ topolojisi :

**Giriş Katmanı :** 16 nöron var.

**Gizli Katman :** 1 gizli katman, 16 nöron var.

**Çıkış Katmanı :** Modelin sınıflandırma yaptığı ve tek bir hedef degişkenimiz olduğu için çıktı katmanı tek bir nörondan oluşur.

1. Başarı sonuçları (ortalamalar):

30 epoch : 86,60% dogruluk, 33,65% kayip

50 epoch : 86,49% dogruluk, 33,70% kayip

100 epoch : 85,73% dogruluk, 38,51% kayip

1. Konfüzyon matrisleri :

Notebook’ta yer almaktadır.

**4.%66-%34 eğitim test ayırarak (5 farklı rassal ayırma ile)**

1. Bulunan en uygun parametre değerleri :

**Öğrenme Hızı:** 0.001

**Epoch Sayısı:** 100 epoch sayısı için ortalama %86,01 doğruluk elde edilmiştir.

**Batch Size:** 16

**Aktivasyon Fonksiyonları:** Giris katmani : Relu, Gizli katman : Relu, Cikis katmani : Sigmoid

**Optimizasyon Algoritması:** Adam optimizasyon algoritması kullanılıyor. Adam (Adaptive Moment Estimation) optimizatörü, derin öğrenmede sık kullanılan ve etkili bir optimizasyon yöntemidir.

1. En uygun ağ topolojisi :

**Giriş Katmanı :** 16 nöron var.

**Gizli Katman :** 1 gizli katman, 16 nöron var.

**Çıkış Katmanı :** Modelin sınıflandırma yaptığı ve tek bir hedef degişkenimiz olduğu için çıktı katmanı tek bir nörondan oluşur.

1. Başarı sonuçları (ortalamalar):

50 epoch : 85, 56% dogruluk, 37, 86% kayip

100 epoch : 86,01% dogruluk, 37, 68% kayip

30 epoch : 85, 91% dogruluk, 36, 35 kayip

1. Konfüzyon matrisleri :

Notebook’ta yer almaktadır.