

Karaciğerde Oluşan Hastalıkların Tespitinde Makine Öğrenmesi Yöntemlerinin Kullanılması

1



Emre DANDIL

Bilecik Ş. Edebali Üniversitesi
emre.dandil@bilecik.edu.tr

+90228 214 1613



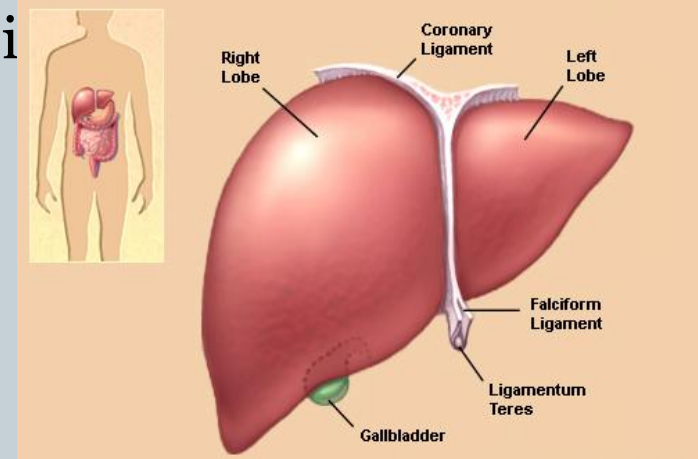
- **Özet**
- **Giriş ve Literatür Bilgileri**
 - Amaç, kapsam...
- **Materyal ve Yöntem**
 - Kullanılan dataset(seri kümesi)
 - Yapay sinir ağları
 - ✦ Çok katmanlı ağ yapısı(BPA Alg.)
 - Bayes sınıflandırma algoritması
- **Deneysel Bulgular**
 - YSA ve Bayes Yöntemlerinin Değerlendirilmesi
- **Sonuç ve Değerlendirme**
 - Çıkarımlar, Yorumlamalar...
 - Gelecekte neler yapılabilir?
- **Kaynaklar**



Ne Yapıldı (Özetçe)?

3

- Bu çalışmada karaciğerde hastalık teşhisi için **Makine Öğrenmesi** tabanlı bir sınıflandırma yazılımı geliştirilmiştir.
- Sistem ile karaciğerde oluşan hastalıkların teşhisi için C# programlama dili ile **Yapay Sinir Ağları(YSA)** tabanlı bir yazılım tasarlanmış ve sınıflandırma sonuçları **Bayes yöntemi** ile karşılaştırılmıştır.
- Çalışmada bir veri setindeki karaciğer hastası olan ve olmayan insanlardan alınan sayısal veriler sınıflandırılmıştır. Ayrıca yazılımın performansı da Matlab yazılımı ile doğrulanmıştır.

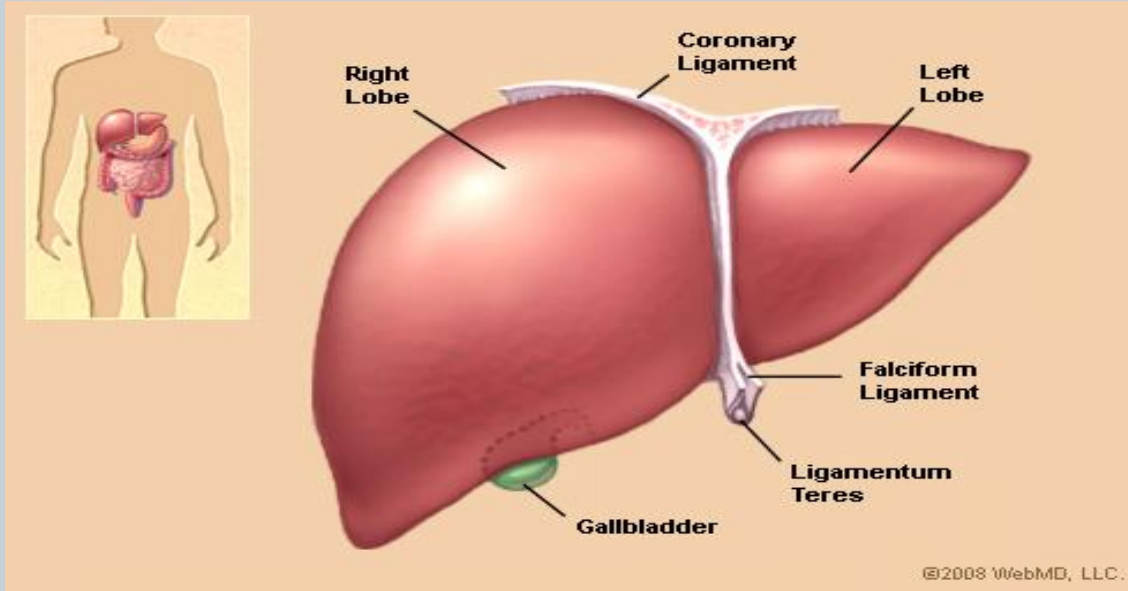


	Bulunan Sonuçlar	İstenilen Sonuçlar	Hata Miktarları
0.1 1 1 1 1 1 0.7 1 1 1 1 0.5 1			
0 0... 0.3 0... 0... 0... 0... 1 1 1 1 1 0.6 1			
0.3 1 1 1 1 1 0.8 0.6 0... 0.5 0.5 0.85 1 0.9 0.1			
0 0... 0... 0... 0.9 0.6 0... 0... 0... 1 1 1 1 1 0.6 0.1			
0... 1 1 1 1 1 0... 1 0... 0.75 0.75 0.75 0.8 0.75 0.5			
0... 1 1 1 1 1 0... 0.7 0... 0.75 0.75 0.75 0.8 0.45 0.5			
0... 1 1 1 1 1 0.9 0... 0.4 0.7 0.5 0.5 0.5 0.9 0.6 0.1			
0 0... 1 1 1 1 0.6 0... 0.2 1 0.75 0.75 0.75 0.9 0.75 0.1			
0 0.3 1 1 1 1 0.5 0.4 0... 0.5 0.5 0.5 0.55 0.9 0.1			
0.1 0.9 1 1 1 1 0... 0... 0... 0.5 0.5 0.5 0.75 0.6 0.5			
0 0... 0... 0.9 1 0... 0 0... 0.8 0.75 0.75 0.75 0.7 0.5			
0 0.3 1 1 1 1 0.2 0.4 1 1 1 1 1 0.9 0.5			
0 0... 1 1 1 1 0... 0... 0.85 0.75 0.75 0.95 0.7 0.1			
0... 1 1 1 1 1 0... 0... 0.8 0.15 0 0.6 0.75 0.7 1			
0.3 1 1 1 1 1 0... 0... 0... 0.5 0.5 0.95 1 0.6 0.5			
	Giriş Ağından	Çıkış Ağından	Ortalama Hata Miktarı
			Genel Hata Miktarı
			Ağrıklık Kaydet
			Ağrıklık Al

Giriş ve Literatür Bilgileri

4

- *Karaciğer metabolizmanın düzenlenmesi, kırmızı kan hücrelerinin dağıtılması* gibi benzer hayati fonksiyonları yerine getirdiği için insan vücudunun en önemli organlarından bir tanesidir .
- Karaciğer hastalıkları vücutta hem başka hastalıklara neden olabilmekte, hem de başlı başına vücut için büyük tehlikeler oluşturabilmektedir.

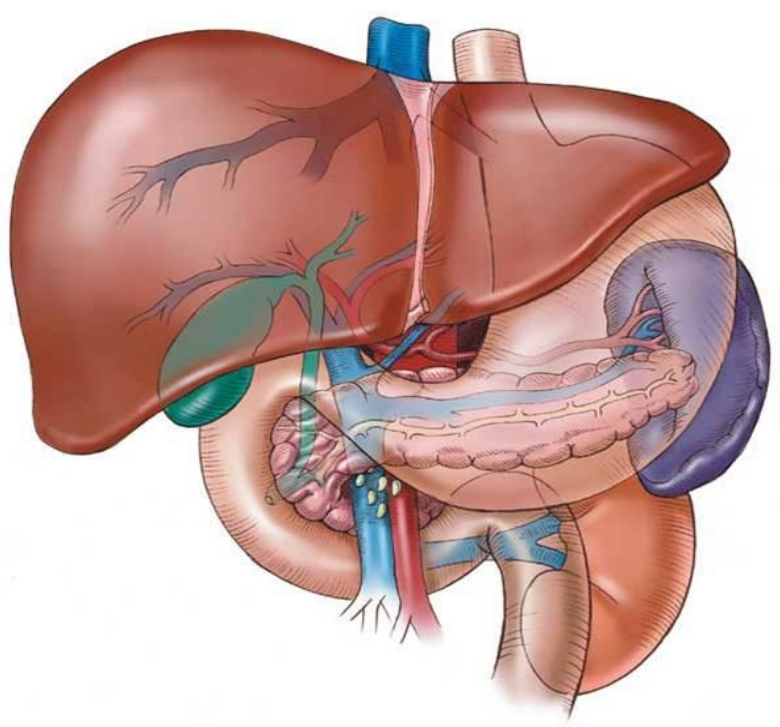


İnsanda karaciğer organı

- Bu sebeplerden ötürü **karaciğer kanseri, siroz, karaciğer tümörü ve HCC** gibi karaciğer hastalıklarının insan vücudu için erken bir aşamada teşhisi ve tedavisi hayati öneme sahiptir.

Giriş ve Literatür Bilgileri

5

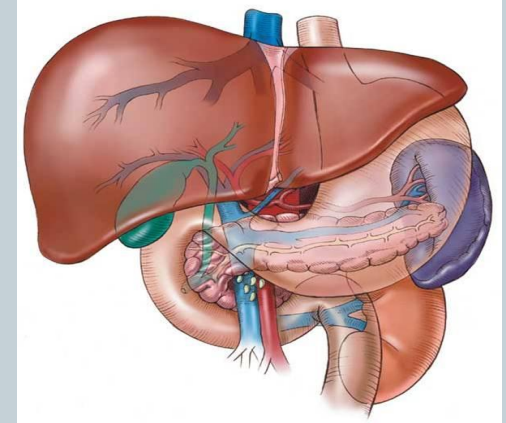


Karaciğer ve yardımcı organlar

- Ancak **hız, doğruluk ve kesin teşhis sonuçları** gibi parametrelere bağlı olarak otomatik sınıflandırma yöntemlerinin kullanılmasına ihtiyaç duyulması biyomedikalde tıbbi bilişim ihtiyacı olarak ortaya çıkmıştır.

- Karaciğer hastalıklarının **ilk zamanlarda yeterli belirtileri olmamasına** karşın, ilerleyen aşamalarda **belirgin olarak** ortaya çıkabilmektedir
- Geleneksel olarak karaciğer hastalıklarının tıbbi olarak teşhis edilmesinde **kanda oluşan enzim düzeyleri ölçümleri yapılır** ve bu sonuçlara göre karar verilir.

- *Karaciğerdeki hastalıkların esnek hesaplama yöntemleri de denilen makine öğrenmesi yöntemleri ile tespitinde günümüze kadar birçok çalışma yapılmıştır:*
- **Destek Vektör Makineleri(DVM)** ile kimyasal karaciğer verilerinin sınıflandırılması, **Naive Bayes ve DVM** yöntemlerinin performans karşılaştırması, **yapay hibrit yöntemler** ile karaciğer hastalık tiplerinin sınıflandırılması gibi benzer çalışmalar literatürde kullanılmıştır.
- Bu çalışmada ise, karaciğerde oluşan hastalıkların teşhisi için literatürde en çok tercih edilen esnek hesaplama yöntemlerinden birisi olan **yapay sinir ağları tabanlı** bir sınıflandırma sistemi kullanılmış Bayes yöntemi ile doğruluk yüzdesi bakımından karşılaştırılmıştır.



Karaciğer

- **Veri Kümesi(Dataset)**
- Bu çalışmada kullanılan veriler için **UCI Machine Learning Respository** veritabanında bulunan karaciğer hastalığı veri kümesi(*Liver Disease Dataset*) kullanılmıştır. (2011 yılına ait datalar)
- Bu veri kümesindeki datalar 416 karaciğer hastası ve 167 hasta olmayan toplamda 583 adet gerçek deneklerden alınmıştır.

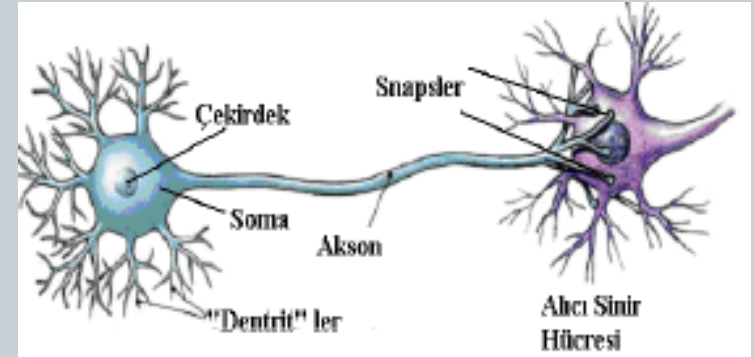
Özellik No	Özellik Bilgisi
1	Hastanın yaşı
2	Hastanın cinsiyeti
3	Toplam bilirubin
4	Direct bilirubin
5	Indirect fosfat
6	Alemin aminotransferaz
7	Aspartat aminotransferaz
8	Toplam protein
9	Albumin
10	Albumin ve Globulin oranı

Veri tabanından alınan datalara ait sayısal özellikler

• Makine Öğrenmesi?

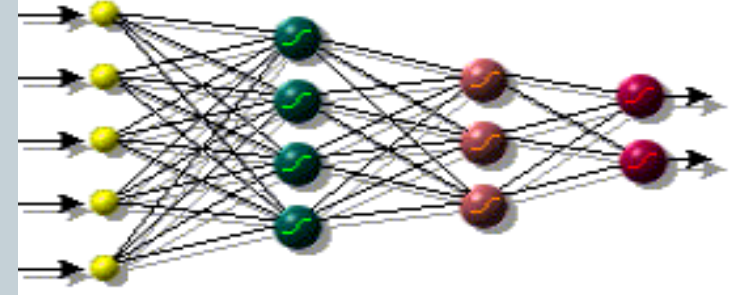
- Yapay zekanın bir dalı olan **makine öğrenmesi**; temelini öğrenmeden alarak ve sayısal dataları kullanarak **örüntü tanıma, sınıflandırma, kümeleme** gibi temel işlemlerin bir bilgisayar(makine) tarafından da yapılabildiğini gösteren yapıdır.
 - Danışmanlı öğrenme(supervised-learning)
 - Danışmansız öğrenme(unsupervised-learning)
 - Destekleyici(Yarışmacı) öğrenme(reinforcement-learning)
 - Birleştirilmiş öğrenmeler(union-learning)
- Bazı makine öğrenmesi yöntemleri
 - ✦ **Yapay sinir ağları**, bulanık mantık, uzman sistemler, genetik algoritmalar, kümeleme (k-means,k-NN) yöntemleri, destek vektör makineleri, karar ağaçları, **Bayes yöntemi**, adaptif (ANFIS) mimariler v.b. olarak sınıflandırılabilir.

- **Yapay Sinir Ağları**
- Yapay Sinir Ağları, insan beyninin işleyişini taklit ederek yeni sistem oluşturmaya çalışan yapay zeka yaklaşımıdır.



Doğrusal olmayan sistem modelleme

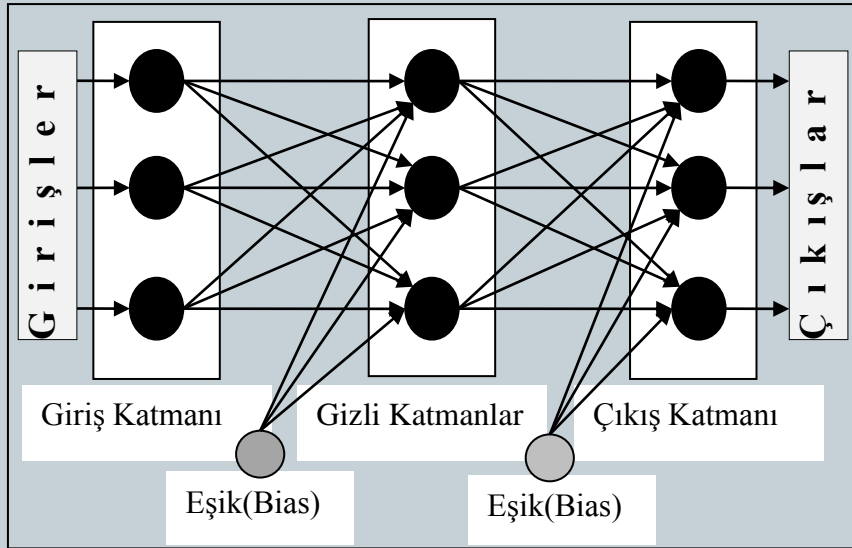
- Zeki kontrol
- Optimizasyon
- Sinyal işleme
- Örüntü tanıma
- Veri sıkıştırma
- Sinyal süzme
- Sınıflandırma



Sinir hücresi ve yapay sinir hücresi

Kullanım alanları ve işlemleri

• Yapay Sinir Ağları



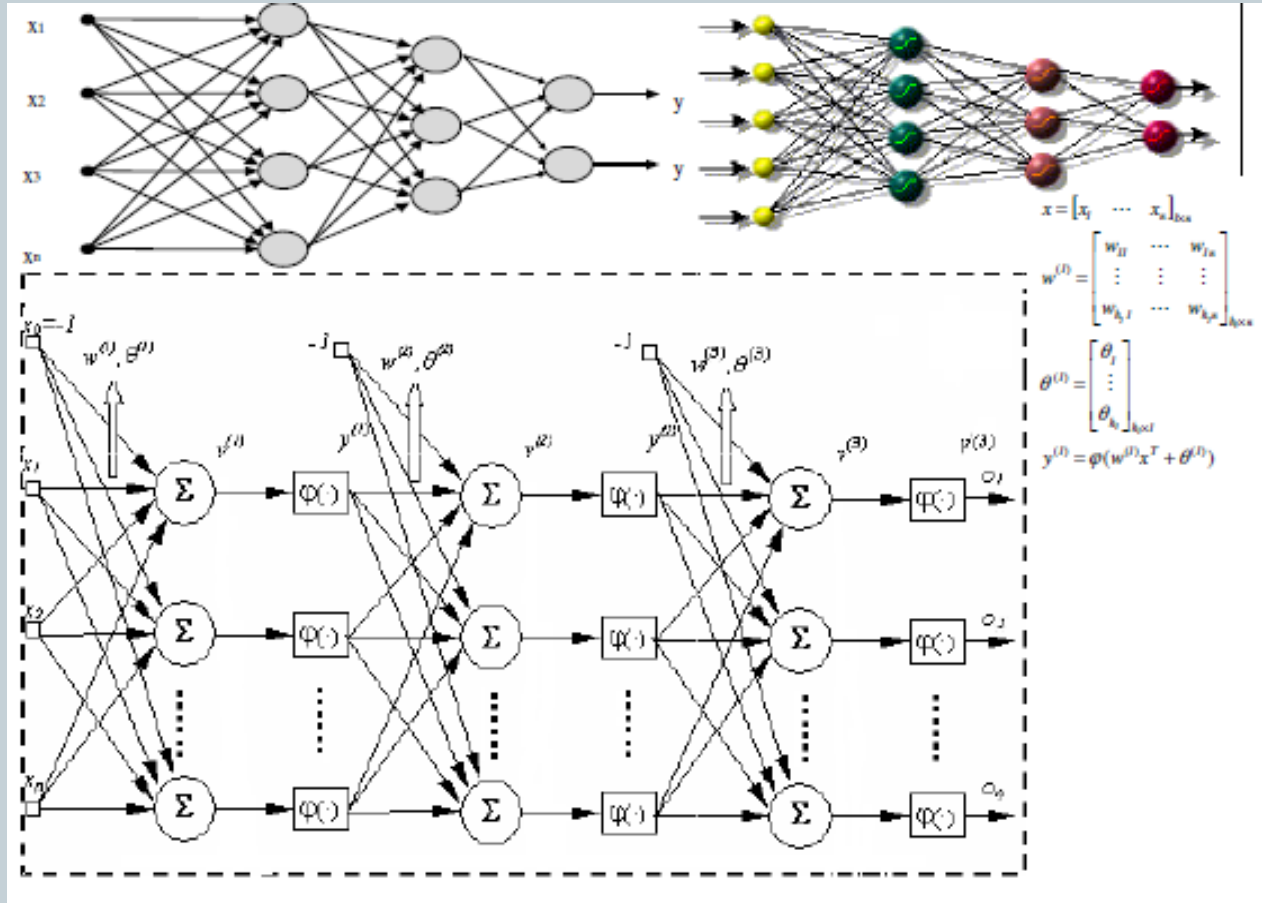
İleri beslemeli geri yayımlı ÇKA yapısı

- İLERİ BESLEMELİ YSA
 - TEK KATMANLI YSA
 - ÇOK KATMANLI YSA
- GERİ DÖNÜŞÜMLÜ YSA AĞI
 - ELMAN AĞI
 - SOM/KOHONEN AĞI
 - HOPFIELD AĞI
 - BOLTZMAN MAKİNASI
 - COUNTERPROPAGATION AĞI
 - COGNITRON AĞI
- RADYAL TABANLI YSA AĞLARI (RBF)
- OLASILIKSAL (Probabilistic) AĞLAR (PNN)
- Hücresel YSA
- Dinamik YSA

•Bu çalışmada da Çok Katmanlı Algılayıcı Modeli kullanılmıştır. İleri beslemeli geri yayımlı bir **Çok Katmanlı Ağ Modeli** şekilde gösterilmiştir.

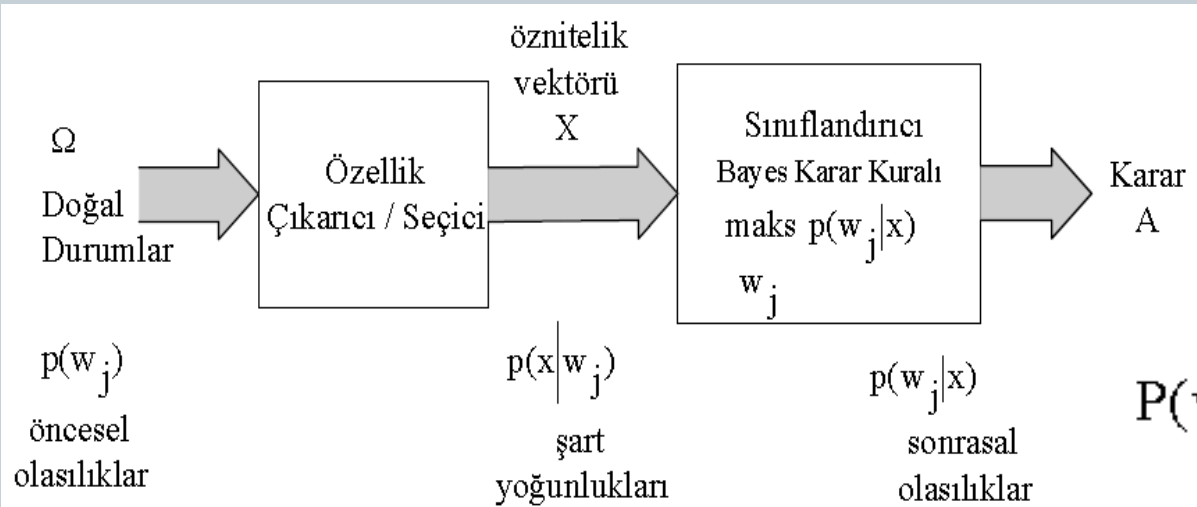
• YSA Back-propagation(geri yayılım) öğrenme algoritması

- YSA ağ yapısında, verilen bir girdi setine karşılık çıktı değerleri verilerek belirtilen öğrenme kuralına göre ağırlık değerleri otomatik olarak değiştirilmektedir.



• Bayes algoritması

- Bayes karar teorisi, örüntü sınıflandırma probleminin çözümü için temel istatistiksel bir yaklaşımı olup, karar problemini olasılık bazda değerlendirmekte ve öncesel bir olasılık değerden sonrasal bir olasılık değer elde etme işlemi olarak tanımlanmaktadır.




$$P(w_k | x) = \frac{P(w_k) \cdot p(x | w_k)}{p(x)}$$

Bayes teoremi

Gerçekleştirilen Yazılım

13


Girişleri Yönet

 Açık

Rakamlar: xds

A1:BG200

Sayfa 1

 Değerleri Al

Normalize Et

Yönetim Ekranı



Öğrenme Hızı: 0.3 Bias: 1


İstenilen Hata: 0.0001

Maximum Adım: 500

Activasyon Fonk.: Sigmoid

Eğitime Katılacak Örnek: 200

 Ağı Eğit  1. Adım

 Test Et

0,1	1	1	1	1	1	0...	0,7	1	1	1	1	1	0,5	1
0	0...	0,3	0...	0...	0...	0...	0...	1	1	1	1	1	0,6	1
0,3	1	1	1	1	1	0,8	0,6	0...	0,5	0,5	0,85	1	0,9	0,6
0	0...	0...	0...	0,9	0,6	0...	0...	0...	1	1	1	1	0,6	0,7
0...	1	1	1	1	1	0...	1	0...	0,75	0,75	0,75	0,8	0,75	0,9
0...	1	1	1	1	1	0...	0,7	0...	0,75	0,75	0,75	0,8	0,45	0,9
0...	1	1	1	1	0,9	0...	0,4	0,7	0,5	0,5	0,5	0,9	0,6	0,6
0	0...	1	1	1	0,6	0...	0,2	1	0,75	0,75	0,75	0,9	0,75	0,6
0	0,3	1	1	1	1	0,5	0,4	0...	0,5	0,5	0,5	0,55	0,9	0,8
0,1	0,9	1	1	1	0...	0...	0...	0...	0,5	0,5	0,5	0,75	0,6	0,9
0	0...	0...	0,9	1	0...	0	0...	0,8	0,75	0,75	0,75	0,75	0,7	0,9
0	0,3	1	1	1	1	0,2	0,4	1	1	1	1	1	0,9	0,9
0	0...	1	1	1	1	0...	0...	0...	0,85	0,75	0,75	0,95	0,7	0,6
0...	1	1	1	1	1	0...	0...	0,8	0,15	0	0,6	0,75	0,7	1
0,3	1	1	1	1	0...	0...	0...	0...	0,5	0,5	0,95	1	0,6	0,9

Bulunan Sonuçlar

İstenilen Sonuçlar


Hata Miktarları


Giriş Ağırlıkları

Çıkış Ağırlıkları

Ortalama Hata Miktarı

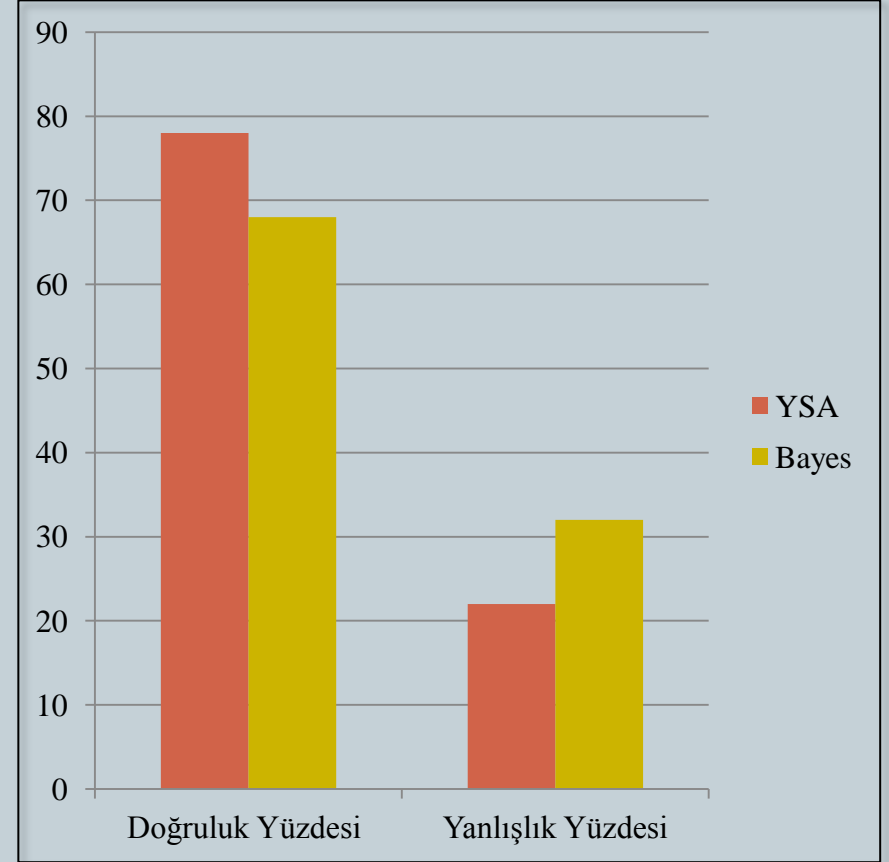
Genel Hata Miktarı

 Ağırlıkları Kaydet

 Ağırlıkları Al

- Gerçekleştirilen YSA yapısında **2** gizli katman ve her katmandaki nöron sayısı **10** olarak belirlenmiştir.
- Öğrenme yöntemi olarak ÇKA yapısında **geri yayılım(back-propagation)** algoritması kullanılmıştır.
- Bunlara ek olarak ağın performansı da **ortalama karesel hata(MSE)** kurallarına göre bulunmuştur.
- Çalışmada kullanılan veri kümesinde 416 adet karaciğer hastası ve 116 adet de karaciğer hastası olmayan deneklere ait toplamda 583 örnek vardır.
- Bu dataların **%70'** i YSA' nın eğitim aşamasında, geri kalan **%30'** luk kısmı ise test verisi olarak kullanılmıştır.

- Analiz sonucunda doğru tespit yüzdesi **Yapay Sinir Ağları yöntemi ile %82** olurken, yanlış tespit yüzdesi ise %18 olarak hesaplanmıştır. Ayrıca yapılan çalışma sınıflama tekniklerinden olan Bayes teoremi ile de kıyaslanmıştır. **Bayes teoremi ile %68 doğruluk yüzdesi, %32 yanlışlık yüzdesi** elde edilmiştir. Bu sonuçlara göre **YSA ile daha başarılı performans sonuçlarına ulaşıldığı** görülmüştür.



- Bu çalışmada, karaciğer hastalığının teşhisinde yardımcı olacak bir YSA tabanlı sınıflandırma geliştirilmiştir. Başarı değerlendirmesinde sınıflandırma yöntemlerinde "**doğruluk sınıflandırma yüzdesi**" performans ölçümleri kullanılmıştır.
- Veri kümesindeki test başarımlarında YSA yöntemi ile **%82** doğruluk yüzdesi elde edilirken, Bayes yöntemi ile doğruluk yüzdesi **%68** oranında hesaplanmıştır. Elde edilen sonuçlar mevcut veri kümesi ile karaciğer hastalığının teşhisinde YSA sınıflandırıcıların eğitim ve test başarımları açısından başarılı sonuçlar verdiği gözlenmiştir.
- Çalışmada daha sonra sistem başarımlarını yükseltmek için daha farklı ön işleme teknikleri kullanılabilir ve farklı sınıflandırma yöntemleri ile bu çalışmanın başarımlarını değerlendirmesi yapılabilir.





Yaralanılan Kaynaklar

17

1. Lin, R. H. ve Chuang, C. H., " A hybrid diagnosis model for determining the types of the liver disease", **Computers in Biology and Medicine**, 40, 665–670. 2010.
2. Parkin, D.M., Bray, F. ve Ferlay, J., "Global cancer statistics 2002" , **CA: A Cancer Journal for Clinicians** 55, 74–108, 2005.
3. Maddrey, W. C. ve Sorrell, M. F., Eugene, S. R., Schiff's Diseases of the Liver, **Lippincott Williams & Wilkins**, 10th Edition Copyright, 2007.
4. Sorich, M. J. vd., "Comparison of linear and nonlinear classification algorithms for the prediction of drug and chemical metabolism by human UDP-Glucuronosyltransferase Isoforms" , **J. Chem. Inf. Comput. Sci.**, 43, 2019-2024, 2003.
5. Huang, L.C. vd., " A comparison of classification methods for predicting Chronic Fatigue Syndrome based on genetic data" , **Journal of Translational Medicine**, 7:81, doi:10.1186/1479-5876-7-8, 2009.
6. A. Frank ve A. Asuncion, UCI Machine Learning Repository, University of California, **School of Information and Computer Science**, 2010. [<http://archive.ics.uci.edu/ml>].
7. C. Öz, R. Köker, S. Çakar, "Yapay Sinir Ağları ile Karakter Tabanlı Plaka Tanıma", **Elektrik Elektronik ve Bilgisayar Mühendisliği Sempozyumu (ELECO'2002)**, Bursa, 2002.
8. Çevik, K. K. ve Dandıl, E., "Yapay Sinir Ağları İçin .Net Platformunda Görsel Bir Eğitim Yazılımının Geliştirilmesi", **Gazi Üniversitesi Bilişim Teknolojileri Dergisi**, Cilt: 5, Sayı: 1, Ocak 2012.
9. Ş. Sağıroğlu, E. Beşdok, M. Erler, Mühendislikte Yapay Zeka Uygulamaları-1 Yapay Sinir Ağları, **Ufuk Yayıncılık**, Kayseri, 2003.
10. E. Öztemel, Yapay Sinir Ağları, **Papatya Yayınevi**, İstanbul, 2003.



Dinlediđiniz için TEŞEKKÜRLER!!!

18

- Sorular ve Cevaplar...????