

Karaciğerde Oluşan Hastalıkların Tespitinde Makine Öğrenmesi Yöntemlerinin Kullanılması

Emre Dandil

Bilecik Şeyh Edebali Üniversitesi, Meslek Yüksekokulu, Bilgisayar Programcılığı Bölümü, Bilecik
emre.dandil@bilecik.edu.tr

Özet: Zararlı gazların vücuda yayılması, aşırı ve sürekli alkol alma, sindirimi zor besinlerin tüketimi ve bazı tıbbi ilaçların kullanımı gibi etmenler karaciğerde hastalık oluşmasına neden olabilmektedir. İnsan vücudu için hayati derecede öneme sahip karaciğerde oluşan hastalıkların erken bir aşamada teşhisi oldukça önemlidir. Bu çalışmada da karaciğerde oluşan hastalıkların teşhisi için Yapay Sinir Ağları(YSA) tabanlı bir sistem geliştirilmiştir. Sistem ile bir veri setindeki karaciğer hastası olan ve olmayan insanlardan alınan veriler başarılı olarak sınıflandırılmıştır. Ayrıca YSA sonuçları Bayes Sınıflandırma yöntemi ile de bulunan sonuçlar başarılı performansı açısından karşılaştırılmıştır.

Anahtar Sözcükler: Karaciğer, hastalık teşhisi, sınıflandırma, yapay sinir ağları, bayes

Liver Disease Diagnosis Based on Artificial Neural Network

Abstract: The spread of harmful gases into the body, making excessive and chronic alcohol, consumption of foods difficult to digest and some of factors such as the use of medical drugs can cause liver disease. Diagnosis of the liver disease at an early stage is important for human body. In this study, a system based on artificial neural network(ANN) is developed for diagnosis of liver disease. The liver disease dataset have been classified as liver disease or not. Also, results of ANN have been compared with Bayes classification results.

Keywords: Liver, disease diagnosis, classification, artificial neural network, bayes

1.Giriş

Karaciğer metabolizmanın düzenlenmesi, kırmızı kan hücrelerinin dağıtılması gibi benzer hayati fonksiyonları yerine getirdiği için insan vücudunun en önemli organlarından bir tanesidir[1]. Karaciğer hastalıkları vücutta hem başka hastalıklara neden olabilmekte, hem de başlı başına vücut için büyük tehlikeler oluşturabilmektedir. 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[2]. Son zamanlarda farklı sınıflandırma yöntemlerinin biyomedikal hastalıkların teşhisinde kullanılması oldukça yaygın bir hale gelmiştir. Bunlara sebep olarak en geçerli açıklama ise bilgisayar tabanlı sınıflandırma yöntemleri ile doğruya yakın sonuç elde etme olacaktır.

Karaciğer hastalıklarının ilk zamanlarda yeterli belirtileri olmamasına karşın, ilerleyen aşamalarda belirgin olarak ortaya çıkabilmektedir[1]. Geleneksel olarak karaciğer hastalıklarının tıbbi olarak teşhis edilmesinde kanda oluşan enzim düzeyleri ölçümleri yapılır ve bu sonuçlara göre karar verilir[2]. Ancak hız, doğruluk ve kesin teşhis sonuçları gibi parametrelere bağlı olarak otomatik sınıflandırma yöntemlerinin kullanılması da ihtiyaç olmuştur. Bu alanda Destek Vektör Makineleri(DVM) ile kimyasal karaciğer verilerinin sınıflandırılması[3], Naive Bayes ve DVM yöntemlerinin performans karşılaştırması[4], yapay hibrit yöntemler ile karaciğer hastalık tiplerinin

sınıflandırılması[1] 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 sistem geliştirilmiştir. Bu sistem sayesinde insanlardan alınan karaciğerle ilgili veriler ile karaciğer hastalığı teşhisi yapılmıştır. Elde edilen performans sonuçları Bayes sınıflandırma algoritması ile karşılaştırılarak, YSA' da daha başarılı bir sınıflandırma yapıldığı gözlenmiştir.

Çalışmanın ikinci bölümünde çalışmada kullanılan materyal ve yöntemler tanıtılmış, üçüncü bölümünde deneysel sonuçlar incelenmiş ve dördüncü bölümünde elde edilen sonuçlar karşılaştırılarak bildiri tamamlanmıştır.

2.Materyal ve Yöntem

Gerçekleştirilen bu çalışmada YSA yöntemi ile karaciğer hastalık teşhisi yapılmıştır ve bulunan sonuçlar başka bir sınıflandırma yöntemi olan Bayes ile karşılaştırılmıştır. Karaciğer hastalarının teşhisi için kullanılan veriler UCI Machine Learning Repository[6] veritabanındaki "Liver Disease Dataset" veri kümesinden alınmıştır.

2.1.Veritabanı(Dataset)

Bu çalışmada kullanılan veriler için UCI Machine Learning Respository[6] veritabanında bulunan karaciğer hastalığı veri kümesi(Liver Disease Dataset) kullanılmıştır. Bu veri kümesindeki datalar 416 karaciğer hastası ve 167 hasta olmayan toplamda 583 adet gerçek deneklerden alınmıştır. Veri kümesinde toplam 10 özellik için denekler üzerinde ölçümler yapılmıştır ve bu özellikler Çizelge 1' de detaylı olarak verilmiştir.

Çizelge 1. Karaciğer hastalık veri kümesi

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

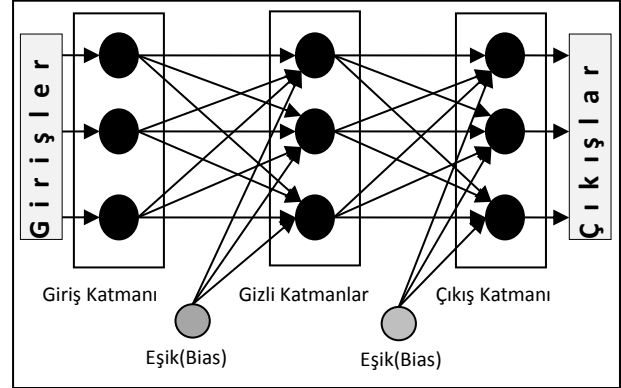
2.2.Yapay Sinir Ağları

Yapay Sinir Ağları, insan beyninin işleyişini taklit ederek yeni sistem oluşturmaya çalışan yapay zeka yaklaşımıdır[7]. Sınıflandırma problemlerinde çok tercih edilen yöntemlerden birisidir. İnsan beynindeki biyolojik sinir hücrelerinin yapısı temel alınarak YSA yapısı oluşturulur. YSA' da aynen beynimizde olduğu gibi öğrenme ve öğrenilen bilgilere göre karar verme mekanizmaları bulunur [8,9,10].

Günümüzde yapay sinir ağlarında Perceptron, Adaline, Çok Katmanlı Algılayıcı Modeli(Hatayı Geriye Yayıma), Vektör Kuantizasyon Problemleri, Hopfield Ağları ve Elman Ağları gibi bir çok öğrenme modeli vardır[10]. Genelde YSA ağ yapısında, verilen bir girdi setine karşılık çıktı değerleri verilerek belirtilen öğrenme kuralına göre ağırlık değerleri otomatik olarak değiştirilmektedir. Eğitim verisinin tamamlanmasından sonra eğitilmiş olan ağ, ağırlık değerlerinin son durumuna göre, verilen herhangi bir veri setinin sonucunu tahmin edebilmektedir. Bu çalışmada da Çok Katmanlı Algılayıcı Modeli kullanılmıştır. İleri beslemeli geri yayımlı bir Çok Katmanlı Ağ Modeli Şekil 1' de gösterilmiştir.

Çalışma kapsamında geliştirilen yapay sinir ağlarında bir karaciğer hastalığı teşhisi için sınıflandırma yapılmıştır. Yapay Sinir Ağlarında gerçekleştirilen sınıflandırma işlemi Matlab ortamında gerçekleştirilmiştir. Gizli katman sayısı 10 olarak belirlenmiştir. Öğrenme yöntemi olarak Levenberg-Marquardt algoritması kullanılmıştır. Bunlara ek olarak ağın performansı da ortalama karesel

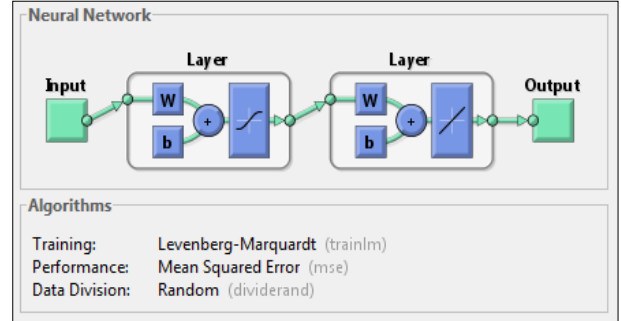
hata(MSE) ve Regresyon kurallarına göre bulunmuştur.



Şekil 1. YSA Çok Katmanlı Algılayıcı Modeli

2.3.Deneysel Çalışmalar

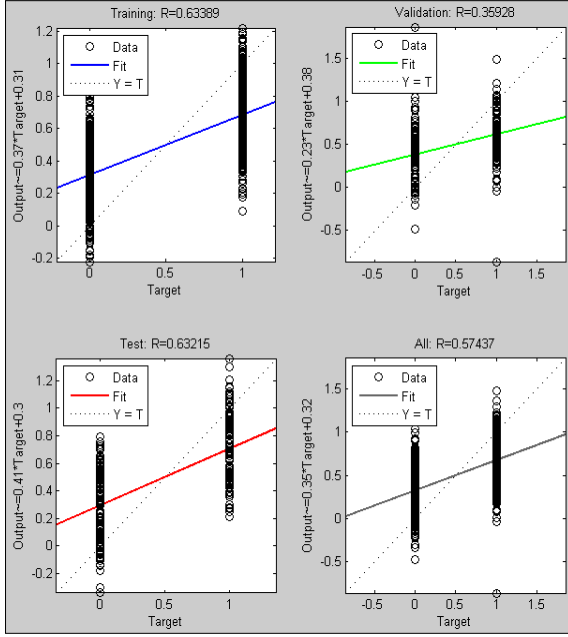
Çalışmada kullanılan veri kümesinde 416 adet karaciğer hastası ve 116 adet de karaciğer hastası olmayan deneklere ait toplamda 583 örnek vardır. Bu dataların %70' i YSA' nın eğitim aşamasında, geri kalan %30' luk kısmı ise test verisi olarak kullanılmıştır. Sistem için geliştirilen ağ yapısına ait bilgiler Şekil 2' de verilmiştir.



Şekil 2. Sistem için kullanılan ağ yapısı

Test işlemi sonucunda performans analizini gerçekleştirmek için "Doğruluk Yüzdesi" ve "Yanlışlık Yüzdesi" kriterleri esas alınmıştır. Çalışmada 1000 epoch için en başarılı sonuçlar elde edilmiştir. YSA için regresyon sonuçları Şekil 3'de gösterilmiştir.

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.



Şekil 3. YSA Regresyon Sonuçları

Her iki sınıflandırma yöntemi için 1000 epoch sonucunda elde edilen doğruluk ve yanlışlık yüzdeleri Çizelge 2' de gösterilmiştir.

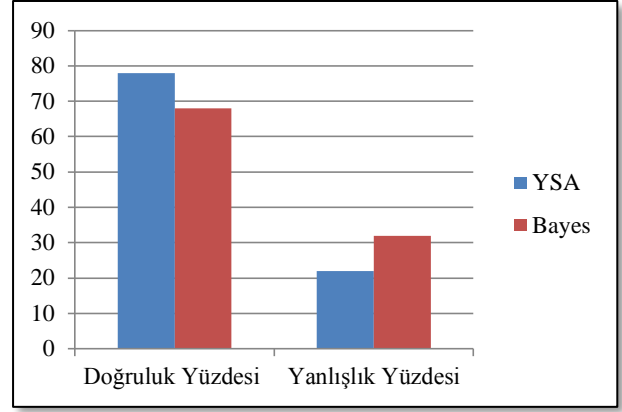
Çizelge 2. 1000 epoch için başarımlar yüzdeleri

Yöntem	Doğruluk Yüzdesi	Yanlışlık Yüzdesi
YSA	82	18
Bayes	68	32

YSA ve Bayes sınıflandırma yöntemlerinin başarımlar yüzdelerinin sonuçları ise Şekil 4' deki grafik üzerinde gösterilmiş ve performans kıyaslamaları yapılmıştır.

3.Sonuçlar

Bu çalışmada, karaciğer hastalığının teşhisinde yardımcı olacak bir YSA tabanlı sınıflandırma geliştirilmesi amaçlanmıştır. Başarı değerlendirilmesinde sınıflandırma yöntemlerinde "doğruluk yüzdesi" ve "yanlışlık 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 başarımlarını açısından başarılı sonuçlar verdiği gözlenmiştir.



Şekil 4. Sınıflandırma yöntemlerinin başarımlar yüzdeleri

Çalışmada daha sonra sistem başarımlarını yükseltmek için daha farklı önışleme teknikleri kullanılabilir ve farklı sınıflandırma yöntemleri ile bu çalışmanın başarımlar değerlendirilmesi yapılabilir.

4.Kaynaklar

- [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] 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] 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 Dandil, 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. Eriş, Mühendislikte Yapay Zeka Uygulamaları-1 Yapay Sinir Ağları, Ufuk Yayıncılık, Kayseri, 2003.
- [10] E. Öztemel, Yapay Sinir Ağları, Papatya Yayınevi, İstanbul, 2003.