

T.C.
NECMETTİN ERBAKAN ÜNİVERSİTESİ
SEYDİŞEHİR AHMET CENGİZ MÜHENDİSLİK
FAKÜTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

İRİS ÇİÇEĞİ TÜRÜNÜN YSA YÖNTEMLERİ
İLE TAHMİNİ

Nilay SOYTOP^{*1}, Emre ÇAVDAR^{*2}, Erkan KARYAĞDI^{*3}

¹soytopnilay@gmail.com, ²emrecvdr20@gmail.com, ³karyagdierkan@gmail.com

ÖZET

İRİS ÇİÇEĞİ TÜRÜNÜN YSA YÖNTEMLERİ
İLE TAHMİNİ

Bilindik veriler kullanılarak sınıflandırma yapmak, birçok alanda yaygın olarak kullanılmaktadır. Fakat gözlemlere dayalı sınıflandırma yapılırken bazı hatalar meydana gelebilmektedir. Bu hatalarını en aza indirmek için yapay öğrenme yöntemleri kullanılmaktadır. Bu çalışmada amaç, yapay sinir ağı (YSA) yöntemlerini kullanarak iris çiçeğinin türünün belirlenmeye çalışılmasıdır. Çalışmada UCI adlı sitedeki en popülerler arasında bulunan iris çiçeği verilerinden ve Matlab dan yararlanılmıştır. Üç katmanlı, ileri beslemeli, dört girişli ve bir çıkışlı YSA modeli kullanılmıştır. YSA modelinin eğitiminde 120 adet eğitim verisi ve testinde de 30 adet test verisi kullanılmıştır. Çalışma sonucunda tahmin değerleri ile gerçek değerler arasında düşük hata oranları tespit edilmiştir. Yapay sinir ağıımız ne kadar eğitilirse bu değerler arası hata oranları o kadar azalmış ve gerçek değere bir o kadar yaklaşmıştır.

Anahtar Kelimeler: Iris çiçeği, Matlab, Yapay sinir ağı, Veri kümeleme

ABSTRACT

ANN METHODS OF IRIS FLOWER TYPE
WITH THE FORECAST

Classification using known data is widely used in many fields. However, some errors may occur when classifying based on observations. Artificial learning methods are used to minimize these errors. The aim of this study is to try to determine the type of iris flower by using artificial neural network (ANN) methods. In the study, iris flower data, which are among the most popular on the UCI website, were used Decisively. ANN model with three layers, feed-forward, four inputs and one output is used. In the training of the ANN model, 120 pieces of training data and 30 pieces of test data were used in the test. As a result of the study, low error rates were found between the estimated values and the actual values. Dec. The more our artificial neural network is trained, the lower the error rates between these values Decayed and the closer it got to the real value.

Keywords: Iris flower, Matlab, Artificial neural network, Data clustering

1. GİRİŞ

Birçok alanda ve akademik çalışmada yaygın olarak veriler kullanılarak sınıflandırma ve kümeleme yapılmaktadır. Veriler üzerinde yapılan bu sınıflandırma ve kümeleme işlemleri, çıplak göz ile yapıldığı zaman yapılan işlemlerde bazı hatalar meydana gelebilmektedir. Son zamanlarda yapılan çalışmalarda, bu hataları azaltmak için çeşitli yapay zeka çalışmaları kullanılmaktadır. Yapay zeka, bilgisayarların örnek veri ve geçmiş deneyimlerini kullanarak öğrenmesi ve tahminlemesine yönelik kurulmuş bir alandır. Mohan ve Subashini'nin yaptığı çalışmada, manyetik rezonans görüntülerinden yararlanarak beyin tümör seviyesinin tespit edilmesine kadar birçok yapay öğrenme yönteminin başarılı bir şekilde sınıflandırılabilirdiğini belirtmiştir. Bu çalışmada, Yapay Sinir Ağı (YSA) yöntemlerini kullanarak, veriler eğitim ve test olarak 2 gruba ayrılmıştır. İris çiçeğinin çanak yaprak uzunluğu (sepal length), çanak yaprak genişliği (sepal width), taç yaprak uzunluğu (petal length), taç yaprak genişliği (petal width) gibi özelliklerine göre ise 4 sınıfa ayrılmıştır. Verilerin işlenmesi ve iris çiçeğinin ait olduğu türün en yakın oranda belirlenmesi hedeflenmektedir. Bu özellikleri kullanılan eğitim ağına giriş parametreleri olarak sunulmuş ve bu değerlere göre sistemin iris çiçeğinin türünü tahmin etmesi sağlanmıştır. YSA modelinin eğitiminde 120 adet eğitim verisi ve testinde de 30 adet test verisi kullanılmıştır.

2.MATERYAL VE METOT

Bu çalışmada veriler UCI veri setinden alınmıştır. Yapay sinir ağı eğitimi, ağın veri tahmini ve sonuç olarak oluşan çıktıyı elde etmek için MATLAB programlama dili kullanılmıştır. MATLAB'a veriler Excel üzerinden çekilmiştir.

2.1. Yapay Sinir Ağları

İnsan beyni çalışma yapısından ötürü bir çok alanda bilimde dikkat çekici çalışmalara konu olmuştur. İnsan beyni içerisinde bulunan milyarlarca nöron ile adeta bir makineye benzetilmektedir. Yapay sinir ağları insan beyninde bulunan nöron hücreleri ve bu hücreler arasında kurulan sinaptik bağların modellenmesi ile ortaya çıkmıştır. Nöronlar arası iletişim ele alınarak oluşturulan modelde düğümler arası çeşitli ağırlıkta bağlantılar mevcuttur. İnsan beyninin yaşayarak, deneyimleyerek ve tecrübe kazanarak öğrendiği bilgi ve olayları yapay sinir ağına da aynı mantıktan yola çıkarak ağına bir çok kez girdi/çıkı verilerinin ağı defalarca tanıtılmasıyla eğitimi sonucu bu tecrübe ve deneyim kazandırılmaya çalışarak doğruluk oranlarına yaklaşılmaya çalışılır. Yapay sinir ağları girdi katmanı, gizli katman ve çıktı katmanı olarak 3 katmandan oluşmaktadır. Girdi katmanına dışardan gelen bilgiler ağırlıklar ile çarpılarak bunun sonucunda aktivasyon fonksiyonu kullanılarak aktivasyon değeri oluşur. Aktivasyon fonksiyonları bir nevi Normalizasyon işlemi gerçekleştirir. Aktivasyon değerinin ortaya çıkması gizli katmanda gerçekleşmektedir. Gizli katmandaki işlemler sonucunda oluşan bu değer, çıktı katmanına gönderilerek dışarıya veya tekrar kendine girdi olarak dönebilmektedir.

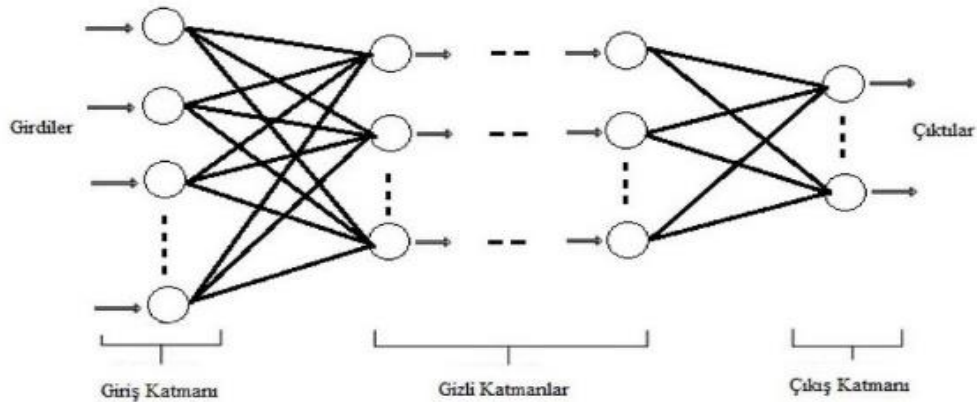


Figure-1 (Yapay Sinir Ağı Yapısı)

2.2. İris Veri Seti

Fisher, iris çiçeğini çeşitli ölçümler neticesinde üç türe ayırarak sınıflandırmıştır. Çiçek türünün sınıflandırılmasında çiçeklere ait 4 belirgin özellikten faydalanılmıştır. Bu özellikler; taç yaprak genişliği, çanak yaprak genişliği, taç yaprak uzunluğu ile çanak yaprak uzunluğudur.

Ölçümlerin sonunda iris çiçeği için elde edilen bu dört belirgin özellik kullanılarak iris çiçeği üç türe göre sınıflandırılmaktadır. Kullanılan veri setinde yer alan değerler şu şekildedir:

- A: Çanak Yaprak Uzunluğu (sepal length)
- B: Çanak Yaprak Genişliği (sepal width)
- C: Taç Yaprak Uzunluğu (petal length)
- D: Taç Yaprak Genişliği (petal width)
- E: İris çiçeğinin türü (setosa, versicolor, virginica)

Çalışmada kullanılan veri seti incelendiğinde A,B,C,D bağımsız değişken olan giriş değerlerini, bağımlı E değişkeni ise çıkış değerini temsil etmektedir. Veri seti içerisinde, önceden ölçümü yapılarak tanımlanmış olan 150 adet iris çiçeği örneği için ölçüm değerleri ve tür değerleri bulunmaktadır. Elde bulunan veriler doğrultusunda sistem eğitilerek yeni gelen verilere bakılarak çiçek türünün tahmin edilmesine yönelik sistem tasarlanmaktadır. Sistemin eğitilmesi için veriler, Eğitim Veri Seti ve test edilmesi için ise Test Veri Seti olarak iki gruba ayrılmıştır. İlk olarak eğitim verileri ile ağımızı eğiterek, ardından eldeki test verileri ile ağımız üzerinde doğru sonuç verip vermeyeceği test edilmektedir.

2.3. Normalizasyon İşlemi

Veri seti düzenlenirken tür kısmı sözel ifadelerden oluştuğu için bu ifadeler sayısal değere dönüştürülmektedir. Çünkü yapay sinir ağlarında değerlerin sisteme sayısal olarak tanıtılması gereklidir. Bundan dolayı 1-3 arası sayısal değerlerle bu ifadeler değiştirilmiştir. Sayısal değerler ve karşılıkları aşağıdaki gibidir:

- 1: setosa
- 2: versicolor
- 3: virginica

Veri setleriyle çalışmalar yapılırken sonucun daha iyi olması adına çalışmada kullanılan bütün verilerin belirlenen bir değer aralığında normalize edilmesi daha iyi sonuçlara ulaşmamızı sağlamaktadır.

2.4. Eğitim ve Test Veri Setlerinin Oluşturulması

Sistem tasarlanırken, ilk olarak sistemin eğitilmesi daha sonra eğitilen bu sistemin test edilmesi gerekmektedir. Çalışmada toplam 150 adet veriden 120 tanesi sistemi eğitmek için kullanılmıştır. Geri kalan 30 adet veri ise sistemin eğitimini test etmek amacıyla Test veri seti olarak kullanılmıştır

3. PROJE VE EKRAN GÖRÜNTÜLERİ

İlk olarak verileri Excel tablomuza ekleme işlemi gerçekleştirildi. 150 adet veri tabloya aktarıldı. Verilerin türlerini bulmamıza yardım edecek veri özellikleri de tabloya eklendi. Bu aşamada veri seti düzenlenirken tür kısmı sözel(string) ifadelerden oluştuğu için bu ifadeler sayısal değere dönüştürülmektedir yani Normalizasyon işlemi gerçekleştirilecektir. Çünkü yapay sinir ağlarında değerler sisteme sayısal olarak tanıtılmak zorundadır aynı zamanda matlab sayısal değerlerle çalışır ve proje matlab üzerinden gerçekleştirildi. Bu yüzden 1-3 arası sayısal değerlerle bu sözel ifadeler değiştirilmiştir.

1	sepal length	sepal width	petal length	petal width	Class	
2	5.4	3.7	1.5	0.2	Iris-setosa	1
3	4.8	3.4	1.6	0.2	Iris-setosa	1

4	5.0	2.0	3.5	1.0	Iris-versicolor	2
5	5.9	3.0	4.2	1.5	Iris-versicolor	2
6	6.5	3.2	5.1	2.0	Iris-virginica	3
7	6.4	2.7	5.3	1.9	Iris-virginica	3
8

Tablo-1 (Veriler ve veri özellikleri)

Ardından verileri test etmek doğruluğunu kontrol etmek amacı ile 30 adet test verisi seçilmiştir.

sepal length	sepal width	petal length	petal width	Class	
5.1	3.5	1.4	0.2	Iris-setosa	1
4.9	3.0	1.4	0.2	Iris-setosa	1
7.0	3.2	4.7	1.4	Iris-versicolor	2
6.4	3.2	4.5	1.5	Iris-versicolor	2
6.3	3.3	6.0	2.5	Iris-virginica	3
5.8	2.7	5.1	1.9	Iris-virginica	3
...

Tablo-2 (Test verileri)

Veriler matlab da Workspace ekranına input-output olarak girilmelidir. Matlab da “import data” kısmını kullanarak Excel den verileri alma işlemi gerçekleştirilmiştir. Daha sonra Excel den input, output ve test verileri seçme işlemi gerçekleştirildi. Her seçme işleminden sonra “import selection” kısmına tıklanarak matlab’a import edildi. İmport işlemlerinden sonra input, output ve test verileri matlab da Workspace ekranında gözükmetedir.

Sonrasında ise nntool Toolbox kullanıldı. Nntool Toolbox’ında veriler satır-sütun olarak çalışır bundan dolayı öncelikle bir transpoze işlemi gerçekleştirilmiştir. Transpoze işleminin kodu aşağıdaki gibidir:

```
input=input';
output=output';
test=test';
nntool
```


Crete diyerek yeni bir ağı oluşturduktan sonra Train kısmından ağı eğitme işlemi gerçekleştirilmiştir. Ağı eğitme işleminde hata oranını düşürmek amacı ile tekrar sayısı olarak 1000 seçilerek ağı eğitime sokulmuştur.

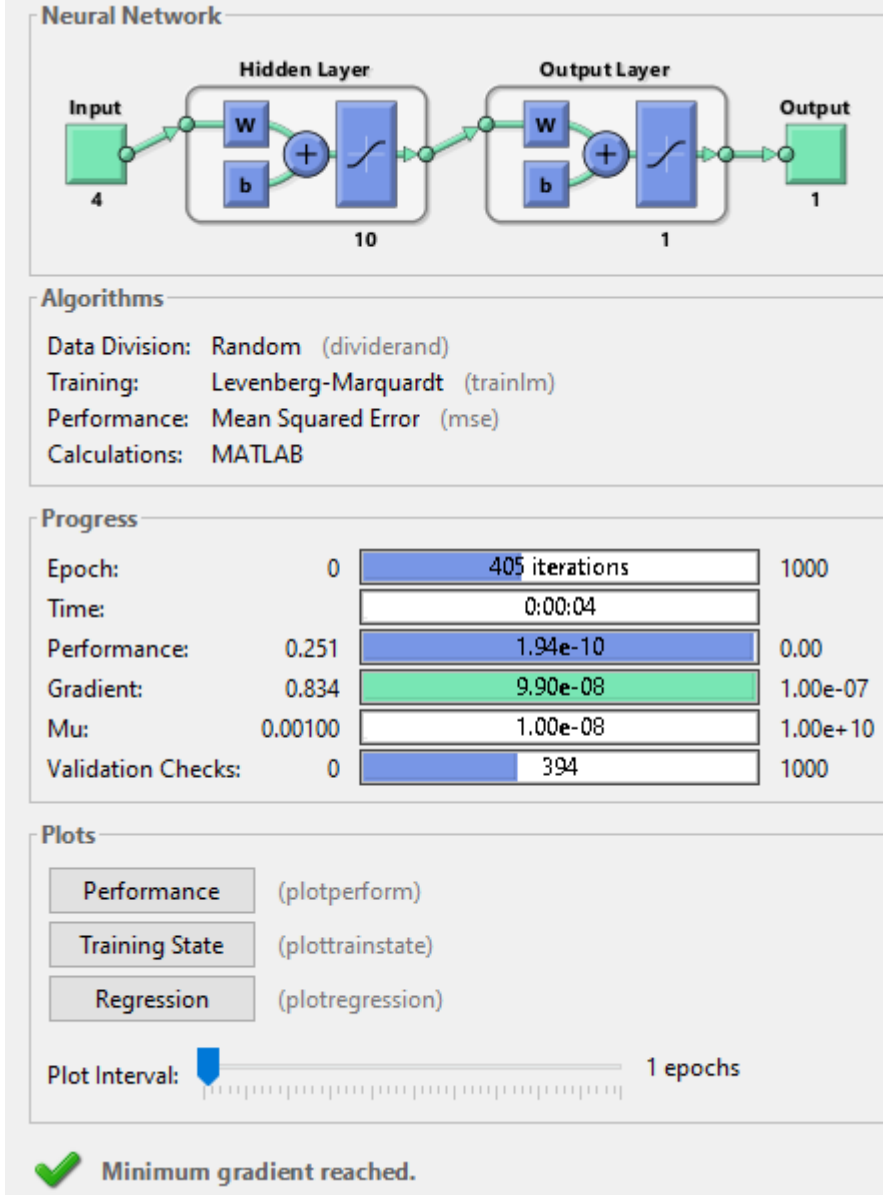


Figure-4 (Ağı eğitme işlemi)

Ardından ağı eğitimi sonrası oluşan çıktının yüzde kaç oranda asıl verilere benzediği gösterilmiştir.

4. SONUÇ

Tüm bu işlemlerden sonra veriler export edilerek eğitim verilerinin asıl verilerle karşılaştırılması yapılarak ne kadar oranda doğru sonuç verilmiş ve ne kadar sapmalar yaşanmış test edilmiştir. Bu çalışmada sonuç olarak, iris çiçeğinin çanak yaprak uzunluğu, çanak yaprak genişliği, taç yaprak uzunluğu ve taç yaprak genişliği özelliklerine göre çiçeğin türünün tahmin edilmesi için Matlab ve YSA yöntemleri

kullanılmıştır. Yapılan tahmin işlemlerinin doğruluğu geçmişte yapılmış proje sonuçları ile karşılaştırılmıştır. YSA yönteminde tahmin işlemi sonucunda bulunan sonuç değeri %0.97708 çıkmıştır. Bulunan değerin 1'e yakın bir sonuç verdiği için kabul edilebilir bir seviyededir. Elde edilen tahmin sonuçları dikkate alındığında, YSA yöntemleri ile yapılan projede iyi sonuçlar elde edildiği görülmektedir. Elde edilen değerler, gerçek "tür" bilgisi ile yapay sinir ağı modelinin eğitimi sonucu elde edilen tahmin sonuçları arasındaki sapmaların çok küçük olduğu gözlemlenmiştir. Sonuç olarak yapılan çalışmada iris çiçeği türlerinin YSA eğitimi ile az bir hata oranıyla birlikte sonuca yakın değerlerle tahmin edilebileceğini göstermek amaçlanmıştır.

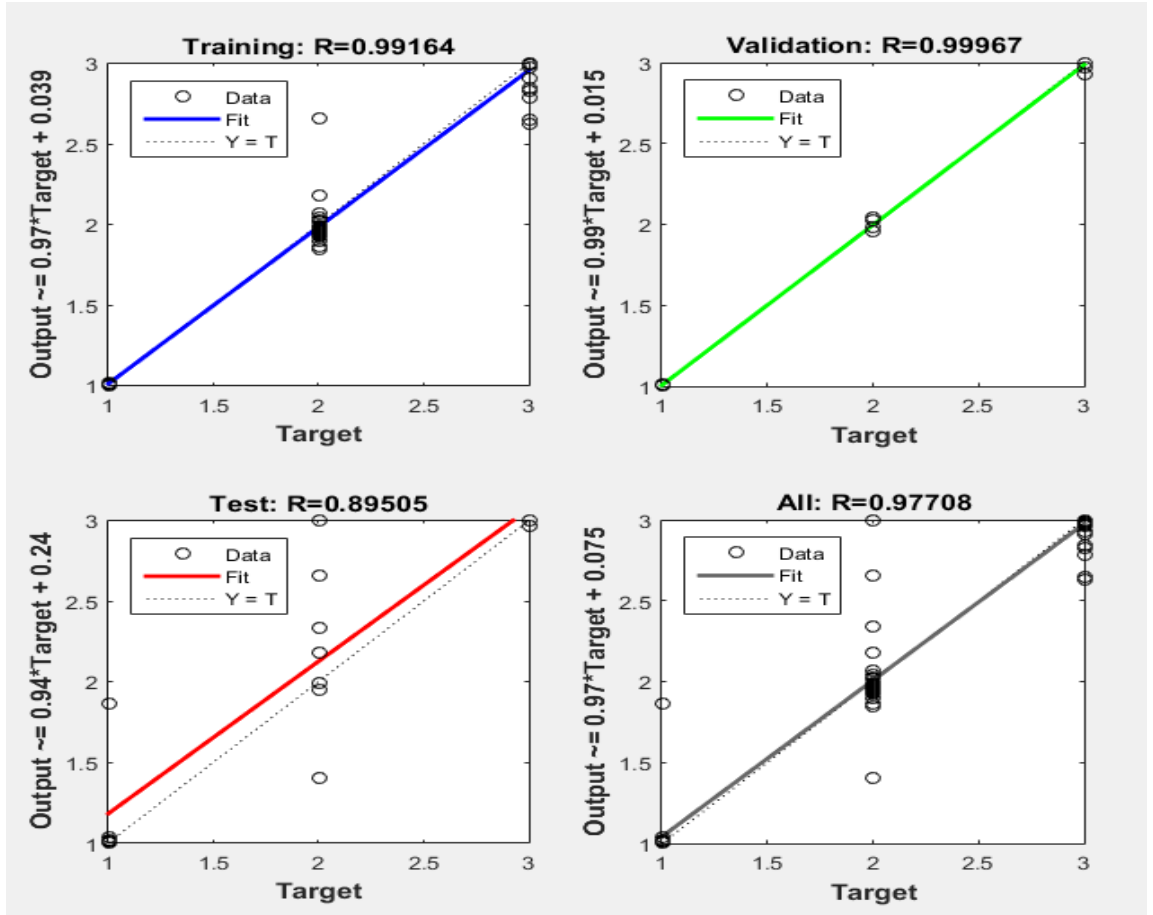


Figure-5 (Yüzde kaç oranla benzer olduğu)

5. KAYNAKLAR

- [1]İris Çiçeği Türünün YSA Yöntemleri ve ANFIS ile Tahmini
Yıl 2020, Cilt 1, Sayı 1, 5 - 11, 01.01.2020 [Serel ÖZMEN-AKYOL](#) [Eyyüp GÜLBANDILAR](#)
- [2]B. S. Khehra and A. P. S. Pharwaha, "Classification of Clustered Microcalcifications using MLFFBP- ANN and SVM," Egypt. Informatics J., vol. 17, no. 1, pp. 11–20, 2016.
- [3]"İris Veri Seti," <https://archive.ics.uci.edu/ml/datasets.php>

[4]Çuhadar, M. and Kayacan, C., “Yapay Sinir Ağları Kullanılarak Konaklama İşletmelerinde Doluluk Oranı Tahmini : Türkiye ’ deki Konaklama İşletmeleri Üzerine Bir Deneme,” Anatolia Tur. Araştırmaları Derg., vol. 16, no. 1, pp. 24–30, 2005.

[5] M. Caner and E. Akarslan, “Mermer Kesme İşleminde Spesifik Enerji Faktörünün ANFIS ve YSA Yöntemleri ile Tahmini Estimation of Specific Energy Factor in Marble Cutting Process Using,” Pamukkale Üniversitesi Mühendislik Bilim. Derg., vol. 15, no. 2, pp. 221–226, 2009.

[6] “İris Çiçeği,” <https://www.veribilimiokulu.com/nedir-bu-iris-cicek-muhabbeti/>

[7] Yapay Sinir Ağı Kullanarak Deprem Eğiliminin Kestirimi. Umut FIRAT