

Sınıflandırma Problemlerinin Karşılaştırılmasında ANFIS ve Basamak Korelasyon Sinir Ağının Kullanımı

Nilgün ŞENGÖZ*¹, Gültekin ÖZDEMİR²

¹Mehmet Akif Ersoy Üniversitesi, Stratejik İşbirliği Proje Danışmanlık Eğitim Uygulama ve Araştırma Merkezi, 15030, Burdur

²Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, 32260, Isparta

(Alınış / Received: 28.02.2016, Kabul / Accepted: 30.09.2016, Online Yayınlanma / Published Online: 04.11.2016)

Anahtar Kelimeler

Yapay zeka,
Yapay sinir ağları,
ANFIS,
Basamak korelasyon sinir ağları,
BKSA

Özet: Yapılan bu çalışmada, sınıflandırma problemleri üzerinde durulmuştur. Bu çalışma farklı yapıda olan 3 adet veri kümesinin (Tohum, Arazi Uydu Görüntüleme ve Kırmızı Şarap Kalite Değerlendirme) incelenmesi ve aynı algoritmaların bu verilere uygulanması itibariyle hem başarı performansı olarak hem de zaman bakımından aralarındaki ilişkiyi gözler önüne sermektedir. Tohum verisi için yapılan karşılaştırmada Uyarlamalı ağ tabanlı bulanık çıkarım sistemi (Adaptive-Network Based Fuzzy Inference Systems-ANFIS) ile Basamak Korelasyon Sinir Ağı BKSA (Cascade Correlation Neural Network-CCNN) test performans yüzdeleri %86.41 ve %88.06 olup, işlem zamanı yönünden bakıldığında ise BKSA 0.59 saniye gibi kısa bir süre içerisinde sınıflandırma yapmıştır. Arazi Uydu Görüntüleme veri kümesinde ise test performans yüzdeleri ANFIS de %100, BKSA da ise %99.92 olmaktadır. Verinin algoritma içerisinde yaptığı işlem süreleri ise ANFIS 800 saniye, BKSA 72 saniyede tüm sınıflandırma işlemini tamamlamıştır. Son veri kümemiz ise Kırmızı Şarap Kalite değerlendirmedir ve test performans olarak bakıldığında ANFIS %99.975, BKSA ise %99.862 gibi birbirleri arasında çokta anlamlı bir fark oluşturmadığı açıkça görülmektedir. İşlem zamanı bakımından ANFIS 85.271 saniye ile BKSA metoduna göre çok hantal kalmaktadır.

Comparison of ANFIS and Cascade Correlation Neural Network Methods for Classification Problems

Keywords

Artificial intelligence,
Neural network,
ANFIS,
Cascade correlation
neural network,
CCNN

Abstract: In this study, it focused on classification problems. This thesis different structure with 3 data set (Seed, Terrain Satellite Imaging and Red Wine Quality) examination and the same algorithms applied to these data as the relationship as well as the performance achievements and time in terms between the eyes reveals. Seed data is for the comparison with the ANFIS and CCNN test performance of 86.41% and 88.06% percent, in the classification process has always CCNN less time when viewed from the side. In the dataset Terrain Satellite Imaging performance ANFIS percentage of 100%, in the CCNN is 99.92%. The duration of the transactions made in the algorithm's data ANFIS 800 seconds, it completed the classification process CCNN all in 72 seconds. Our last data set is red wine quality evaluation and performance when viewed as ANFIS% 99 975, the CCNN is clearly not constitute a very significant difference between each other like 99 862%. According to the terms of processing time with the BKS method ANFIS 85, 271 seconds is far too cumbersome.

1. Giriş

Örüntü tanıma, sayıca fazla kategorilere veya sınıflara ayırma işlemleri amacını güden bir disiplin olarak yer almaktadır. Uygulamaya bağlı olarak araştırmacı tarafından verilen nesnelere; resim, sinyal dalgacıkları

veya sınıflandırılacak herhangi bir veri kümesi olabilir. Endüstride gerçekleşen otomasyon ile yeni bilgilerin idare edilmesi ve verilerden anlamlı bilgi çıkarılma durumu daha önemli hale gelmesi mühendislerin bu alana doğru araştırma yapmaya zorlamış ve karar vermede çoğu makine zekâsı

sistemleri örüntü tanıma da ayrılmaz bir parça haline dönüşmüştür.

Uygulayıcı metotlar üzerinde hangisinin en optimal, hızlı ve hata payı en düşük olanı bulmasıyla sisteme girilen veri kümeleri üzerinden en uygun algoritmanın seçilmesidir.. Sadece tek algoritma bazlı incelemeler gerçekçi sonuçlar elde etme olanağını büyük ölçüde kısıtlayacağı için, farklı tarzlarda yapılan uygulamalar da hem sonuç bazlı hem de uygulama derinliği açısından karşılaştırma olanağı bulunmaktadır. Farklı algoritmaların karşılaştırılacağı veri kümeleri birbirinden farklı tarzda olması önem arz etmektedir. Ayrıca uygulayıcının çıktısı sonuçlarından hem eğitim hem de test performansları yüzdeleri ve sistem üzerinde uygulama üzerinde geçen süre de bir o kadar önem arz etmektedir.

Çalışmada kullanılacak olan algoritmalarından bir tanesi Basamak Korelasyon Sinir Ağı (BKSA)'dır. İlk defa Fahlman ve Lebiere (1990) tarafından geliştirilen BKSA, yapay sinir ağları içerisinde öğreticili/denetimli öğrenme algoritmasıdır [1]. Augusteijn ve Clemens (1994) görüntü sınıflandırma ve segmentlere ayırmada [2], McKenna vd (1991) Cervical Hücrelerin sınıflara ayrılmasında [3], Augusteijn ve Skujca (1993) doku tabanlı unsur tanımda bu algoritmayı kullanmışlardır [4].

Daha sonraki yıllarda BKSA farklı metotlar ile karşılaştırılmaya başlanmıştır. Burke vd (1997) tıbbi sonuç tahmini için lojistik regresyon, geri yayımlı sinir ağı, eşlenik gradyan iniş sinir ağı, regresyon ağacı gibi metotlar ile karşılaştırılmıştır [5]. Ayrıca, Prampero ve Carvalho (1999) araç silüetleri sınıflandırılmasında Çok Katmanlı Perseptron (MLP) ağı ile Radyal Tabanlı Fonksiyonu (RTF) kullanarak karşılaştırma gerçekleştirmişlerdir [6].

Çalışmada karşılaştırma yapılacak olan bir ikinci algoritma ise Bulanık Sinir Ağ Yapısı yani ANFIS yöntemidir. Bu modeli, Jang (1997) geliştirmiş ve doğrusal olmayan fonksiyonların modellenmesinde, kaotik zaman serilerinin tahmininde kullanmıştır [7]. Şenol ve Yıldırım (2009) tiroit hastalıklarının teşhisi için, ANFIS, Fuzzy-MLP, MLP, Fuzzy-RBF, RBF, Fuzzy-CSFNN, CSFNN gibi uzman sistemlerin başarılarını karşılaştırmıştır [8]. Sharma ve Mukharjee (2013) beyin tümör bulma problemi için ANFIS metodu ile birlikte bulanık C ortalama (FCM) ve K-en yakın komşu algoritmaları kullanmıştır [9]. Ziasabounchi ve Askerzade (2014) kalp krizi tahmini için ANFIS metodunu kullanmışlardır [10]. Horlama ile ilgili seslerin büyük lyapunov üstel ve entropi kullanarak sınıflandırma problemi için Ankişhan ve Yılmaz (2013) ANFIS metodu ile Destek Vektör Makinasını birlikte kullanmış ve karşılaştırma yapmışlardır [11].

Tüm bu tahmin ve sınıflandırma karşılaştırmaları içerisinde basamak korelasyon sinir ağı ile bulanık sinir ağı algoritmalarının aynı makale içerisinde

görülmeyeceği tespit edilmiş olup, literatürdeki boşluğu doldurmak üzere çalışma konusunun bu araştırma olabileceği düşünüldü. Yapılan kapsamlı literatür çalışmasından da anlaşıldığı üzere sınıflandırma, tahmin, görüntü işleme vb. problemlerde genellikle birbirine yakın metotlar karşılaştırılmış ve çalışmalar bu şekilde ilerlemiştir. Çalışmanın amacı ortaya konmak istenen bir sınıflandırma problemi için iki farklı metodun (Basamak Korelasyon Sinir Ağı ile Bulanık Sinir Ağı) karşılaştırılması ve hangi metodun daha iyi sınıflandırma yapacağına dair hesapların ortaya konmasıdır. Bu iki farklı türde metodun aynı veri kümesi üzerinde kullanılması literatür açısından yeni bir bakış açısı kazandıracaktır..

2. Materyal ve Metot

Çalışmanın amacına yönelik olarak, örüntü sınıflandırma için UCI Machine Learning (<https://archive.ics.uci.edu/ml/datasets.html>) internet sitesinden güncel olan 3 adet birbirinden farklı veri kümeleri (Tohum, Arazi Uydu Görüntüleme ve Kırmızı Şarap Kalite Değerlendirme) üzerinde Basamak korelasyon sinir ağı ile Bulanık-sinir ağı metodları uygulanmıştır.

2.1. Basamak korelasyon sinir ağı (BKSA)

Basamak Korelasyon öğrenme algoritması geri yayılım öğrenme algoritmasının getirdiği problemler ve kısıtlamalar yüzünden geliştirilmiştir [12].

Bir ağındaki sabit topoloji de ağırlıkları ayarlamak yerine, Basamak Korelasyon ilk önce minimum ağı ile başlar, sonra otomatik olarak eğitir ve yeni gizli katmanları tek tek ekleyerek çok katmanlı bir yapı oluşturur. Eğer yeni gizli katman ağı eklenirse, giriş tarafındaki ağırlıklar dondurulur. Eklenmiş olan bu katman kalıcı bir hal alarak ağıda özellik sezinleyici (feature-detector) olur ki bu durum daha karmaşık özellik sezinleyiciler için yeni çıkışlar üretme veya yeni yaratımlar için, olanak sağlar. Basamak Korelasyon yapısı var olan diğer algoritmalara nazaran bazı avantajlara sahiptir: çabuk öğrenir, ağı boyutuna ve topolojisine kendi karar verir, eğitim kümesi değişse dahi kurulmuş olan yapı kendini korur ve ağındaki bağlantıları aracılığıyla hata sinyallerinin hiçbirini geri yayılım gerektirmez.

Fahlman ve Lebiere (1990) tarafından bulunan bu algoritmanın 11 temel adımları aşağıda sunulmuştur:

Adım 1: Problem tarafından verilen tüm giriş ve çıkışlar birbiriyle bağlantılı olacak şekilde başlanır.

Adım 2: Tüm bağlantılar karesel hata değeri (E_s) anlamlı ölçüde düşmeyecek şekilde bilinen bir öğrenme algoritması ile eğitilir.

$$E_s = \frac{1}{2} \sum_{p=1}^{N_p} \|y_p - t_p\|^2 = \frac{1}{2} \sum_{p=1}^{N_p} \sum_{i=1}^m [y_{i,p} - t_{i,p}]^2 \quad (1)$$

(Fahlman ve Lebiere, 1990)

m çıkış sayısını, N_p eğitim seti büyüklüğünü, $y_{i,p}$ sinir ağından i'ninci çıkışı, $t_{i,p}$ ise ilgili hedefi belirtir.

Adım 3: Tüm ağıın dış girişlerinden ve eğer varsa var olan önceki gizli katmanlarından aday bir katman oluşturulur. Bu aday katman henüz aktif olan ağıın çıkışına bağılı değıildir.

Adım 4: S_c olarak gösterilen (Denklem 2 korelasyonu maksimize edecek şekilde katman (ağırlıkları) eğitilir. Öğrenme, sıradan bir algoritma ile çok fazla zaman alır; korelasyon daha fazla ilerleyemediğı zaman eğitim durur ve bunun denklemi ise aşığıda belirtilmiştir;

$$S_c = \sum_{i=1}^m \left| \sum_{p=1}^{P_{max}} (z_{o,p} - \bar{z}_o) (E_{i,p} - \bar{E}_i) \right| \quad (2)$$

(Fahlman ve Lebiere, 1990)

$z_{o,p}$ aday gizli katmanın çıkışını, $E_{i,p}$ ise adım 2 de hesaplanan ($E_{i,p} = y_{i,p} - t_{i,p}$) çıkışların artık hata değıerini belirtmektedir. Bir niceliğın üstündeki çizgi eğitim seti üzerindeki ortalamayı gösterir.

Adım 5: Aday katmanı çıkışlara bağılanır ve giriş değıerleri dondurulur. Bu sayede aday katman fazladan giriş katman olarak hareket eder.

Adım 6: Adım 2'de de belirtildiğı üzere, giriş-çıkış bağılantıları tekrardan karesel hata değıer (E_s)'ini minimize edecek şekilde eğitilir.

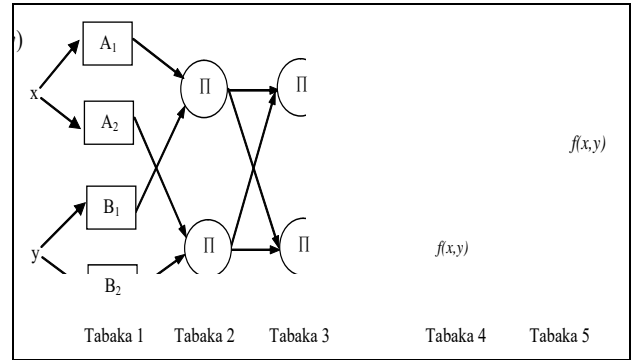
Adım 7 ve Adım 10: Adım 3'ten Adım 6'ya kadar olan kısmı bir seferde gizli katman ekleyerek tekrar edilir.

Adım 11: E, net hatası verilen bir değıerin, ϵ , altına düştüğü zaman eğitim durdurulur. Tek bir aday katman yerine, aday katmanların olduğı bir havuzdan her biri farklı başlangıç ağırlık kümesi kullanmak daha iyidir. Hepsini aynı giriş sinyallerini alır ve her bir eğitim modeli için aynı artık hata görülür. Çünkü onlar birbirleriyle etkileşime geçmez veya aktif ağı etkilemez, aynı anda eğitilmiş olabilirler. Sadece adaylardan hangisinin korelasyon skoru iyiyse o yüklenir. Aday havuz sistemini kullanımı eğitimi kötü yönde etkileyecek ve onu durdurabilecek bir katmanın yüklenmesi riskini ortadan kaldırıyor. Fahlman ve Lebiere (1990), her bir havuzun içerisinde dört veya sekiz aday katmanın olma durumunun yeterli olacağını belirtmiştir [13].

2.2. Sinirsel bulanık mantık çıkarım sistemi (ANFIS)

YSA ve Bulanık Mantık yöntemlerinin birleşiminden oluşan Sinirsel Bulanık Mantık Çıkarım Sistemi (ANFIS) bir modelleme yöntemi önerilmekte ve

kullanılmaktadır. Bulanık mantık ile modellemede en önemli adım giriş/çıkış değıişkenlerinin üyelik derecelerinin belirlenmesi olarak gösterilmektedir. ANFIS, YSA'nın öğrenme yeteneğini kullanarak, eğitim veri kümesinden bulanık set ve kuralları belirlemekte ve kural tabanlı bir modelleme gerçekleştirmektedir. ANFIS, üç katmanlı ileri beslemeli bir YSA şeklinde değıerlendirilebilir. Bu ağı yapısında, birinci katman, giriş değıişkenlerinin uygulandığı ve YSA yardımıyla üyelik fonksiyonlarının belirlendiğı katman, ikinci katman bulanık kuralların oluşturulduğı ve çıkarım işleminin gerçekleştirildiğı katman ve en son katman ise çıkarım sistem sonucunun elde edildiğı katmandır [7].



Şekil 1. İki girişli bir Sugeno Bulanık Çıkarım Sistem Yapısı (Jang, 1997)

Katman 1: Bu katmanda her bir düğüm, giriş değıişkenlerinin üyelik fonksiyonları düğüm fonksiyonlarını oluşturur. Her bir düğüm çıkışı Denklem 3 ve 4 gibi hesaplanır;

$$O_{1,i} = \mu_{A_i}(x) \quad i = 1,2 \quad (3)$$

$$O_{1,i} = \mu_{B_{i-2}}(y) \quad i = 3,4 \quad (4)$$

Denklem (5)'de verilen Genelleştirilmiş çan tipi üyelik fonksiyonu kullanılarak, düğüm çıkışı hesaplanmıştır. Burada, bulanık çıkarım sisteminin lineer olmayan giriş parametreleridir.

$$OP_i^1 = \mu_{A_i}(x) = \frac{1}{1 + \left(\frac{x - c_i}{a_i} \right)^{2b}} \quad (5)$$

Katman 2: Bu katmandaki her bir nöron AND/OR operatörleri kullanılarak giriş sinyalleri ile çarpılır. Düğüm çıkışı için ateşleme kuvveti denklem (6) yardımıyla hesaplanır;

$$O_{2,k} = w_k = \mu_{A_i}(x) \mu_{B_j}(y), \quad i=1,2 \quad j=1,2 \quad k=1,..4 \quad (6)$$

Katman 3: Bu katmanda ateşleme kuvveti normalize edilir. i^{th} düğüm için normalize edilmiş ateşleme kuvveti aşığıdaki gibi hesaplanır.

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, \dots, 4 \quad (7)$$

(Jang, 1997)

Katman 4: Bu katmanda, her bir düğümün model çıktısı üzerindeki katkısı hesaplanır.

$$O_{4,i} = \overline{w}_i \cdot f_i = \overline{w}_i \cdot (p_i x + q_i y + r_i) \quad i = 1, \dots, 4 \quad (8)$$

(Jang, 1997)

Burada, \overline{w} katman 3'ün çıkışı; $\{p_i, q_i, r_i\}$ bulanık çıkarım sisteminin çıkış parametreleridir.

Katman 5: Bu katman çıkış katmanı olarak adlandırılır ve ANFIS'in genel çıkışı hesaplanarak durulaştırma işlemi ile bulanık değerden kesin değere dönüştürülür.

$$f(x, y) = \frac{w_1(x, y)f_1(x, y) + w_2(x, y)f_2(x, y)}{w_1(x, y) + w_2(x, y)} = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2} \quad (9)$$

$$Q_{5,i} = f(x, y) = \sum_i \overline{w}_i \cdot f_i = \overline{w}_1 f_1 + \overline{w}_2 f_2 = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (10)$$

(Jang, 1997)

3. Bulgular

Bu çalışmada örüntü sınıflandırma için UCI Machine Learning internet sitesinden güncel olan 3 adet birbirinden farklı veri kümeleri üzerinde BKSA ile ANFIS metodları uygulanmıştır. Bu çalışma yapılırken gerekli görülen Bilgisayar Programlama dilleri kullanılmıştır. Oluşturulan tüm tablolar yapılan işlemler sonuçlarını göstermektedir.

İlk incelenen veri kümesi Tohum veri kümesi olup, geometrik şekillerine göre ölçülmüş 7 parametre mevcuttur. Bunlar; kernel buğdayının alanı, çevresi, uzunluğu, genişliği, oyuğu, yoğunluğu ve asimetric katsayısına göre 207 tane örnek mevcuttur. Buradaki önemli ayrıntı ise, ilk 69 veri Kama, sonraki 69 veri Rosa ve son 69 veri ise Canadian olan 3 sınıf mevcuttur.

Tablo 1. Tohum adlı veri kümesinin sonuç değerleri

Model	Eğitim Performansları (%)	Test Performansları (%)	Zaman (saniye)
ANFIS	98.076923	86.407767	1.145
BKSA	98.365385	88.058252	0.598

Arazi uydu görüntü veri kümesi, bir uydu görüntüsünde 3x3'lük komşu piksellerin oluşturduğu çoklu-spektral değerlerini içermekte olup, her bir komşusunda merkezi piksel ile sınıflandırma oluşturmaktadır. Amacı ise verilen çoklu-spektral değerleri üzerinde bu sınıflandırmayı tahmin

etmektir. Veri kümesi içerisinde piksellerin sınıfı sayı olarak verilmiştir.

Bu veri kümesinde 6 adet sınıf vardır, bunlar; kırmızı toprak, pamuk mahsulü, gri toprak, nemli gri toprak, bitki anızı ile toprak ve son olarak da çok nemli gri topraktır. Bunların her birine sayı verilmiş olup sırasıyla 1'den 6'ya kadar gitmektedir. Toplamda 6435 veri olup, her bir sınıfın 37 adet özneliği bulunmaktadır.

Tablo 2. Arazi Uydu Görüntüsü adlı veri kümesinin sonuç değerleri

Model	Eğitim Performansları (%)	Test Performansları (%)	Zaman (saniye)
ANFIS	100	100	800.951
BKSA	99.972032	99.928505	72.166

En son olarak incelenen veri kümesi ise Kırmızı Şarap Kalite Değerlendirmedir. UCI Machine Learning sitesinden en iyi 10 verisinden bir tanesi olup, Mayıs/2004 ile Şubat/2007 tarihleri arasında 1600 adet veri toplanmıştır. Girdi değişkenleri (fizikokimyasal deneylerine dayanmaktadır) olarak sırasıyla; sabit asitlik, uçucu asitlik, sitrik asit, artık şeker, klorürler, kükürt dioksitsiz, toplam kükürt dioksit, yoğunluk, pH, sülfatlar ve alkoldür. Çıktı değişkeni (duyusal verilere dayalı) ise kırmızı şarabın kalitesini gösteren 0 ile 10 arasında puanlamadır.

Tablo 3. Kırmızı Şarap Kalite Değerlendirme veri kümesinin sonuç değerleri

Model	Eğitim Performansları (%)	Test Performansları (%)	Zaman (saniye)
ANFIS	100	99.975	85.271
BKSA	99.9625	99.862	4.0086

4. Tartışma ve Sonuç

Yapay zeka çalışmaları gün geçtikçe günümüz dünyasında daha fazla önem kazanmakta olup neredeyse tüm sistemlere uygulanabilirliği üzerinde çalışmalar yapılmaktadır. Büyük veri kümeleri içerisinde istenilen bilgiye ulaşmak önem arz etmektedir. Bu yüzden bu tür çalışmalara örnek oluşturma açısından elde olan verilerin sınıflandırılma işlemi yapay zeka metodları ile irdelenmiş ve çalışılmıştır. UCI Machine Learning internet sitesinden elde edilen verilerin normalizasyon işlemlerine tabi tutulması gerekmektedir. Normalizasyon tekniklerin seçimi veri için önem arz etmekte olup, ağır performans yüzdelerini direkt olarak etkilemektedir. Bunun nedeni ise normalizasyon işlemi sisteme girilen verilerin transfer durumunda fonksiyonun en aktif olan alanından aktarılmasını sağlar. Böylelikle yüksek değerlenen kümülatif toplamaların yaratacağı olumsuzlukların engellenmesi sağlanır. Log-Sigmoid normalizasyonunun seçilme nedeni yapılan literatür taramasında da görüldüğü üzere (Yavuz ve Deveci,

2012) elimizdeki veriler pozitif olduklarından dolayı bu metod seçilmiştir.

Bu çalışma farklı yapıda 3 adet veri kümesinin incelenmesi ve aynı algoritmaların bu verilere uygulanması itibarıyla hem başarı performansı olarak hem de zaman bakımından aralarındaki ilişkiyi gözler önüne sermektedir. Tohum verisi 207 adet her biri 69 sınıftan oluşan 3 sınıflı bir yapıya haizdir. Test ve eğitim kısmı 104 ve 103 adet olup kullanılan bilgisayar programlama dilinde tesadüfi olarak atanmıştır. Elde edilen sonuçlara bakılacak olursa ANFIS ile BKSA test performans yüzdeleri (%86.41 , %88.06) anlamlı bir fark oluşturmamıştır. Eğitim performans yüzdelerine bakılacak olursa eğer ANFIS %98.07, BKSA ise %98.36 olarak gerçekleşmiştir. Test performans değerleri gibi eğitim performans yüzdeleri arasında da anlamlı bir fark yoktur. Bu nedenle diğer bir karşılaştırma kriteri olan işlem zamanlarına bakmakta fayda görülmüştür. İşlem zamanı açısından BKSA 0.598 saniye ile 1.145 saniye olan ANFIS işlem zamanından daha kısa sürede gerçekleştirmiştir. Hem daha kısa sürede hem de performans açısından neredeyse %3'lük bir başarısı olan BKSA yine de elde olan verilerin eğitim kısmı sayıca az olduğundan %95 ve üstü aralığına girememiştir.

Diğer bir veri kümesi StatLog yani Uydu görüntüleme ile elde edilen bir veri kümesidir. 6435 adet veri olup, her bir 6 sınıfın 37 tane özniteliği vardır. Bir önceki veri kümesine nazaran daha fazla eğitim ve test kümesi ayrıldığından test performans yüzdeleri ANFIS de %100, BKSA da ise %99.92 olmaktadır. Eğitim performansları yüzdelerine bakılacak olursa eğer ANFIS yine %100'lük bir başarı sağlamış, BKSA ise %99.97'lik rakamla aralarında önemli ölçüde fark olmadığını ortaya koymuştur. Bu bağlamda işlem süresi açısından bakılacak olursa eğer; ANFIS 800 saniye iken BKSA 72 saniyede tüm sınıflandırma işlemini tamamlamıştır. Buradan da açıkça görüleceği üzere ANFIS her ne kadar %100'lük bir başarı sergilemiş olsa da BKSA %99.92 ile hem zaman açısından hem de performans başarı yüzdesi olarak ANFIS algoritmasından çok daha iyi bir örüntü tanıma sergilemiştir. Bu durumdan da anlaşılacağı üzere eğer sisteme eğitim ve test bakımından daha fazla sayıda veri yüklenirse sistem çok daha iyi bir performans sergilemekte ve araştırmacılara daha iyi sonuçlar vermektedir.

Son olarak incelenen veri kümesi ise 10 adet sınıftan oluşan 11 özniteliğe sahip Kırmızı Şarap Kalite değerlendirme veri kümesidir. Aynı internet sitenin en iyi ve eksiksiz olan verilerinden bir tanesi olan bu küme üzerinde yine aynı metodlar denenmiş hem performans hem de işlem zamanı açısından değerlendirilmeye tabi tutulmuştur. Bu veri kümesinin seçim nedeni ise önceki kümelerin sınıfları 3 ve 6 iken bunda 10 adet sınıf olması algoritmayı zorlayıcı etkenlerden bir tanesidir. Diğer kısıtlardan biri ise öznitelik sayısı (11 adet) StatLog veri kümesi

kadar (37 adet) olmasa da Tohum veri kümesinden (7 adet) fazla olması sebebiyle başka bir zorlayıcı etken olarak karşımıza çıkmaktadır. Sonuç itibarıyla performans yüzdeleri bakımından ANFIS eğitimde %100'lük, BKSA ise %99.9625 ile başarı sağlamıştır. Test performansları bağlamında ANFIS %99.975, BKSA ise %99.862 gibi birbirleri arasında çokta anlamlı bir fark oluşturmadığı açıkça görülmektedir. Performans yüzdeleri bakımından karşılaştırmanın anlamlı olmadığı görüldüğünden her bir algoritmanın işlem zamanlarına bakılması gerekmekte ve bununla bir sonuca varılabileceği düşünülmektedir. ANFIS hem bulanık hem de sinir ağlarının birleşimiyle oluşan bir algoritma olduğu için denenen diğer 2 veri kümesinde olduğu gibi işlem zamanında çok fazla süre harcamaktadır. Her bir sınıfın öznitelğine bağlı kalarak oluşturulan kural tabanlı bir sistem olan ANFIS 85.271 saniye ile BKSA'ya nazaran 4 hantal kalmaktadır. Bu durum zaman öncelikli çalışılan sistemlerde bir dezavantaj haline gelebilmektedir. BKSA ise 4 saniye gibi kısa bir sürede performans bakımından ANFIS metodundan geri kalmayarak hızlı bir şekilde örüntü sınıflandırma işlemini sorunsuzca halledebilmiştir. Bu da bu sistemin kullanımının ne kadar doğru bir yaklaşım olduğunu açıkça göstermektedir.

Yapılan çalışmalar neticesinde Basamak Korelasyon Sinir Ağı her açıdan karşılaştırılan metodlara göre daha iyi sonuçlar vermektedir. Burada önemli olan bu algoritmalar üzerinden başka araştırmacılar çalışma yapacağı zaman önemli bir hususu dikkate almalıdırlar. Buda şudur ki; elde edilen veri kümesi sayısal anlamda ne kadar fazla olursa hem eğitim için hem de test performansları bir o kadar iyi sonuçlar vermektedir. Tohum veri kümesinde 207 adet veri ile çalıştırılan algoritma %90'nın altında bir performans yüzdesi göstermiştir. Fakat hem öznitelik hem de eğitim ve test veri sayısı fazla olan veri kümelerinde performans yüzdeleri %100'lük bir başarı sağlamıştır. Eğer az sayıda veri kümesi ile çalışmak zorunda kalınırsa veya veri toplama ile ilgili bazı sıkıntılar meydana gelirse, çapraz doğrulama (cross-validation) yöntemi ile elde edilen veriler çoğaltılabilir ve performans yüzdeleri istenilen seviyeye çıkartılabilir.

Bir başka önemli husus ise algoritmaların işlem zamanıdır. Eğitim ve Test başarı yüzdeleri neredeyse aynı çıkan veri kümelerinde karşılaştırma adımının ikinci aşamasında her bir metodun sistem üzerinde geçirdiği işlem zamanları dikkate alınmıştır. ANFIS metodunun iki önemli yapıyı (Bulanık Mantık-Sinir Ağı) bir arada bulundurması sebebiyle her ne kadar performans yüzdelerinde başarılı bir seviye gösterse de Basamak Korelasyon Sinir Ağına (BKSA) göre çok hantal ve ağır işlemektedir. İşte bu yüzden ki; BKSA metodu ANFIS yapısından daha hızlı ve etkin çalışarak veri kümeleri üzerindeki başarılarını gözler önüne sermektedir. Çalışmanın amacına yönelik olan her bir farklı veri kümelerini hem performans hem de işlem zamanları açısından incelenip ortaya koyulması bu konu üzerinde

araştırma yapacak olan kişilere ışık tutması açısından büyük önem arz etmektedir. Bu çalışma konusu ile BKSA metodunun daha ileriki aşamalarda yüz tanıma, zaman serisi tahminlerinde, üretim sistem kurma aşamalarında, hat dengeleme problemlerinde ve biyo-teknolojik sistemlerin uygulanma ve kurulma aşamalarında bir yol çizilebileceği düşünülmektedir.

Kaynakça

- [1] Fahlman, S.E., Lebiere, C., 1991. The Cascade-Correlation Learning Architecture; Report CMU-CS-90-100; Carnegie Mellon University: Pittsburgh, PA, USA, s. 1-13
- [2] Augusteijn, M.F., Clemens, L.E., 1994, A performance evaluation of texture measures for image classification and segmentation using cascaded correlation architecture, IEEE Inter. Conf. Neural Networks IEEE World Congress Comput. Intell. 7, s. 4300-4305
- [3] McKenna, S.J., 1991 A Comparison of Neural Network Architectures for Cervical Cell Classification, IEEE, s. 105-109
- [4] Augusteijn, M., Skujca, T., 1993, Identification of human faces through texture-based feature recognition and neural network technology, In Proceedings of the IEEE International Conference on Conf. Neural Networks, s. 392-398
- [5] Burke HB, Goodman PH, Rosen DB, vd., 1997, Artificial neural networks improve the accuracy of cancer survival prediction., Cancer, s. 857-862
- [6] Prampero, P., Carvalho, A., 1999, Classifier Combination for Vehicle Silhouettes Recognition, Seventh International Conference on Image Processing and its Applications, IPA'99, s. 67-71
- [7] Jang, J.S.R, Sun, C.T., Mizutani, E., 1997, Neuro-fuzzy and soft computing, Prentice Hall, Upper Saddle River
- [8] Şenol, C., Yildirim T., 2009, Thyroid and Breast Cancer Disease Diagnosis using Fuzzy Neural Networks, 6th International Conference on Electrical and Electronics Engineering (ELECO'2009)
- [9] Sharma, M., Mukharjee, S., 2013, Brain Tumor Segmentation Using Genetic Algorithm and Artificial Neural Network Fuzzy Inference System (ANFIS), Advances in Computing and Information, Springer - Verlag Berlin Heidelberg, s. 329-339
- [10] Ziasabounchi N., Askerzade I., 2014, ANFIS Based Classification Model for Heart Disease Prediction, International Journal of Electrical and Computer Sciences. IJECS-Ijens
- [11] Ankişhan, H. ve Yılmaz, D., 2013, Comparison of SVM and ANFIS for snore related sounds classification by using the largest Lyapunov exponent and entropy, Computational and Mathematical Methods in Medicine
- [12] Rumelhart, D. E., Hinton, G. E., Williams, R. J. 1986, Learning Internal Representations by Error Propagation, Rumelhart, D. E. ve McClelland, J. L. J. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press
- [13] Schmitz, A., Courouble, F., Hefazi, H., Besnard, E., 2010, Modified Cascade Correlation Neural Network and its Applications to Multidisciplinary Analysis Design and Optimization in Ship Design, Machine Learning
- [14] Augusteijn, M.F., Clemens, L.E., 1994, A performance evaluation of texture measures for image classification and segmentation using cascaded correlation architecture, IEEE Inter. Conf. Neural Networks IEEE World Congress Comput. Intell. 7, s. 4300-4305
- [15] McKenna, S.J., 1991 A Comparison of Neural Network Architectures for Cervical Cell Classification, IEEE, s. 105-109