

## Tıbbi Karar Destek Sistemlerinin Yöntemsel Olarak Değerlendirilmesi Üzerine Bir Çalışma

Yüksel Yurtay<sup>1</sup>, Gülşah Ak<sup>1</sup>, Nihal Zuhal Bacınoğlu<sup>1</sup>

<sup>1</sup> Bilgisayar ve Bilişim Bilimleri Fak., Bilgisayar Müh. Bölümü, Sakarya Üniversitesi, Türkiye

### Özet

Tıbbi karar destek sistemleri alanında bugüne kadar pek çok çalışma yapılmıştır. Bu çalışmaların başarı yüzdesi gittikçe artmakta, dolayısıyla bu alandaki çalışmalar yükselen bir ivme kazanmaktadır. Birçok kişiden alınan tahliller, tanı bilgileri ve yapılan operasyonlardan elde edilebilecek veri topluluklarından, çeşitli algoritmalar aracılığıyla karar desteği sağlayacak sistemler oluşturmak mümkündür. Tıbbi karar destek sistemlerinin hazırlanmasında kullanılan yöntemler incelendiğinde, yapay sinir ağları, yapay bağışıklık sistemi ve veri madenciliği tekniklerinin sıklıkla kullanıldığı görülmektedir.

Bu çalışmada, değişik hastalık tanılarına ilişkin olarak geliştirilmiş tıbbi karar destek sistemlerinin, içerdiği teknikler açısından incelemesi yapılmış ve bu tekniklerin performansları değerlendirilmiştir. Benzer çalışmalar arasında karşılaştırma yapılarak, kullanılan teknikler kıyaslanmıştır.

**Anahtar Kelimeler:** Tıbbi Karar Destek Sistemleri, Yapay Sinir Ağları, Yapay Bağışıklık Sistemi, Veri Madenciliği.

### A Study on the Evaluation of Medical Decision Support Systems Methodology

#### Abstract

Many studies have been conducted until this day in the field of medical decision support systems. Success percentage of these studies is increasing, so studies in this field is gaining momentum. Analyzes which get from many people, diagnostic information and the data communities can be obtained from operations, decision support systems is possible to build with various algorithms. When used methods in the preparation of medical decision support systems are examined, observed that artificial neural networks, artificial immune system and data mining techniques are often used. In this study, the diagnosis of different diseases related developed medical decision support systems, according to included techniques are examined and performances of those techniques were evaluated. By comparison of between similar studies, the techniques used were compared.

**Keywords:** Medical Decision Support Systems, Artificial Neural Networks, Artificial Immune System, Data Mining.

## 1.Giriş

Doğru kararlar verebilmek için bilgiye ihtiyacımız vardır. Teknolojinin gelişmesiyle ve bilgi boyutunun artmasıyla bilgisayar programları beklenmedik durumlarda karşılaşınca kesin olmayan yargılar hakkında karar vermede ve bilgiye gereksinim duyulan karar noktalarında yetersiz kalmaktadır. Karar Destek Sistemleri karar verme aşamasında karar vericilere yardımcı olan ve kendi kendine karar veremeyen sistemlerdir.

Çağımızın en önemli araştırma alanları sağlık ve tıptır. Bu alanda daha etkili, daha hızlı, daha doğru karar vermek ve karar vermeyi kolaylaştırmak için tasarlanmış değişik model ve uygulamaları içeren sistemlere Tıbbi Karar Destek sistemleri denir. Yaptığımız bu çalışmada Tıbbi Karar Destek sistemlerinden yola çıkılarak kullanılan teknikleri şu şekilde sıralayabiliriz: Veri Madenciliği, Yapay Sinir Ağları ve Yapay Bağışıklık Sistemleri.

Veri madenciliği konusunda birçok tanım yapılabilir. Basit bir tanım yapmak gerekirse, veri madenciliği, büyük ölçekli veriler arasından “değeri olan” bir bilgiyi elde etme işidir. Bu sayede veriler arasındaki ilişkileri ortaya koymak ve gerektiğinde ileriye yönelik kestirimlerde bulunmak mümkün görünmektedir. Bunun anlamı, veri madenciliği, bir kurumda üretilen tüm verilerin belirli yöntemler kullanarak var olan yada gelecekte ortaya çıkabilecek gizli bilgiyi su yüzüne çıkarma süreci olarak değerlendirilebilir[1]. Bu açıdan, veri madenciliği tekniğinin tıbbi karar destek sistemleri için önemli bir yere sahip olduğunu söyleyebiliriz.

Tıbbi veri madenciliğinin diğer alanlardan farkı verilerin heterojen olmasıdır. Tıbbi veri madenciliği metotları veri kaynaklarını hem teknik hem de sosyal sebeplerden ötürü oluşan kayıp değerlerin yayınlaşmasını heterojen yapıda ele almalıdır[2]. Bu çalışmada ele alınan 5 makalelerde, veri madenciliği yöntemlerinden sınıflandırma ve kümeleme yöntemleri kullanılmıştır.

Yapay Sinir Ağları, insan sinir hücresini taklit etmeye yönelik bir girişimin sonucu olan Yapay Sinir Hücre’lerinin gruplanmasıyla oluşmuştur. Yapay Sinir Ağları, insan beyninin özelliklerinden olan öğrenme yoluyla yeni bilgiler türetebilme, yeni bilgiler oluşturabilme ve keşfedebilme gibi yetenekleri herhangi bir yardım almadan otomatik olarak gerçekleştirmek amacıyla geliştirilen bilgisayar sistemleridir[3].Yapay Sinir Ağları’nın temel amacı bilgisayarların öğrenmesini sağlayıp, benzer olaylar karşısında benzer kararlar vermeye çalışmaktır. Çalışmamızda tıbbi alandaki altı farklı makalede Yapay Sinir Ağları yöntemlerinden tahmin amaçlı, kategori ayırma amaçlı ve optimizasyon amaçlı metotlar arasından tahmin amaçlı ve kategori ayırma amaçlı metotlar kullanılmıştır.

Yapay Bağışıklık Sistemleri biyolojik temelli hesaplama yöntemi olarak 1990’da ortaya çıkmıştır[4].Yapay Bağışıklık Sistemleri, bağışıklık sisteminden esinlenen ve gerçek hayat problemlerini çözmek için geliştirilen zeki ve uyum sağlayabilen sistemlerdir[5]. Öğrenme algoritmasının sistem özellikleri insan vücudundaki mikropları tanıma ve yok etme

yeteneğinden ilham almıştır[6]. Bu çalışmada yapay bağışıklık sisteminin tıbbi tanı amaçlı kullanımına ilişkin 5 makale yer almaktadır.

## 2. Geliştirilmiş Çalışmalar

### 2.1. Veri Madenciliği Alanında Yapılan Çalışmalar

2010 yılında yapılan bir çalışmada ise Parkinson Hastalığının tanısı için çoklu sınıflandırma metotları karşılaştırılmıştır[7]. Bu çalışmanın amacı %25 gibi yüksek bir yanlış teşhis ihtimali olduğu için sağlıklı bireylerin ayırt edilmesidir. Burada dört bağımsız yöntem kullanılmış ve karşılaştırılarak bir çalışma yapılmıştır. Veri Madenciliği yöntemlerinden karar ağaçları ve regresyon bu yöntemlerdendir. Parkinson hastalığı veri kümesi sağlıklı kişiler ve hastaları karakterize etmek için kullanılacak özelliklere sahiptir. Daha önce de bahsedildiği gibi, veritabanı 23 sütun ve 197 satır oluşmaktadır. Bu veri kümesi, NCVS(National Centre for Voice), Denver, Colorado ile işbirliği içinde kaydedilen konuşma sinyalleriyle Oxford Üniversitesi'nden Max Little tarafından oluşturulmuştur. Orjinal çalışmada genel ses bozuklukları için özellik çıkarma yöntemleri yayınlanmıştır. Bu veri kümesi, 23'ü Parkinson hastası(PD) 31 kişinin biyomedikal ses ölçümleri dizisinden oluşur. Tablo içindeki her bir sütun , belirli bir ses ölçüsüdür ve her satır bu bireylerin (adı sütununda) 195 ses kaydından birine karşılık gelir. Verilerin temel amacı, durum sütununa göre sağlıklı için 0 ve PD için 1 olarak ayarlanmasıdır. Veri ASCII CSV biçimindedir. Her satır bir ses kaydı örneğine karşılık bir CSV dosyasını içerir.

Bu çalışmada ROC(Receiver Operating Characteristic=Alıcı İşletim Karakteristliği) analizi kullanılmıştır. ROC analizi, bir duyarlılık ve seçicilik değeri kullanılarak tanı koymanın getirdiği sakıncaları ortadan kaldırmak için geliştirilmiş istatistik değerlendirme yöntemidir [8]. Kısacası ROC doğru pozitiflerin, yanlış pozitiflere olan oranıdır. Tablo 1'de ROC analizinde kullanılan parametreler gösterilmiştir.

**Tablo 1.** ROC analizi için kullanılan parametreler

Test Sonucu	Gerçek Durum		
	Pozitif	Negatif	Toplam
Pozitif	Pozitif Durum(DP)	Yanlış Pozitif(YP)	(DP+YP)
Negatif	Yanlış Negatif(YN)	Doğru Negatif(DN)	(YN+DN)
Toplam	(DP+YN)	(YP+DN)	(DP+YN+YP+DN)

$$\text{Doğruluk} = \frac{DP+DN}{DP+DN+YN+YP}$$

$$\text{Duyarlılık} = \frac{DP}{DP+YN}$$

$$\text{Seçicilik} = \frac{DN}{DN+YP}$$

Tanı testlerinde olumlu ya da olumsuz kararın doğruluk derecesi önemlidir. Pozitif ya da negatif kararların her biri için doğruluk düzeyini gösteren ölçütler vardır. Duyarlılık, testin,

gerçek pozitif durumlar içinden pozitif olan durumları ayırma yeteneğidir, seçicilik ise testin gerçek negatif durumlar içinden negatif olan durumları ayırma yeteneği olarak ifade edilebilir [9]. [7] çalışmasında uygulanan sınıflandırıcıların puan hesaplaması için çeşitli performans değerlendirme testleri kullanıldı. Regresyon 88,6'lık derecesiyle karar ağacının 84,3'lük derecesine göre daha iyi bir skor elde etmiştir. Öte yandan regresyon düğümü eğitim bölümünde karar ağacına göre daha kötüdür.

2011 yılında baş ağrısı tanısı üzerine yapılan bir başka çalışmada veri madenciliğinde en iyi açık kaynak kodlu sistemlerden biri olan RapidMiner kullanılmıştır[10]. Bu çalışmada aurasız muhtemel migren için daha hızlı analiz ve sonuç üreten bir karar destek sistemi geliştirilmiştir. K-Means algoritması ile tanı kümeleri oluşturulmuştur. Çalışmada kullanılan veri kümesindeki hesaplamalara göre aurasız migren, aurasız muhtemel migren ve migren yok olarak 3 küme vardır. K-Means Algoritması basit bir iterasyon yöntemidir ve verilen veri kümesi içindeki nesnelere kullanıcının belirttiği sayıda kümeye böler. Baş ağrısı veri seti www.migbase.com web sitesinde 2012 Mart ayında yayınlamıştır ve 353 öğrencinin bilgisi cinsiyet, alkol, yaş, tütün ve günlük bilgisayar kullanımı süre gibi sorular muhtemel aurasız migren için sorulmuştur. Bu anketler Uluslararası Baş ağrısı Derneği kriterlerine göre hazırlanmıştır. Sonuçlar Uzman nörolog tarafından avantaj ve dezavantajları incelenerek değerlendirilmiştir. Yapılan çalışma sonucunda 65 öğrencide baş ağrısı tanısı yokken 288 öğrencide var olduğu tespit edilmiştir.

2012 yılında kadınlardaki demir eksikliğine bağlı olarak kansızlık tanısına ilişkin bir çalışma yapılmıştır[11]. Zonguldak Devlet Hastanesi'nin 2010 yılına ait 2640 tane verisi kadın hastalarda demir eksikliğine bağlı kansızlık tanısı için laboratuvar kan analiz sonuçları kullanılmıştır. Bu verilerde 567 kadında kansızlık varken 2073 kadında ise kansızlık yok tanısı bulunmuştur. Bu çalışmada Gini algoritması kullanılmıştır. Gini Algoritması ikili bölünmeler şeklinde gerçekleşen bir sınıflandırma yöntemi olup, ikili yineleme bölümler için en iyi bilinen kurallardandır. Her bir ağaç farklı bir stil ile gelişir. Algoritma nitelik değerlerinin sol ve sağda olmak üzere ikili bölünmeler şeklinde ayrılması temeline dayanır[1]. Hesaplanan her düğüm sol ve sağ bölümleri için ayrı ayrı gerçekleştirilir. Çalışmanın performansı için ROC(Receiver Operating Characteric) analizinden yararlanılmıştır. Bu çalışmada 2640 adet veri temizlendikten sonra geriye 2599 veri kalmıştır. Bu veriler 3 kutuya rastgele olarak ayrılmıştır. 2 kutu eğitim ve 1 kutu test olasılıklarının tümünü değerlendirilecek ortalama doğruluk değerleri elde edilmiştir. Bu çalışmada kadınlardaki demir eksikliğine bağlı kansızlık tanısı amacı için Karar ağacı Gini algoritması ile geliştirilmiştir ve % 92.97 doğrulukla tanı sağlanmıştır.

Bir başka çalışmada veri madenciliği yöntemlerinden karar ağacı yöntemi kullanılarak Tiroid hastalığının teşhisi amaçlanmıştır [12]. Karar ağaçlarının oluşturulmasında öğrenme yöntemlerinden Gini algoritması, değerlendirilmesinde ise ROC analizi kullanılmıştır. Öğrenme seti UCI veri tabanındaki Tiroid veri setidir. Veri setinde 6 nitelik değeri gösterilmiştir. Bunlar 5 giriş 1 sınıf niteliği olup sınıflar, “normal”, ”hyper” ve “hypo” olarak belirlenmiştir [13]. Sınıflardaki veri sayıları normal=150, hyper=35, hypo=30 olarak

verilmiştir. Bu çalışmada 215 adet veri 3 bölüme ayrılmıştır. Bunların 2 bölümü öğrenme ve 1 bölümü test olacak şekilde 3 kez Gini algoritması kullanılmıştır Genel doğruluk değeri % 94,5, genel duyarlılık değeri % 85,3 ve genel seçicilik değeri de % 97.3 olarak elde edilmiştir.

2013 yılında yapılan çalışmada ise diyabet hastalığının belirlenmesi için farklı yöntemler karşılaştırılmıştır[14]. Bu çalışmada veri kümesi Arizona çevresinde yaşayan Pima yerlilerinden alınmıştır. Pima yerlileri homojen bir grup olduklarından bu insanlardan alınan veriler diyabetikte yoğun bir çalışma konusu olmuştur. Çalışmada kullanılan veri seti, 8 giriş, 1 çıkış verisi içermektedir. Çıkış bilgileri ya da sınıflandırma değerleri, 0 ise diyabet yok, 1 ise diyabetir [15]. Bu çalışmada veri madenciliği yöntemlerinden karar ağacı ve karar ağacı yöntemi altında da Gini algoritması kullanılmıştır. Ayrıca ROC analizi yapılmıştır. Bu çalışmada kullanılan yöntem sonucunda doğruluk da % 65.97, duyarlılık da % 44.71 ve özgünlük de % 77.78 başarı elde edilmiştir.

## **2.2.Yapay Sinir Ağları**

Yapay sinir ağları kullanılarak yapılmış bir çalışmada, 2006 yılında etkin tanı ve bulanık tabanlı reçete ile kadın hastalıklarında homeopatik tıbbi sistem kullanılarak sinir ağı kavramına dayalı bir karar destek sistemi çalışması geliştirilmiştir[16].Bu çalışmanın amacı, hastalığın birincil ve ikincil belirtilerine dayanarak jinekolojik bir hastalık için ilaç önerecek bir karar destek sistemi geliştirmektir. Çok katmanlı ağırlık odaklı ve ileri beslemeli bir yapay sinir ağı sistemi mevcuttur. Sistem gereksinimlerinin ayrıntılı bir analizi yapılmıştır. Gerekli kural ve koşullar, veritabanı gibi gereksinimlerin belirlenmesinden sonra tasarım aşamasına geçilmiştir. Sinir ağı sisteminin tasarımında ilk adım programın çekirdeğini oluşturan uzman bilgi koleksiyonu olmuş ve sistem bu bilgi etrafında inşa edilmiştir. İnsan beyni nöronlarını taklit edecek şekilde tanecikli bir biçimde veritabanında saklanmış ve karar verme kuralları için birkaç nöron birleştirilerek oluşturulmuştur. Böylece, insan beyninde üretilen sinaps taklit edilmiştir. Çalışma sonucunda klinik karar destek sistemi Jinekolojik hastalıkların tedavisi için tıp homeopatik sistemi ile ilgili hızlı ve doğru kararlar için kullanılmak üzere öngörülmüştür. Karar destek sistemi benzer ve canlı her iki veri tipiyle test edilmiştir. Karar destek sistemi tarafından tanı ile ilgili alınan kararlar ve manuel yollarla hesaplanan ilaçların reçete ile uyumlu olduğu görülmüştür. Aynı zamanda yeterli belirti girilmediğinde aktif olan sistemin bulanık bileşeninin beklendiği gibi çalıştığı tespit edilmiştir.

2009 yılında yaşlı hastalarda baş ağrısı prognozunun tahmini için bir çalışma yapılmıştır. Çalışmadaki veri seti Mart 1999 ve Ocak 2005 tarihleri arasında Mersin Üniversitesi Tıp Fakültesi'ne baş ağrısı kliniğine başvuran 341 hastadan oluşturulmuştur. Bu veri setinin % 75 (257) eğitim seti olarak alınırken % 25'de (84)test seti için alınmıştır [17].Çalışmada sinir ağları uygulandığı zaman içinde baş ağrısı tam olarak iyileşen yaşlı hastaları belirleyip ROC analizi ile sınıflandırma performansı incelenmiştir. Yaşlı hastalarda baş ağrısı prognozunu tahmin etmek için çok tabakalı perseptron sinir ağı kullanılmıştır. Veri seti eğitim ve test

verisine ayrıldıktan sonra geriye yayılım algoritması kullanılmıştır. Araştırmanın sonucunda sinir ağı modellerinin hem eğitim hem de test verilerinde yüksek performansa sahip oldukları görülmüştür. Uygulanan metotlar sonucunda tamamen iyileşmiş hastaların oranı 0.10' dan daha az olduğu için, tüm modellerin hassasiyetlerinin (Eğitim seti için 0.33, 0.39 ve 0.50; test seti için 0.25, 0.00 ve 0.50; genel set için 0.47, 0.37 ve 0.36) özgünlüklerine göre daha az olduğu görülmüştür (Eğitim seti için 0.99 1.00 ve 0.99 ; test seti için 0.99, 0.97 ve 0.97 ; genel set için 0.99, 0.99, ve 0.99 dır). Eğitim seti ve genel veri setindeki her bir dönem için pozitif tahmini değer 0.67 ve 1.00 arasında ve negatif tahmini değer ise 0.93 ve 0.96 arasında bulunmuştur.

2009 yılında yapılan başka bir çalışmada da tiroit hastalığının tanısında kullanılan sinir ağları üzerinde karşılaştırma yapılmıştır[18]. Çalışmadaki veri seti UCI 'den alınmıştır[13]. 3 farklı türde sinir ağı kullanılmıştır. Bunlar çok katmanlı sinir ağı(MLNN), olasılıklı sinir ağı(PNN) ve öğrenme vektör niceleme sinir ağıdır(LVQ-NN). 3 kat ve 10 kat çapraz doğrulama tekniği sinir ağı modellerinin doğruluğunu karşılaştırmak için yapılmıştır. Tiroit hastalığı veri kümesi için bu çalışmadan elde edilen sınıflandırma doğrulukları Tablo 2'de verilmiştir. Tablo 2'de görüldüğü gibi 3-kat ve 10- kat doğrulama yaklaşımları bu çalışmada kullanılan aynı yöntemler için hemen hemen benzer sınıflandırma sonuçlarını vermiştir.

**Tablo 2.** Tiroit hastalığı veri seti için sinir ağları sınıflandırma doğrulukları

Yöntem	Sınıflandırma Doğruluğu (%)
MLNN with LM (3 * FC)	92.96
PNN (3 * FC)	94.43
LVQ (3* FC)	89.79
MLNN with LM (10 * FC)	93.19
PNN (10* FC)	94.81
LVQ (10 * FC)	90.05

2010 yılında Parkinson hastalığının tanısı için çoklu sınıflandırma metotları karşılaştırılmıştır ve bu çalışmada Yapay Sinir Ağları yöntemlerinden sinir ağı ve DMNeural kullanılmıştır[7]. Sinir ağı düğüm özellik alanını sınıflandırmak için kullanılmıştır. Sinir ağı mimarilerinin bir çok türü vardır ancak çok katmanlı ileri beslemeli sinir ağı türü öngörü için en yaygın kullanılan türdür. Çok katmanlı ileri beslemeli ağlarda, nöronlar katmanlar halinde düzenlenmiştir ve diğer katmanlar ile nöronlar arasında bir bağlantı vardır. Geri yayılım öğrenme algoritması olarak, ileri beslemeli tek gizli katmanlı sinir ağı kullanılmıştır. Çalışmada kullanılan algoritmanın türevleri Levenberg-Marquardt (LM), ölçekli eşlenik (SCG) ve Pola-Ribiere eşlenik (CGP) algoritmaları vardır. Çalışmada kullanılan sınıflandırma yöntemlerinden ROC analizi ile elde edilen en iyi sonucu, hem eğitim sınıflandırması bölümünde %100 oran ile hem de test sınıflandırması bölümünde %92,9'luk oran ile Sinir Ağı yöntemi vermiştir.

2013 yılında yapılan bu çalışmada diyabet hastalığının tahmin edilmesinde farklı yöntemlerin karşılaştırılması yapılmıştır[14]. Veri madenciliği başlığı altında anlatılan [14] nolu makaledeki bu çalışmada kullanılan veri seti hakkında bilgi verilmiştir. Çalışmada olasılıklı

sinir ağı(PNN) , öğrenme vektör niceleme sinir ağı(LVQ), ileri beslemeli sinir ağı(FFN), basamaklı sinir ağı(CFN), dağıtılmış zaman gecikmeli sinir ağları(DTDN), zaman gecikmeli sinir ağları(TDN) olmak üzere altı tane sinir ağı veri seti üzerinde uygulanmıştır. ROC analizi ile sınıflandırılması yapılmıştır. Aşağıdaki Tablo 3’de kullanılan sinir ağı yöntemlerinin sonuçları verilmiştir.

**Tablo 3.** Sinir ağı yöntemleri sonuç analizleri

Yöntem	Doğruluk( %)	Duyarlılık(%)	Özgünlük(%)
PNN	72.00	63.33	76.88
LVQ	73.60	54.44	84.38
FFN	68.80	54.44	76.88
CFN	68.00	62.22	71.25
DTDN	76.00	53.33	88.75
TDN	66.80	41.11	81.25

2013 yılında yapılmış bir diğer çalışmadaki amaç, sağlıklı insanlar için gelecekte kanser türlerine yakalanma risklerini belirlemek ve pilot kanser türleri belirleyerek ön teşhis üzerinde çalışmaktır[19]. Meme kanseri, akciğer kanseri ve kolon kanseri pilot kanser türleri olarak seçilmiştir. Bulanık mantık modelinin performans ölçümünü test etmek için Onkoloji Hizmetlerinden, hem kanser tanısı konulmuş bireyleri hem de sağlıklı bireyleri içeren toplam 120 veri kümesi meme kanseri için, 140 veri kümesi akciğer kanseri için ve 110 veri kümesi kolon kanseri için alınmıştır. Çalışmanın bir diğer amacı düzenlenen verilerin ışığında sağlık alanında bulanık mantık modelinin kullanılabilirliğini araştırmak ve Mamdani yöntemi ile önerilen bulanık mantık yönteminin performans farklılıklarını değerlendirmektir. ROC analizi doğruluk oranının, kolon kanseri için 0.83, meme kanseri için 0.81 ve akciğer kanseri için 0.80 olduğu tespit edilmiştir. Yapılan çalışmada bulanık mantık yöntemi ile meme kanserinde %80.83, akciğer kanserinde %80, kolon kanserinde %82.72 başarı sağlanmıştır. Bu çalışma sayesinde insanların kansere yakalanmalarına karşı önlemler alma şansına sahip olabilecekleri ve kansere yakalanma riskinin azabileceği görülmüştür

### **2.3.Yapay Bağışıklık Sistemleri**

Bu alanda yapılan bir çalışmada, 2007 yılında hepatit hastalığının tahmini için temel bileşenler (PCA) ve yapay bağışıklık sistemi(AIRS) kullanılmıştır[20]. Çalışmadaki veri Bu çalışma iki aşamadan oluşmuştur. İlk aşamada veri setlerinin özellik sayısı 19’dan 5’e temel bileşenler analizi ile azaltılmıştır. Daha sonrasında veri seti [0,1] aralığı arasında normalize edilmiştir. Normalize edilmiş giriş değerleri AIRS sınıflandırma sistemi kullanılarak sınıflandırılmıştır. Sistem performansı sınıflandırması doğruluk, duyarlılık ve özgüllük açısından analiz edilmiştir. Bu kriterler 10 kat çapraz doğrulama ile elde edilmiştir. Doğruluğu % 94.12, duyarlılığı % 100 ve özgüllüğü ise % 94.44 bulunmuştur. Test sonuçları hekimin doğrudan tanısından bulunduğu sonuçlar ile karşılaştırılmıştır ve uyumlu olduğu görülmüştür.

2007 yılında yapay bağışıklı sistemi ve Hızlı Fourier Dönüşümü- Welch Yönteminden elde edilen karotis arter doppler sonogramlarının maksimum hattını aterosklerozun ve sağlıklı bireylerin arasındaki farkı ayırt etmek için kullanılması üzerine bir çalışma yapılmıştır[21]. Bu çalışmada iki amaç öne çıkmıştır. Bunlardan birincisi, ateroskleroz araştırmalarını ilerletmek için çok daha etkili bir Doppler tekniği kullanımı ve önde gelen bir tanı sistemi geliştirmek olmuştur. İkinci hedef ise tıbbi sınıflandırma sisteminde gerçek dünyada yapay bağışıklık sistemini kullanmak ve problemdeki bu yapay zeka alanının etkisini göstermektir. Tüm veriler rastgele birbirini dışlayan ve yaklaşık olarak eşit büyüklükte 10 altkümeye bölünmüştür. Bütün veri kümesi 60 hasta ve 54 sağlıklı bireyin verisini içermektedir. Sınıflandırma probleminde kullanılan yapay bağışıklık sisteminin yalnızca ikili sistem birimleri ile 10 kat çapraz doğrulama yöntemi (CV) kullanılarak önemli ölçüde sınıflandırma zamanının azaldığı görülmüştür. % 99.33'lük sınıflandırma doğruluğuna ulaşılmıştır

2009 yılında yapılan başka bir yapay bağışıklık sistemi tiroit hastalığının tanısında kullanılmıştır[22]. Bu çalışmada bilgi kazancı yapay bağışıklık tanıma sistemi (IG-AIRS) olan yapay bağışıklık sistemi kullanılmıştır. IG-AIRS, sekil-uzay gösterimdeki Öklit uzaklık hesaplanmasında tüm özellikleri dikkate alarak olumsuz etkileri en aza indirmiştir. Bu çalışma iki basamaktan oluşur. İlk basamakta veri özelliklerinin bilgi kazanç değerleri elde edilmiştir. İkinci aşamada ise bu değerlere yapay bağışıklık tanıma sistemi (AIRS) Öklit uzaklığına dayalı bilgi kazancı uygulanmıştır. Çalışmada kullanılan veri seti [13] den alınmıştır. Yapay bağışıklık sistemi 10-kat CV yöntemiyle % 95.90'lık sınıflandırma doğruluğuna ulaşmıştır.

2009 yılında yayınlanan bir makalede bilgi kazancı tabanlı yapay bağışıklı tanıma sistemi (IG-AIRS) adında yeni bir yapay bağışıklık sisteminin tıbbi uygulaması yapılmıştır[23]. Tıbbi veri için önerilen mikroorganizma veri seti sistemin performans analizlerinde uygulanmıştır. Mikroorganizma veri seti Cyranose 320 Elektronik Burun kullanılarak elde edilmiştir. Çalışmada önerilen sistem beş kat çapraz doğrulama yöntemi ile % 92.35'lik sınıflandırma doğruluk derecesine ulaşılmıştır. Sonuç olarak IG-AIRS, laboratuvar testlerine dayanan mikroorganizma türlerinin sınıflandırmasında yardımcı olmuştur.

2013 yılında yapay bağışıklık sistemi kullanılarak yapılan bir diğer çalışmada ise, diyabet hastalığının tanısı amaçlanmıştır[14]. Veri madenciliği başlığı altında anlatılan [14] nolu makaledeki bu çalışmada kullanılan veri seti hakkında bilgi verilmiştir. Yapay bağışıklık sistemi sınıflandırıcısı, duyarlılık % 52.22, doğruluk % 68.8 ve özgüllük % 78.13 olarak sonuçlar sağlamıştır.

### 3.Sonuç

Sınıflandırıcıların kullanımının yaygınlaşmakta olduğu alanlardan biri de tıp uygulamalarıdır. Birçok kişiden alınan tahliller ve yapılan operasyonlardan elde edilebilecek veri topluluklarından tıbbi karar destek sistemlerinin geliştirilmesi mümkündür. Bu çalışmada, 2006-2013 yılları arasındaki tıbbi sınıflandırıcıların kullanıldığı makalelerin bazıları üzerinde



bir inceleme yapılmış, bu makalelerde yer alan veri madenciliği, yapay sinir ağları ve yapay bağışıklık sistemi yöntemlerinin kullanımı ve performansları değerlendirilmiştir. Bu yöntemlerin, diyabet, tiroit, kanser çeşitleri, Parkinson, migren gibi hastalıkların tanısında yüksek başarı seviyesine ulaştığı görülmüştür.

### **Kaynakça:**

- [1] Özkan Y. Veri Madenciliği Yöntemleri. Papatya Yayıncılık;2008.
- [2] Krzysztof j. Cios, G. Willism Moore. Uniqueness of medical data mining. Artificial Intelligence in Medicine 2002; 26:1-24.
- [3] Öztemel E. Yapay Sinir Ağları. Papatya Yayıncılık; 2012.
- [4] O. Engin, A. Döyen. Artificial Immune Systems and Applications in Industrial Problems. G.U. Journal of Science, Vol. 2004; 17, pp. 71-84.
- [5]Machado, R.B., Boukerche, A., Sobral, J.B.M.,Juca, K.R.L. ve Notare, M.S.M.A., A Hybrid Artificial Immune and Mobile Agent Intrusion Detection Based Model for Computer Network Operations, 19th IEEE International Parallel and Distributed Processing Symposium 2005.
- [6] L. N. Castro, J. Timmis, Artificial Immune Systems: A New Computational Intelligence Approach 1st ed., London, Springer, 2002.
- [7] Das R., A comparison of multiple classification methods for diagnosis of Parkinson Disease , Expert Systems with Applications 2010; 37:1568-1572.
- [8]Tomak. L., Bek.Y. , İşlem Karakteristik Eğrisi Analizi ve Eğri Altında Kalan Alanların Karşılaştırılması, Journal of Experimental and Clinical Medicine 2010; Vol:27,no:2, s:58-65.
- [9] Dirican A. , Evaluation of the diagnostic test's performance and their comparisons, Cerrahpaşa J Med, 2001; 32(1):25-30.
- [10] Yurtay N., Çelik U., Yurtay Y. Headache Diagnosis with K-Means Algorithm. Awerprocedia Informationn Technology & Computer Science.2012; s:1074-1081.
- [11] Yurtay Y., Yılmaz Z., Çiftçi Ö.,Ayar K., Kadınlarda Demir Eksikliğine Bağlı Kansızlık Tanısına İlişkin Bir Veri Madenciliği Çalışması,ISTEC,2012-Dubai,s:762-766.
- [12] Yurtay N., Adak MF, Dural D., Serttaş S. A study on use of decision tree method in the diagnosis of thyroid disease. ISTEC Dubai 2012;s:768-774;.
- [13] UCI , <http://archive.ics.uci.edu/ml/machine-learning-databases/thyroid-disease/new-thyroid.names>
- [14] Bozkurt M.R., Yurtay N., Yılmaz Z., Sertkaya C., Comparison of different methods for determining diabetes disease, TÜBİTAK 2013;doi:10.3906/elk-1209-82.
- [15] UCI Machine Learning, "Pima Indians Diabetes Data Set", Available: <http://archive.ics.uci.edu/ml/machine-learning-databases/pima-indians-diabetes>, Accessed: 12.07.2012
- [16] Mangalampalli A., Mangalampalli S.M., Chakravarthy R., Jain A.K.,A neural based clinical decision-support system for efficient diagnosis and fuzzy-based prescription of gynecological disease using homoeopathic medicinal system, Expert Systems with Applicarions2006;30:109-116.
- [17] Taşdelen B., Helvacı S., Kaleağası H, Özge A., Artifical Neural Network Analysis for Prediction of Headache Prognosis in Elderly Patients. Tübitak 2009;39(1):5-12.
- [18] Temurtas F., A comparative on thyroid disease diahnosis using neural network, Expert Systems with Application 2009;36:944-949.
- [19] Yılmaz A., Ayan K. Cancer risk analysis by fuzzy logic approach and performance status of the model. Turkish Journal of Electrical Engineering & Computer Sciences 2013;

doi:10.3906/elk-1108-22.

[20] Polat K., Güneş S., Prediction of hepatitis disease based on principal component analysis and artificial immune recognition system. *Applied Mathematics and Computation* 2007;189:1282-1291.

[21] Latifoğlu F., Şahan S., Kara S., Güneş S. Diagnosis of atherosclerosis from carotid artery Doppler signals as a real-world medical application of artificial immune systems. *Expert Systems with Applications* 2007; 33:786-793.

[22] Kodaz H., Özşen S., Arslan A., Güneş S., Medical application of information gain based artificial immune recognition system(AIRS): Diagnosis of thyroid disease, *Expert Systems with Applications* 2012;36:3086-3092.

[23] Kara S., Aksebzeci B.H., Kodaz H, Güneş S., Kaya E., Özbilge H., Medical application of information gain-based artificial immune recognition system(IG-AIS): Classification of microorganism species. *Expert Systems with Applications* 2009;36:5168-5172.