

SAKARYA ÜNİVERSİTESİ

# Veri Madenciliği Uygulamaları

---

Hafta 8

Yrd. Doç.Dr. Nilüfer YURTAY



# Sınıflandırma- Yapay Sinir Ağları ve Yapay Bağışıklık Sistemi

## 8.1 Giriş

Yapay sinir ağları insan beyninin çalışma sisteminin yapay olarak benzetimi çabalarının bir sonucu olarak ortaya çıkmıştır.

En genel anlamda bir YSA insan beynindeki bir çok sinir hücresinin yada yapay olarak basit işlemcilerin birbirlerine değişik etki seviyeleriyle bağlanmasından oluşan karmaşık bir sistem olarak düşünülebilir.

YSA kavramı beynin çalışma ilkelerinin sayısal bilgisayarlar üzerinde taklit edilmesi fikri ile ortaya çıkmıştır. İlk çalışmalar beyni oluşturan biyolojik hücrelerin (literatürdeki ismi ile nöronların) matematiksel olarak modellenmesi üzerinde yoğunlaşmıştır. İnsan beyni geleneksel bilgisayarlardan tamamen farklı bir yolla işlem yapar. Oldukça kompleks, non-lineer ve paralel dağıtık bir yapıya sahiptir. YSA, beynin belirli bir işi veya fonksiyonu gerçekleştirme yöntemini/yolunu modellemek için dizayn edilen bir yapıdır. YSA, tecrübe veya eğitim yolu ile elde edilen bilgiyi saklamak ve kullanmak için uygun hale getirmeye yarayan basit işlem elemanlarından oluşan paralel dağıtık bir yapıdır. YSA'nın gücü paralel işlem yapabilme, öğrenme ve genelleme kabiliyetinden ileri gelir.

YSA, karşılaştığımız çeşitli problemlerde, beynin matematiksel gücünü, matematiksel işleyişini makineler/bilgisayarlara kazandırmaya, bu güçten faydalanmaya ve modellemeye çalışır. YSA, öğrenme, ilişkilendirme, sınıflandırma, genelleme, özellik belirleme, örüntü tanıma, optimizasyon, vb. konularda başarılı bir şekilde kullanılmaktadır. Bir YSA yapısının çözebileceği problem uzayı, insan beyninin çözebildiği problem uzayının oldukça kısıtlanmış bir alt kümesidir. İnsan beyni görüntü tanıma, algılama, öğrendiklerinden yeni bilgi türetebilme, keşfedebilme vb. konularda günümüz bilgisayarlarından oldukça hızlıdır ve daha üstündür. YSA yapıları, beynin biyolojik yapısının anlaşılmasına yardımcı olabilmek için bir model olarak ta kullanılabilir.

İnsan beyninin çalışma mekanizmasını taklit etmeye çalışan YSA sistemleri günümüzde birim işlem zamanı nanosaniyeler mertebesinde olan silikon mantık kapıları ile gerçekleştirilebilirler de insan beyninin birim işlem zamanı milisaniyeler mertebesindeki sinirlerinin toplu biçimde ele alındıkları işlevselliklerinden çok uzakta kalırlar. Beyin enerjisi kullanımı bakımından oldukça verimli çalışan bir sistemdir. Dijital bilgisayara kıyasla işlem başına milyarlarca kez daha az enerji harcar.

Beynin yüksek seviyeli bilgiyi nasıl sembolize ettiği bilinmemekle beraber, kendi arasında yüksek yoğunlukta bağlı bir çok yavaş üniteler kullandığı bilinmektedir. YSA, beynin bu yüksek yoğunlukta paralel ve yüksek yoğunluklu ara bağlantılı yapısını modellemeye çalışır. Beynin temel inşaa blokları olan biyolojik sinirler, silikon mantık kapılarından daha yavaştır. Sinirler milisaniyeler mertebelerinde işlerken, silikon mantık kapıları nanosaniyeler mertebelerinde (belki daha düşük sürede) işler

YSA, iki yönden beyne benzer :

- Bilgi ağ tarafından bir öğrenme işlemi yolu ile kazanılır.
- Sinaptik ağırlıklar olarak bilinen sinirler arası bağlantı güçleri, bilgi saklamak için kullanılır.

Beyin, en hızlı dijital bilgisayardan kat kat daha hızlı olarak örüntü tanıma, algılama, kavrama, sezme, vb. görevleri yerine getirir;

- Çevre ile etkileşimimizi mümkün kılan görme sistemimizi ele alalım. Örneğin bildik olmayan bir sahnede bildik bir yüzün tanınması 100-200 ms' yede mümkün olurken, çok basit bir komplekslikteki bir işin çözümü geleneksel bilgisayarlarda günler olmasa da saatlerce zaman almaktadır.
- Diğer bir örnek bir yarasanın sonar sisteminin etkinliğidir. Bir yarasa sonarı hedefin uzaklığı, hedefin göreceli hızı ve büyüklüğü, hedefin değişik uzuvlarının büyüklüğü, hedefin gökyüzündeki açısı, yer yüzünden yüksekliği ve hedefin belirleyici özellikleri gibi birçok bilgiyi sağlar. Hedefin yerinin saptanmasındaki tahmin ve başarı oranı bile radar veya solar mühendislik hesaplarıyla yakalanamayacak imkansızlıktadır.

Yapay Sinir Ağları (Uygulama örnekleri):

- Otomatik pilot, uçak kontrol sistemleri, uçuş simulatörleri,
  - Otomatik araç yol rehberi,
  - Kredi başvuru değerlendirme, çek ve döküman okuma,
  - Kredi kart aktivite değişim (kayıp kart) kontrolü,
  - Hedef saptama, radar, sonar, görüntü sinyal işleme,
  - Makina görmesi, proses kontrol, nonlineer modelleme,
  - Animasyon, özel Efektler,
  - Üretim proses kontrolü, ürün optimizasyonu, görsel kalite kontrol,
  - EEG, ECG analizi, kanserli hücre analizi,
  - Robot yöre kontrolü, robot kol kontrolörleri,
  - Konuşma analizi, text konuşma dönüşümü,
  - Veri ve görüntü sıkıştırma, gerçek zamanlı tercüme,
- .....

Bulanık Mantık (Fuzzy Logic) ise, insan düşünmesini ve mantık yürütmesini modellemeye ve karşılaşılan problemlerde ihtiyaç doğrultusunda kullanmayı amaçlar. Bilgisayarlara, insanların özel verileri işleyebilme, deneyimlerinden ve öngörülerinden yararlanarak çalışabilme yeteneğini vermeye çalışır. İnsan mantık sistemini bilgisayarlara/makinelere uyarlanarak, bilgisayarların/makinelerin problem çözme kabiliyetini artırmaya çalışır. Bulanık mantık kullanan sistemler, insanın yapabildiği gibi, kesin olmayan dilsel bilgilere bağlı olarak etkin sonuç çıkarabilirler. Bulanık mantık, sistemin matematiksel modellenmesi ihtiyacını giderir. Bulanık mantık uygulama örnekleri;

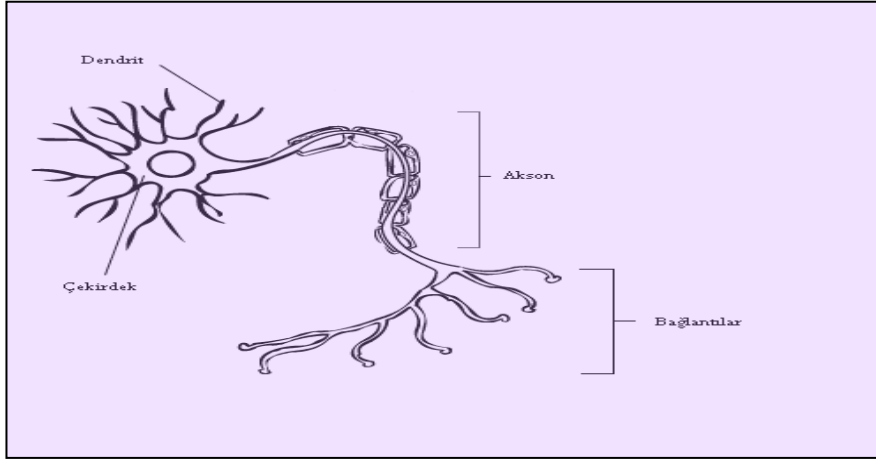
- Çamaşır makineleri, buzdolabı, kameralar, vb tüketici ürünleri.
  - Motor kontrolü, araç güç devri ve iletim kontrolü, batarya şarj cihazı, vb otomotiv ve enerji ürünleri.
  - Isı denetimi, kimyasal işlemler, çimento fırını arıtma ve damıtma sistemleri vb endüstriyel proses kontrol ürünleri.
  - Robotik ve üretim ile ilgili diğer konular.
  - Bulanık veri yapıları vb zeki bilgi sistemleri.
- .....

Yapay sinir ağlarını daha iyi anlayabilmek için, öncelikle insanın biyolojik sinir yapısını incelemeliyiz..

## 8.2 Yapay Sinir Ağlarının Mimarisi ve Yapı Elemanları

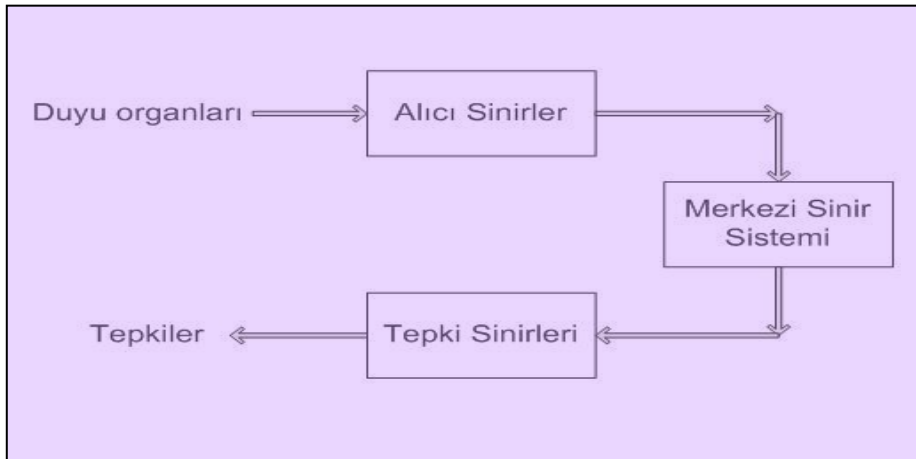
Yapay sinir ağı biyolojik sinir ağlarının modellenmesi olduğu için yapay sinir ağlarının çalışmasını anlayabilmek için öncelikle biyolojik sinir sisteminin yapısına bakmak gerekir. Biyolojik sinir sisteminin yapı taşı olan sinir hücreleri nöronlar, yapay sinir ağlarının da yapı taşıdır.

Biyolojik sinir sisteminin temel yapı taşı olan nöronların yapısı dört ana bölümden oluşmaktadır; dendrit, akson, çekirdek ve bağlantılar. Dendritlerin sinir hücresinin ucunda bulunan ve ağaç kökü görünümüne sahip bir yapıya sahiptir. Dendritlerin görevi bağlı olduğu diğer nöronlardan veya duyu organlarından gelen sinyalleri çekirdeğe iletmektir. Çekirdek dendrit tarafından gelen sinyalleri bir araya toplayarak ve aksona iletir. Toplanan bu sinyaller akson tarafından işlenerek nöronun diğer ucunda bulunan bağlantılara gönderilir. Bağlantılar ise yeni üretilen sinyalleri diğer nöronlara iletir.



Şekil 8.1 Biyolojik sinir ağının yapısı

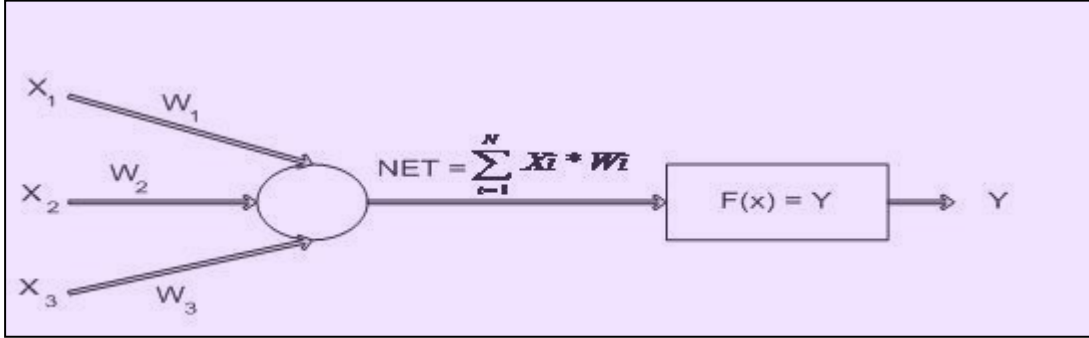
Bir insanın beyinde yaklaşık olarak 10 milyar sinir hücresi ve bu nöronların birbirleriyle yaptığı bağlantı sayısının ise 60 trilyon olduğu tahmin edilmektedir. Bu sinirler girdi bilgilerini duyu organlarından alırlar. Daha sonra alıcı (taşıyıcı) sinirler bu sinyalleri işleyip bir sonraki sinire aktararak sinyalin merkezi sinir sistemine kadar ulaşmasını sağlar. Merkezi sinir sistemi bu sinyalleri alıp yorumladıktan sonra tepki sinyallerini üretir. Bu sinyaller de tepkilerin oluşacağı organlara tepki sinirleri vasıtasıyla iletilir. Bu sayede duyu organlarından gelen bilgilere karşı tepki organlarına uygun işaretler sinir sistemi vasıtasıyla yollar.



Şekil 8.2 Merkezi sinir sisteminin Görevi

Bir yapay sinir hücresi beş bölümden oluşmaktadır;

- Girdiler
- Ağırlıklar
- Birleştirme fonksiyonu
- Aktivasyon fonksiyonu
- Çıktılar



**Şekil 8.3 Yapay Sinir Ağı**

**Girdiler;** Girdiler nöronlara gelen verilerdir. Girdiler yapay sinir hücresine bir diğer hücreden gelebileceği gibi direk olarak dış dünyadan da gelebilir. Bu girdilerden gelen veriler biyolojik sinir hücrelerinde olduğu gibi toplanmak üzere nöron çekirdeğine gönderilir.

**Ağırlıklar;** Yapay sinir hücresine gelen bilgiler girdiler üzerinden çekirdeğe ulaşmadan önce geldikleri bağlantıların ağırlığıyla çarpılarak çekirdeğe iletilir. Bu sayede girdilerin üretilen çıktı üzerindeki etkisi ayarlanabilmektedir. Bu ağırlıkların değerleri pozitif, negatif veya sıfır olabilir. Ağırlığı sıfır olan girdilerin çıkıl üzerinde herhangi bir etkisi olmamaktadır.

**Birleştirme Fonksiyonu;** Birleştirme fonksiyonu bir yapay sinir hücresine ağırlıklarla çarpılarak gelen girdileri toplayarak o hücrenin net girdisini hesaplayan bir fonksiyondur.

**Aktivasyon Fonksiyonu;** Birleştirme (toplama ) fonksiyonundan çıkan NET toplam hücrenin çıktısını oluşturmak üzere aktivasyon fonksiyonuna iletilir. Aktivasyon fonksiyonu genellikle doğrusal olmayan bir fonksiyon seçilir. Yapay sinir ağlarının bir özelliği olan “doğrusal olmama” aktivasyon fonksiyonlarının doğrusal olmama özelliğinden gelmektedir.

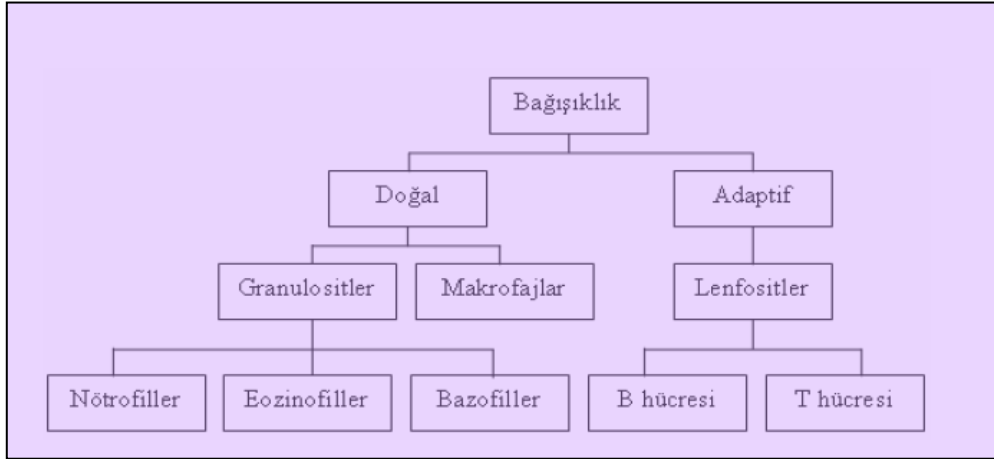
Aktivasyon fonksiyonu seçilirken dikkat edilmesi gereken bir diğer nokta ise fonksiyonun türevinin kolay hesaplanabilir olmasıdır. Geri beslemeli ağlarda aktivasyon fonksiyonunun türevi de kullanıldığı için hesaplamaların yavaşlamaması için türevi kolay hesaplanır bir fonksiyon seçilir.

### 8.3 Yapay Bağışıklık Sistemi

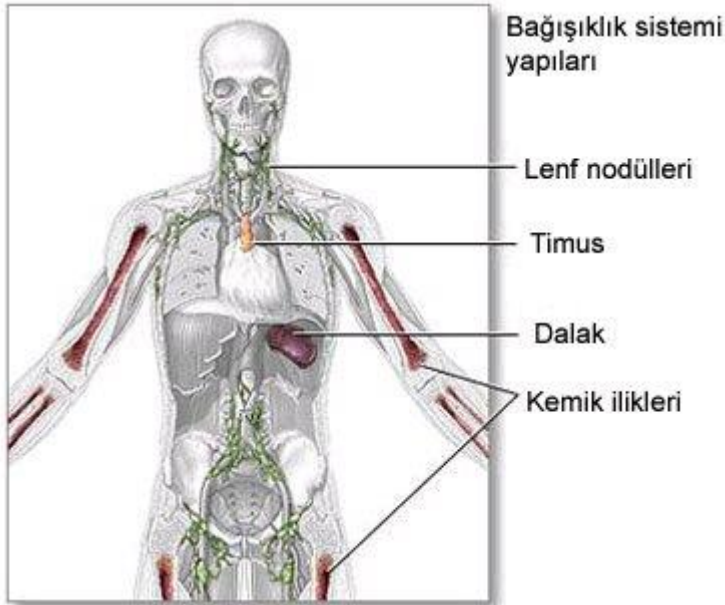
Vücudumuzun içinde bağışıklık sistemi adı verilen şaşırtıcı ve bir o kadar da ilginç savunma mekanizması vardır. Bağışıklık sistemi insanoğlunu "mikrop" diye tanımlanan, enfeksiyona yol açabilen virus, bakteri, mantar ve parazit gibi mikrororganizmaların zarar verici etkilerine karşı korur.

İnsan vücudu çevresinde bulunan çok sayıdaki mikrobun saldırısına uğrar ve bu organizmalar vücudumuza girebilmek için uğraş verir. Sağlıklı bir vücut; karşılaştığı hastalık etkenleriyle ve yabancı maddelerle çoğunlukla "çaktırmadan" başeder. Mikroplarla başedemediğimiz durumlarda da "hasta" oluruz.

Bağışıklık sisteminin görevi de; öncelikle bu organizmaların vücuda girmelerini engellemek veya girer ise vücuda girdikleri yerde yutmak, yayılmalarını engellemek ya da geciktirmektir. Bağışıklık sistemi bu görevlerini, yaşam süresi boyunca sürdürür ancak bazı koşullarda bağışıklık sistemi zayıflar.



Şekil 8.4 Bağışıklık Sistemi



Şekil 8.5 Bağışıklık Sistemi Yapıları

Yapay bağışıklık sistemleri, doğal bağışıklık sisteminden esinlenerek, doğal sistemin prensiplerinin mühendislik açısından incelenerek karmaşık problemlerin çözümünde ve sistemlerin gelişiminde kullanılan sezgisel yöntemlerden biridir. Bu konuda Farmer (1986) tarafından, bağışıklık sistemini diğer yapay zeka yaklaşımları ile ilişkilendirme üzerine yapılan çalışma öncü çalışma olarak kabul edilebilir. Bu alanda öncülük yapmış çalışmalar arasında Ishida (1996); Hunt & Cook (1996); Dasgupta (1996) ve Hofmeyr & Forrest (1999) tarafından yapılmış çalışmalar da sayılabilir.

Yapay bağışıklık sisteminde prensip B hücrelerinin temsilidir. Bağışıklık sisteminin örüntü tanıma bakış açısından en önemli karakteristiği B ve T hücrelerinin antijenleri tanıyabilme yeteneğine sahip yüzey

alıcı molekülleri taşımasıdır. B hücresi örüntü tanıma ve makine öğrenmesi için temel kavramdır. Özellikle başlıca tanıma mekanizmaları sık sık antikor-antijen bağlanımını modellemektedir. Uyarılmış bir B hücresinin özellik vektörü ve bir eğitim verisinin özellik vektörü arasındaki *Öklid uzaklığı*ni kullanarak antikorlar ve antijenler arasındaki bağlanmanın derecesine karar verilen çalışmalar söz konusudur (Şekil 8.6).

YBS, yapay sinir ağları ve genetik algoritmalarla olduğu gibi doğadan taklit edilerek ortaya çıkarılmış genel amaçlı bir sezgisel yöntemdir. Canlılardaki savunma mekanizması özetlenip modellenerek oluşturulmuş bu optimizasyon algoritmasının, mühendislik alanındaki bir çok kompleks problemin çözümünde nasıl performans göstereceği incelenmiştir. Yapay sinir ağları ve genetik algoritmalar kadar eski olmayan bu yöntemin popülerliği gittikçe artmaktadır. Çünkü her iki yönetime göre de avantajları olduğu yapılan uygulamalarda gözlemlenmiştir. Ancak görünen o ki; uygulama bazlı seçim yapmak ve bunun üzerine artıları ve eksileri göz önüne almak daha sağlıklıdır. YBS, genetik algoritmalar ile benzerlikler göstermektedir. Çözüm uzayının kodlanması, uygunluk fonksiyonunun belirlenmesi, mutasyon vb. ortak operatörler ve süreçler her iki yöntemde de ortaktır. Çaprazlama operatörü YBS’de kullanılmamaktadır. YBS, lokal optimumlara daha hızlı yönelmekte ancak bir genetik algoritma kadar global optimuma yakınsamayabilmektedir. YBS ise doğası gereği örüntü tanıma, bilgisayar güvenliği, ağ güvenliği ve dinamik is-programlama vb. alanlarda genetik algoritma ve diğer optimizasyon tekniklerine göre daha iyi sonuçlar vermiştir. YBS, yapay sinir ağlarında olduğu gibi hafıza ve öğrenme özelliklerine sahiptir. Her iki sistemde içerisinde çok sayıda ve farklı hücrelerden oluşmaktadır. Her iki sistem için de belirleyici bir eşik değeri vardır ve bu değer kesin tanıma sağlar. Ancak yapıları incelendiğinde farklı alanlarda benzer performanslar için üretildikleri göze çarpmaktadır<sup>1</sup>.



**Şekil 8.6 Bağışıklık Sisteminin İşleyişi**

Yapay bağışıklık sistemi kapsamında çeşitli algoritmalar geliştirilmiştir:

- *Klonal seçim algoritması* antijenik uyarıcı için bağışıklık tepkisinin temel özelliklerini açıklamak için bağışıklık sistemini kullanan bir algoritmadır.

<sup>1</sup> Yapay Bağışıklık Sisteminde Bulanık Yaklaşımcı kural Tabanı Optimizasyonu, Erkan DUMAN Erhan AKIN

- *Klonsal Seçim Algoritması* iki temel esas üzerine şekillendirilmiştir: Birincisi sadece antijeni tanıyan hücreler çoğalma için seçilirler, ikincisi seçilen ve çoğalan hücreler duyarlılık olgunlaşması işlemine tabi tutularak, antijene olan duyarlılıkları artırılır.
- De Castro ve Von Zuben duyarlılık olgunlaşmasındaki işlemleri baz alarak *Klonsal Seçim Algoritmasını* ortaya atmış ve oluşturdukları algoritmayı karakter tanıma, optimizasyon gibi problemlere uygulayarak performansını analiz etmişlerdir<sup>2</sup>.

Yapay Bağışıklık Sistemi aşağıdaki konu başlıklarında uygulama alanı bulmuştur<sup>3</sup>:

- Robotics
- Control
- Optimization:
- Neural Network yaklaşımlar
- Anomali Saptama
- Agent-Based yaklaşımlar
- Öğrenme
- Örüntü tanıma
- Diğer uygulamalar
  - Açık WebServer kordinasyonlar
  - Çizelgeleme
  - Veri madenciliği
  - Sınıflama

Yapılan bir çalışmada beyin MRI sınıflandırmak için yapay bağışıklık sistemlerinin klonsal seçim algoritması kullanılmış ve çok sayıda tespit için sürede azalma sağlanmış, böylece doktorlara daha verimli ve doğru bir teşhis imkanı verilmiştir<sup>4</sup>.

Yapay Bağışıklık Tanıma Sistemi kullanılarak başka bir çalışmada Karotis arter Doppler sinyallerinin bulanık görünümündeki tanı hatalarını giderebilecek bir sistem geliştirilmiştir<sup>5</sup>.

Bir başka çalışmada ise mikroorganizmaların sınıflandırılmasında faydalanılmıştır<sup>6</sup>.

Statlog kalp hastalığı ve Pima Hint Diyabeti konusunda yapılan sınıflandırmada %80 lerde başarı elde edildi<sup>7</sup>.

Bir diğer çalışmada, yapay bağışıklık sisteminin kablosuz ağlarda farklı bir uygulaması üzerinde durulmuştur. Yapılan uygulamalardan farklı olarak kullanıcı tabanlı yerine ağ tabanlı YBS uygulanmıştır. Ağ tabanlı ve kullanıcı tabanlı YBS uygulanmasının bilgisayarları ve ağı meşgul etme oranlarına bakılmıştır. Saldırı tespitinin önemli olmasının yanında bilgisayarların ve ağda performanslı bir şekilde çalışması önemli bir kriterdir. YBS nin uygulanma farklılığına göre ağdaki ve bilgisayarlardaki performansı değiştirdiği görülmüştür. Kullanıcı tabanlı YBS de kullanıcıların

<sup>2</sup> (De Castro L.N., Von Zuben F.J., 2000a, "The Clonal Selection Algorithm with Engineering Applications", In The Proceedings of Genetic and Evolutionary Computation Conference (GECCO), Workshop on Artificial Immune Systems and Their Applications, Las Vegas, USA, 36-37.)

<sup>3</sup> (Leandro Nunes de Castro, (VI-Brazilian Symposium on Neural Networks, SBRN 2000)

<sup>4</sup> EURASIP Journal on Advances in Signal Processing Volume 2008 (2008), Article ID 547684, 8 pages doi:10.1155/2008/547684

<sup>5</sup> Medical application of Artificial Immune Recognition System (AIRS): Diagnosis of atherosclerosis from carotid artery Doppler signals, Computers in Biology and Medicine 37 (2007) 1092 – 1099

<sup>6</sup> Medical application of information gain-based artificial immune recognition system (IG-AIRS): Classification of microorganism species, Expert Systems with Applications 36 (2009) 5168–5172

<sup>7</sup> The Medical Applications of Attribute Weighted Artificial Immune System (AWAIS): Diagnosis of Heart and Diabetes Diseases



bilgisayarlarında önemli oranda performans düşüklüğü gözlenmiştir. Bunun yanında ağ performans düşüş oranları izlenmiştir. Ağ tabanlı YBS uygulamasında ise kullanıcı bilgisayarlarında çok büyük performans kaybı olmadığı görülmüştür. Ağ tabanlı ve kullanıcı tabanlı YBS nin ağ performans kaybı oranlarına bakıldığında aralarında çok büyük bir farkın olmadığı görülmüştür. Bunun sebebi olarak ağ üzerindeki paket trafiğinde herhangi bir değişikliğin olmamasıdır. Ağ tabanlı ve kullanıcı tabanlı uygulamanın her ikisinde de ağ üzerindeki paketler tüm kullanıcılara ve erişim noktasına ulaşmaktadır. Bu sebeple erişim noktası üzerinden YBS uygulanması daha performanslı olduğu sonucu ortaya çıkmıştır<sup>8</sup>.

---

<sup>8</sup> Kablosuz Yerel Alan Ağlarında Yapay Bağışıklık Sistemi ile Saldırı Tespiti ve Performans Analizi, Erhan Akbal, Burhan Ergen