**YAPAY SİNİR AĞLARI**

Bilgisayar teknolojilerinin gelişmesiyle birlikte insanoğlu neredeyse tüm işlemlerini bu yenilikçi teknolojiler üzerinde yapmakta ve yeni yöntemlerin bulunmasına olanak sağlamaktadır. İşte bu yüzdendir ki, 1980’li yıllarda ortaya atılan makinanın insan gibi düşünebilmesi fikri ortaya atılmış, 1990’lı yıllara gelindiğinde ise Yapay Sinir Ağları teknolojisi iyiden iyiye hızlanmış ve büyük bir gelişme görülmüştür.

Yapay Sinir Ağları, Yapay Zekâ kavramı altında oluşan bir alt başlık olup, bu konuya ilgi duyan araştırmacıların odak noktası haline gelmiştir.

Bu bölümde Yapay Sinir Ağları teorik açıdan anlatılmış olup, alt başlıklar halinde uygulamaları yer almıştır.

**Yapay Sinir Ağları Tanımı ve Tarihçesi**

Yapay sinir ağları örneklerle ilgili bilgiler toplamakta, genellemeler yapmakta ve daha sonra hiç görmediği örnekler ile karşılaştırılınca öğrendiği bilgileri kullanarak o örnekler hakkında karar verebilmektedir. Yapay sinir ağları bu öğrenebilme ve genelleme özellikleri nedeniyle günümüzde birçok bilim alanında geniş uygulama olanağı bulmakta ve karmaşık problemleri başarı ile çözebilme yeteneğini ortaya koymaktadır (Ergezer vd, 2003).

Diğer bir tanıma göre yapay sinir ağları; insan beyninden esinlenerek geliştirilmiş, ağırlıklı bağlantılar aracılığıyla birbirine bağlanan ve her biri kendi belleğine sahip işlem elemanlarından oluşan paralel ve dağıtılmış bilgi isleme yapıları; bir başka deyişle, biyolojik sinir ağlarını taklit eden bilgisayar programlarıdır (Elmas, 2003).

Yapay sinir ağlarının dayandığı ilk hesaplama modelinin temelleri 1940'ların başında araştırmalarına başlayan W.S. McCulloch ve W.A. Pitts'in, 1943 yılında yayınladıkları bir makaleyle atılmış olmuştur. Daha sonra 1954 yılında B.G. Farley ve W.A. Clark tarafından bir ağ içerisinde uyarılara tepki veren, uyarılara adapte olabilen model oluşturulmuştur. 1960 yılı ise ilk neural bilgisayarın ortaya çıkış yılıdır. 1963 yılında basit modellerin ilk eksiklikleri fark edilmiş, ancak başarılı sonuçların alınması 1970 ve 1980'lerde termodinamikteki teorik yapıların doğrusal olmayan ağların geliştirilmesinde kullanılmasına kadar gecikmiştir. 1985 yapay sinir ağlarının oldukça tanındığı, yoğun araştırmaların başladığı yıl olmuştur (Mehra Pankaj Wah W Benjamin, 1992).

Yapay sinir ağlarının kullanım alanları: Kontrol ve sistem tanımlama, görüntü ve ses tanıma, tahmin ve kestirim, arıza analizi, tıp, haberleşme, trafik, üretim yönetimi olarak sayılabilir(Pirim, 2006). Çizelge 1’de geleneksel algoritmalar ile YSA’ların karşılaştırma tablosu görülmektedir.

Çizelge 1 Geleneksel Algoritmalar ile YSA'ların Karşılaştırılması (Pirim, 2006)

|  |  |
| --- | --- |
| **Geleneksel Algoritmalar** | **Yapay Sinir Ağları** |
| Çıkışlar, koyulan kurallara girişlerin uygulanması ile elde edilir.  | Öğrenme esnasında giriş çıkış bilgileri verilerek, kurallar koyulur.  |
| Bilgiler ve algoritmalar kesindir.  | Deneyimden yararlanır.  |
| Hesaplama; merkezi, eş zamanlı ve ardışıktır.  | Hesaplama; toplu, eş zamansız ve öğrenmeden sonra paraleldir.  |
| Bellek paketlenmiş ve hazır bilgi depolanmıştır.  | Bellek ayrılmış ve ağa yayılmıştır.  |
| Hata toleransı yoktur.  | Hata toleransı vardır.  |
| Nispeten hızlıdır.  | Yavaş ve donanıma bağımlıdır.  |

**Biyolojik Sinir Hücresi**

Biyolojik sinir ağlarının temel elemanları, biyolojik sinir hücreleridir. İnsan beyninin korteks kısmında yer alan sinir hücresi sayısı yaklaşık olarak 1011 olup her hücre sayısı, 1000-10000 arasında değişen başka hücrelerle karşılıklı ilişki içerisindedir. Şekil 1.’de görüldüğü gibi, bir sinir hücresinin temel elemanları hücre gövdesi, dendrit ve akson’dur.



Şekil 1. Sinir Hücresi (Elmas, 2003)

Sinir hücresine diğer sinir hücrelerinden gelen uyarımlar, dendritler aracılığıyla hücre gövdesine taşınır ve hücre içi aktivasyonun/kararlılık halinin bozulmasıyla oluşan bir kimyasal süreç içerisinde diğer hücrelere aksonlarla iletilir; uyarımların diğer sinir hücrelerine taşınabilmesinde akson uçları ile dendritler arasındaki sinaptik boşluklar (sinaps) rol oynar. Sinaptik boşluk içinde yer alan “sinaptik kesecikler”, gelen uyarımların diğer hücrelere dendritler aracılığıyla geçmesini koşullayan elemanlardır. Sinaptik boşluğa, “sinaptik kesecikler” tarafından sağlanan nöro-iletken maddenin dolması uyarımların diğer hücrelere geçişini koşullar. Hücrelere gelen uyarımlarla uyumlu olarak hücreler arasındaki mevcut sinaptik ilişkilerin değişimi veya hücreler arasında yeni sinaptik ilişkilerin kurulması “öğrenme” sürecine karşılık gelir (Koç vd, 2004).

Çizelge 2 Biyolojik Sinir Yapısı İle Yapay Sinir Ağı Karşılıkları (Koç, 2004)

|  |  |
| --- | --- |
| **Biyolojik Sinir Sistemi** | **Yapay Sinir Sistemi** |
| Nöron | İşlemci eleman |
| Dentrit | Toplama fonksiyonu |
| Hücre gövdesi | Transfer fonksiyonu |
| Aksonlar | Yapay nöron çıkışı |
| Sinapslar | Ağırlıklar |

**Yapay Sinir Yapısı**

Biyolojik sinir ağlarının sinir hücresi olduğu gibi yapay sinir ağlarının da yapay sinir hücreleri vardır. Yapay sinir hücreleri mühendislik biliminde proses elemanları olarak da adlandırılmaktadır. Öztemel (2006) tüm bu proses elemanlarını (Şekil 2) ki bunların da 5 temel elemanı vardır :



Şekil 2. Yapay nöron yapısı (Öztemel, 2006)

**Girdiler:** Bir yapay hücresine dış dünyadan gelen bilgilerdir. Bunlar ağın öğrenmesini istenen örnekler tarafından belirlenir.

**Ağırlıklar:** Bir yapay hücreye gelen bilginin önemini ve hücre üzerindeki etkisini gösterir. Şekildeki ağırlık wi1, xi1 girdisinin hücre üzerindeki etkisini göstermektedir. Ağırlıkların büyük ya da küçük olması önemli veya önemsiz olduğu anlamına gelmez. Bir ağırlığın sıfır olması o ağ için en önemli olay olabilir.

**Toplama Fonksiyonu (Birleştirme Fonksiyonu):** Bu fonksiyon, bir hücreye gelen net girdiyi hesaplar. Bunun için değişik fonksiyonlar kullanılmaktadır. En yaygın olanı ise ağırlıklı toplamıdır. Burada her gelen bilgi kendi ağırlığı ile çarpılarak toplanır. Böylece ağa gelen net girdi bulunmuş olur. Aşağıdaki tabloda bazı toplama fonksiyonlarına yer verilmiştir.



Şekil 3 Bazı Toplama Fonksiyonları (Çayıroğlu, 2015)

**Aktivasyon Fonksiyonu:** Bu fonksiyon hücreye gelen net girdiyi işleyerek hücrenin bu girdiye karşılık üreteceği çıktıyı belirler. 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 hesaplamanın yavaşlamaması için türevi kolay hesaplanır bir fonksiyon seçilir. Şekil 4‘te yapay sinir ağlarında kullanılan aktivasyon fonksiyonları listelenmiştir. Günümüzde en yaygın olarak kullanılan “Çok katmanlı algılayıcı” modelinde genel olarak aktivasyon fonksiyonu olarak “Sigmoid fonksiyonu” kullanılır.



Şekil 4 Aktivasyon Fonksiyonları (Çayıroğlu, 2015)

**Hücrenin Çıktısı:** Aktivasyon fonksiyonu tarafından belirlenen çıktı değeridir. Üretilen çıktı dış dünyaya veya başka bir hücreye gönderilir. Hücre kendi çıktısını kendisine girdi olarak da gönderebilir. Bir proses elemanının birden fazla çıktısı olmasına rağmen sadece bir çıktısı olabilir (Öztemel, 2006).

Temel hesaplama elemanı (model nöron) genellikle ‘*düğüm’* veya ‘*birim’* olarak tanımlanır. Bu harici bir kaynaktan veya diğer birimlerden girdi/giriş (input) alır. Her bir girdi ilişkili bir ağırlığa, *w,* sahip olup bu sinaptik öğrenme modeli şekline getirilebilir. Ağ girdisi, ağırlıklandırılmış girdilerin toplamı bazı fonksiyonları hesaplar:

|  |  |
| --- | --- |
| $$y\_{i}=f\left(\sum\_{j}^{}w\_{ij}y\_{j}\right)$$ | **(**1) |

Bu çıkış, sırayla, diğer birimlere giriş olarak hizmet edebilir. Ağırlıklandırılmış toplam $\sum\_{j}^{}w\_{ij}y\_{j}$ *i* birimine *’net girdi’* olarak tanımlanır, genellikle *neti* olarak yazılır*.* Ayrıca w*ij* değeri *j* girdisinden *i* girdisine ağırlığı ifade etmektedir. ‘*f’* fonksiyonuna girdinin ‘*aktivasyon fonksiyonu*’ denir. En basit anlamda, *f* birim fonksiyonudur ve birimin çıkışı onun net girdisidir. Buna *doğrusal (lineer) birim* denir.

Yapay sinir ağları ile ilgili geliştirilen modeller; tek katmanlı algılayıcılar (TKA), perceptron ve ADALINE/ MADALINE’dir.

**Tek Katmanlı Algılayıcılar**

Tek katmanlı yapay sinir ağları sadece girdi ve çıktı (Ç) katmanlarından oluşur. Çıktı üniteleri bütün girdi ünitelerine (X) bağlanmaktadır ve her bağlantının bir ağırlığı (W) vardır. İki girdi ve bir çıktıdan oluşan tek katmanlı bir yapay sinir ağı Şekil 5 te verilmiştir.

****

Şekil 5 Tek Katmanlı Algılayıcı Modeli (Öztemel, 2006)

Bu ağlarda süreç elemanlarının değerlerinin ve dolayısıyla ağın çıktısının sıfır olmasını önleyen bir de eşik değeri (Φ) vardır ve değeri daima 1’dir. Ağın çıktısı ağırlıklandırılmış girdi değerlerinin eşik değeri ile toplanması sonucu bulunur. Bu girdi ile bir aktivasyon fonksiyonundan geçirilerek ağın çıktısı hesaplanır. Tek katmanlı algılayıcılarda çıktı fonksiyonu doğrusal bir fonksiyondur ve 1 veya -1 değerlerini almaktadır. Eğer çıktı 1 ise birinci sınıfa -1 ise ikinci sınıfa kabul edilmektedir (Öztemel, 2006).

**Basit Algılayıcılar**

Basit algılayıcılar modeli 1950’li yılların sonlarında Rosenblatt tarafından ortaya atılmıştır. Esas olarak, basit algılayıcılar sınıflandırma amacıyla kullanılan, McCulloch ve Pitts tarafından geliştirilmiş daha karmaşık bir modeldir. Bir araya gelmiş sinir hücrelerinin miktarına bağlı olarak, basit algılayıcılar ile değişik sayıda sınıflandırma problemi çözebilir. Doğru bir sınıflandırma için sınıfların düzlemsel olarak ayrılması gerektiği gösterilmiştir (Nygren, 2004, Minsky ve Papert,1969).

Basit algılayıcılar bir sinir hücresinin birden fazla girdiyi alarak bir çıktı üretmesi prensibine dayanmaktadır. Ağın çıktısı bir veya sıfırdan oluşan mantıksal değerdir. Çıktının hesaplanmasında eşik değer fonksiyonu kullanılır. Ağın yapısı Şekil 6 da verilmektedir.

****

Şekil 6 Basit Algılayıcı Model (Elmas, 2003)

**Adaline**

1959’da, Stanford üniversitesinden Bernard Widrow, basit nöron benzeri elemanlara dayanan ve “Adaline” (Adaptive Linear Nöron) olarak adlandırılan bir adaptif lineer elemanı geliştirmiştir. Adaline yapısı tüm sinir ağlarının en basitidir ve öğrenme için danışmanlı öğrenmeyi kullanır. Adaline ve iki tabakalı biçimi olan “madaline” (Multiple Adaline); ses tanıma, karakter tanıma, hava tahmini ve adaptif kontrol gibi çok çeşitli uygulamalar için kullanılmıştır. Daha sonraları adaline, ayrık bir çıkı\_ yerine sürekli bir çıkış üretmek için geliştirilmiştir. Widrow, telefon hatları üzerindeki ekoları elimine etmeye yarayan adaptif filtreleri geliştirmede, adaptif lineer eleman algoritmasını kullanmıştır. Bununla ilk defa YSA’lar gerçek bir probleme uygulanmıştır. Adaline birçok uygulama için oldukça iyi çalışmasına rağmen lineer problem uzayıyla sınırlıdır. Lineer transfer fonksiyonu kullanırlar. Giriş ve istenilen çıkış desenlerinin tekrar tekrar ağa uygulanmasıyla eğitim gerçekleştirilir. Desenlerin doğru sınıflara ayrılmasıyla, hatalar minimize edilerek öğrenme gerçekleştirilir. Eğitimden sonra adaline, yeni girişleri kazandığı deneyime göre sınıflandırabilir (Kıraç, 2011).

**Çok Katlı Algılayıcılar (ÇKA)**

Algılayıcı ve Adaline yöntemleri doğrusal olmayan çözümler üretemediği için hem mimari hem de eğitim algoritması açısından iyileştirilmiş Çok Katmanlı Algılayıcı (MLP) ağı önerilmiştir. Mimari açıdan doğrusal olmayan aktivasyon fonksiyonuna sahip birçok nöronun birbirine hiyerarşik olarak bağlandığı bir yapıya sahip olan MLP, Algılayıcı ve Adaline yöntemlerinin avantajları yanı sıra geri-yayılım adındaki öğrenme sistemini kullanmaktadır ve genel olarak yapay sinir ağları ileri beslemeli ve geri beslemeli ağlar olarak ikiye ayrılmaktadır.

**İleri Beslemeli Ağlar**

İleri beslemeli sinir ağları tek yönlü sinyal akışı için izin verir. Ayrıca, ileri beslemeli sinir ağları çoğu katmanlar halinde organize edilmektedir. Üç katmanlı ileri beslemeli sinir ağı bir örneği Şekil 7'de gösterilmiştir. Bu şebeke giriş düğümleri, iki gizli katman ve bir çıkış katmandan oluşur. Tipik aktivasyon fonksiyonları Şekil 4’te gösterilmiştir. Bu sürekli aktivasyon fonksiyonları katmanlı ağlar gradyan tabanlı eğitimi için izin verir (Wilamowski, 2003).

****

Şekil 7 İleri Beslemeli Sinir Ağlara Örnek (Öztemel, 2006)

İleri beslemeli yapay sinir ağında, hücreler katmanlar şeklinde düzenlenir ve bir katmandaki hücrelerin çıkışları bir sonraki katmana ağırlıklar üzerinden giriş olarak verilir. Giriş katmanı, dış ortamlardan aldığı bilgileri hiçbir değişikliğe uğratmadan ara (gizli) katmandaki hücrelere iletir. Bilgi, ara ve çıkış katmanında işlenerek ağ çıkışı belirlenir. Bu yapısı ile ileri beslemeli ağlar, doğrusal olmayan statik bir işlevi gerçekleştirir. İleri beslemeli 3 katmanlı yapay sinir ağının, orta katmanında yeterli sayıda hücre olmak kaydıyla, herhangi bir sürekli fonksiyonu istenilen doğrulukta yaklaştırabileceği gösterilmiştir. En çok bilinen algoritma olan geriye yayılım öğrenme algoritması, bu tip yapay sinir ağların eğitiminde etkin olarak kullanılmaktadır. Ağa, hem örnekler hem de örneklerden elde edilmesi gereken çıktılar (beklenen çıktılar) verilmektedir. Ağ kendisine gösterilen örneklerden genellemeler yaparak problem uzayını temsil eden bir çözüm uzayı üretmektedir. Daha sonra gösterilen benzer örnekler için bu çözüm uzayı sonuçlar ve çözümler üretebilmektedir (Öztemel, 2006).

**Geri Beslemeli Yapay Sinir Ağları**

Geri beslemeli Yapay Sinir Ağları (YSA)’ da, en az bir hücrenin çıkışı kendisine ya da diğer hücrelere giriş olarak verilir ve genellikle geri besleme bir geciktirme elemanı üzerinden yapılır. Geri besleme, bir katmandaki hücreler arasında olduğu gibi katmanlar arasındaki hücreler arasında da olabilir. Bu yapısı ile geri beslemeli YSA, doğrusal olmayan dinamik bir davranış gösterir. Dolayısıyla, geri beslemenin yapılış şekline göre farklı yapıda ve davranışta geri beslemeli YSA yapıları elde edilebilir. Aşağıda bulunan Şekil 8’de iki katmanlı ve çıkışlarından giriş katmanına geri beslemeli bir YSA yapısı görülmektedir (Kabalcı, 2015).



Şekil 8 Geri Beslemeli YSA (Kabalcı, 2015)

**Geri Yayılım Algoritması (BackPropagation)**

Geri yayılım algoritması, sinir ağının denetimli sınıfına giren genel bir algoritmadır. Daha öncede belirtildiği gibi girişlerle çıkışlar arasındaki hata sinyali bulunarak, ağırlıklar bu hata sinyaliyle güncellenmektedir. Hata yani e(t), arzu edilen çıkış t(t) ile sinir ağının çıkışı y(t) arasındaki farktır.

$e(t) = t(t) – y(k) ; t=1,..,m $ (2)

Geri yayılım öğrenme algoritmasının temel yapısı, zincir kuralı kullanılarak, ağ üzerindeki tüm ağırlıklara E hata fonksiyonunun etkilerini yaymaktır. Böylelikle toplam hata değerini minimize etmektedir.

|  |  |
| --- | --- |
| $$E\_{Top}= \lim\_{t\to \infty }\left(\frac{\sum\_{t=1}^{t}E^{t}}{t}\right)$$ | (3) |

Herhangi bir ‘t’ denemesinde Et değeri küçültülebilirse sistemin hatasının azalacağı Eşitlik 3’te kolaylıkla gözlenmektedir. Sistem hatasındaki azalmayı temel olarak alarak ve bu azalmayı destekleyecek şekilde ağ üzerindeki ağırlık değerleri yeniden belirlenecek yöntemler kullanılmaktadır. Eğitme işlemi ve eğitimden sonraki test işlemi bu akışa göre yapılır. Bu algoritma ile xi. giriş için, *i* ve *j* kat işlem elemanları arasındaki ağırlıklardaki *wji*(t) değişikliği hesaplanır. Bu ifade,

 (4)

olarak verilir. Eşitlik 4’de *η* öğrenme katsayısı, α momentum katsayısı ve *δj* ara veya çıkış katındaki herhangi bir *j* nöronuna ait bir faktördür. Çıkış katı için bu faktör aşağıdaki şekilde verilir.

 (5)

Burada yj(t) ise j işlemci elemanının hedef çıkışıdır. Ara katlardaki (İşlem Elemanları – Nöronlar ) İşlem Elemanları (İE) için ise bu faktör,

 (6)

olarak verilir. Ara katlardaki İE’ler için herhangi bir hedef çıkış olmadığından, Eşitlik 5 yerine Eşitlik 6 kullanılır. Bu duruma bağlı olarak çıkış katından başlayarak *δj* faktörü, bütün katlardaki İE’ler için hesaplanır. Daha sonra Eşitlik 4’deki formüle bağlı olarak, bütün bağlantılar için ağırlıkların güncelleştirilmesi gerçekleştirilir. Geri yayılım algoritmasında kullanılacak aktivasyon fonksiyonu birkaç önemli karakteristiğe sahip olmalıdır. Aktivasyon fonksiyonu, sürekli, türevi alınabilir ve tekdüze bir şekilde azalmayan bir fonksiyon olmalıdır. Bu fonksiyonun türevinin kolay alınıyor olması tercih sebebidir. Genellikle, fonksiyonun minumum ve maksimum asimtotlar arasında uzanması beklenir (Yüksek, 2007).



Şekil 9 Geri Yayılma Algoritması (Yüksek, 2007)



Şekil 10 İleri beslemeli çok katmanlı sinir ağı: x, z, y vektörleri ile V,

W matrisleri (Yüksek vd, 2007)