Yapay Sinir Ağları

Makine öğrenmesi sınıflandırma uygulamaları kapsamında incelenebilen yapay sinir ağları (YSA), insan beyninin özelliklerinden esinlenerek öğrenilen bilgiler ışığında yeni bilgiler oluşturabilme ve keşfetme yetenekleri sağlayarak bunları kendi kendine yapabilmesi amacıyla geliştirilen bilgisayar sistemleridir. YSA normal programlama tekniği ile çözülmesi imkansız olan problemleri çözmeyi mümkün hale getirir.

YSA'nın genel özellikleri şöyledir:

- YSA makine öğrenmesi gerçekleştirir.
- Bilgileri saklar ve örnekler ile öğrenebilir.
- Eğitimde verilen bilgiler sayesinde görülmemiş örnekler hakkında da bilgi üretebilir.

Yapay sinir hücresi

Yapay sinir ağları yapay sinir hücrelerinden oluşur ve bu hücrelere proses elemanı denir. Prosesler beş elemandan oluşur. Girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktısıdır.

Yapay sinir ağı yapısı

Yapay sinir hücreleri bir araya gelerek yapay sinir ağlarını oluştururlar. Ağ yapısı; girdi katmanı, ara katmanlar ve çıktı katmanından oluşur.

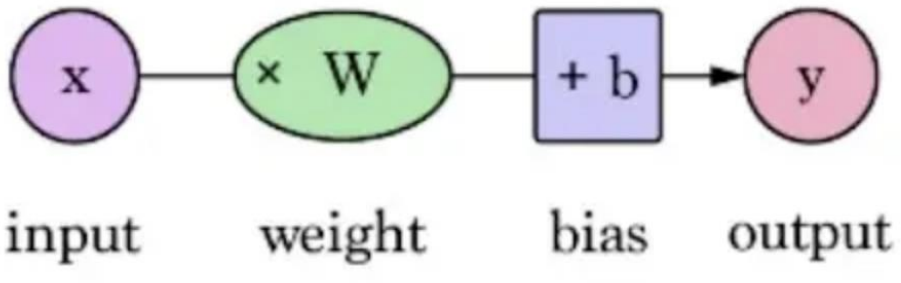
YSA da bir girdi ve bir çıktı seti bulunur. Eğitilmek istenen YSA ya öğretilecek bilgi bir vektör haline çevrilerek verilir. Doğru çıktıyı üretecek şekilde sistem parametreleri ayarlanır.

YSA da öğrenme 2 aşamada yapılır. İlk aşamada sistem eğitilirken üretilecek çıktı değerleri verilir. Çıktı değerinin doğruluğuna göre ağın bağlantılarının ağırlıkları değiştirilir. Ağın eğitimi bittikten sonra sistemin performansını ölçmek için test verisi sisteme sunulur. Test aşamasında ağırlık değerlerinde herhangi bir değişiklik yapılmaz eğitimde kullanılan ağırlıklar kullanılır ve bu karşılaşmadığı test verileri için çıktılar üretir. Elde edilen çıktıların doğruluk değerlerine göre sistemin performansı değerlendirilir.

Tek katmanlı algılayıcı (perceptron)

Yapay sinir ağlarıyla ilgili geliştirilen tek katmanlı yapılardan biri de perceptrondur (algılayıcı). Bu, şimdiye kadar oluşturulan ilk sinir ağıdır. Giriş katmanında 2 nörondan ve çıkış katmanında 1 nörondan oluşur. Bu yapı 2 grubu ayırt etmek için basit bir sınıflandırıcı oluşturmaya izin verir. Bir perceptron, ikili sınıflandırıcıların denetimli öğrenimi için bir algoritmadır. Bu algoritma, nöronların eğitim setindeki öğeleri birer birer öğrenmesini ve işlenmesini sağlar.

Tek katmanlı ve çok katmanlı olmak üzere 2 tür algılayıcı vardır. Tek katmanlı algılayıcılar yalnızca doğrusal olarak ayrılabilir desenleri öğrenebilir. Çok katmanlı algılayıcılar iki veya daha fazla katmana sahip ileri beslemeli sinir ağlarıdır ve daha fazla işlem gücüne sahiptir. Perceptron algoritması, doğrusal bir karar sınırı çizmek için giriş sinyallerinin ağırlıklarını öğrenir. Bu, doğrusal olarak ayrılabilir +1 ve -1 sınıfı arasında ayrım yapmanızı sağlar.



Şekilde görüldüğü üzere tek katmanlı algılayıcı yapısında girdi katmanı ve bu girdilerin ağırlıkları, belirlenmiş bir bias değeri ve bunların toplamı da çıktı (y) olarak hesaplanır. Burada y değeri x değerinin ne kadar doğru hesaplandığını gösterir. Çıktı skorunu iyileştirmek için ağırlık (w) ve bias (b) değerleri güncellenerek kullanılır. Derin öğrenme ya da yapay sinir ağlarındaki amaç en iyi skoru verecek olan ağırlık ve bias değerlerini hesaplamaktır.

Tek katmanlı algılayıcılar doğrusal olmayan problemlerin çözümünde yeterli gelmezler ve çok katmanlı algılayıcılar devreye girer.

Temel olarak perceptron yapısı giriş değerleri, giriş değerlerinin ağırlıkları, bias değerleri, net toplam ve aktivasyon fonksiyonlarından oluşmaktadır. Perceptron, tek katmanlı bir sinir ağıdır ve çok katmanlı bir algılayıcı da sinir ağları (Neural Network) olarak adlandırılır.

Perceptron'un çalışma mantığı şu şekildedir.

- Tüm x girdileri, ağırlıkları w ile çarpılır.
- Tüm çarpılan değerleri toplanır ve bunlara ağırlıklı toplam denir.
- Bu ağırlıklı toplama doğru etkinleştirme fonksiyonu uygulanır.

Ağırlıklar, belirli düğümün gücünü gösterir.

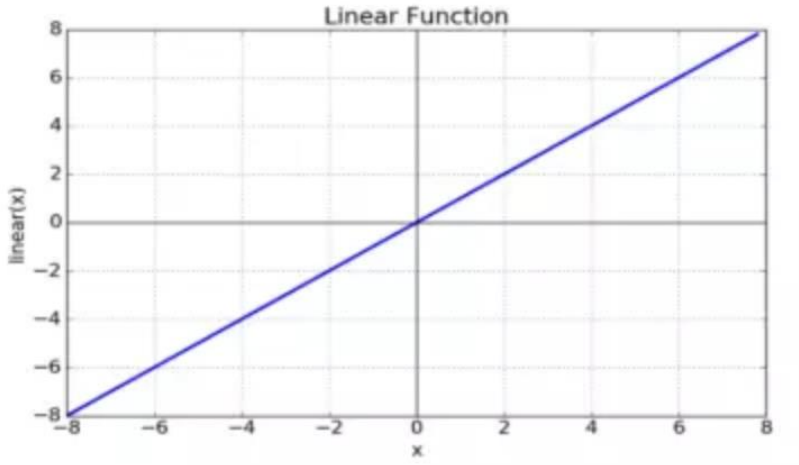
Sapma değeri, etkinleştirme işlevi eğrisini yukarı veya aşağı kaydırmanıza olanak tanır.

Kısaca, aktivasyon fonksiyonları girdi değerlerini (0, 1) veya (-1, 1) gibi gerekli değerler arasında eşlemek için kullanılır. Perceptron genellikle verileri iki bölüme ayırmak için kullanılır. Bu nedenle, doğrusal ikili sınıflandırıcı olarak da bilinir.

Aktivasyon fonksiyonları doğrusal (linear) ve doğrusal olmayan (non-linear) olmak üzere 2 ye ayrılır. Evet veya hayır gibi sinir ağlarının çıktısını belirlemek için kullanılır. Elde edilen değerleri kullanarak aktivasyon fonksiyonuna göre 0 ila 1 veya -1 ila 1 değerleri arasına eşler.

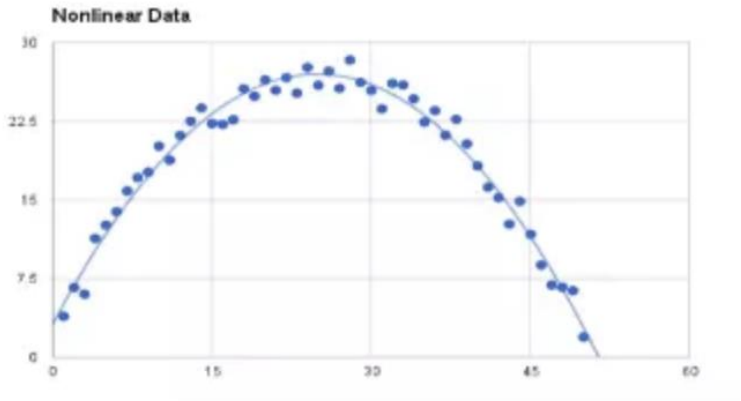
Aktivasyon Fonksiyonları

Doğrusal aktivasyon fonksiyonu



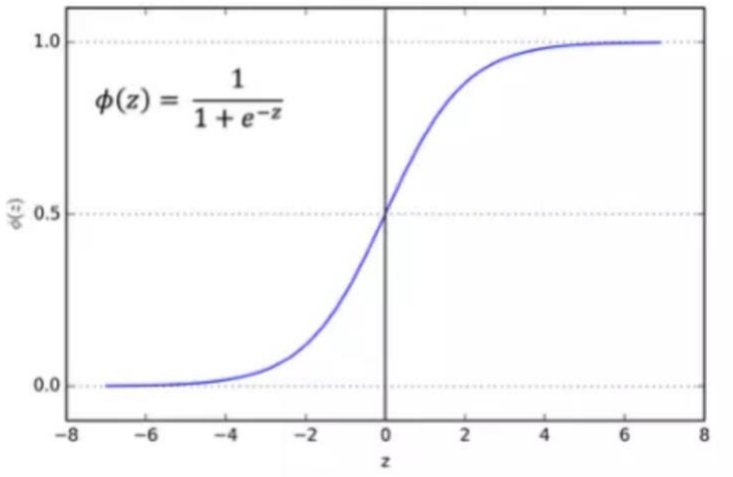
Doğrusal fonksiyon bir çizgi veya doğrusaldır. Bu nedenle, fonksiyonun çıktısı herhangi bir aralık arasında sınırlanmayacaktır. Sinir ağlarıyla beslenen olağan verilerin karmaşıklığında veya çeşitli parametrelerinde kullanılmaz. Türevlenebilir fonksiyon değildir, doğrusal olduğu için ve öğrenme işleminde geriye besleme yapıldığında sabit bir sayı olacağından öğrenme gerçekleşmeyecektir.

Doğrusal olmayan aktivasyon fonksiyonu



Doğrusal olmayan etkinleştirme işlevleri en çok kullanılan etkinleştirme işlevleridir. Doğrusal olmamak, grafiğin Şekildeki gibi görünmesine yardımcı olur. Modelin genelleştirilmesini veya çeşitli verilerle uyum sağlamasını ve çıktılar arasında ayrım yapmasını kolaylaştırır. Doğrusal olmayan aktivasyon fonksiyonları esas olarak aralıklarına veya eğrilerine göre bölünür.

Sigmoid veya lojistik aktivasyon fonksiyonu



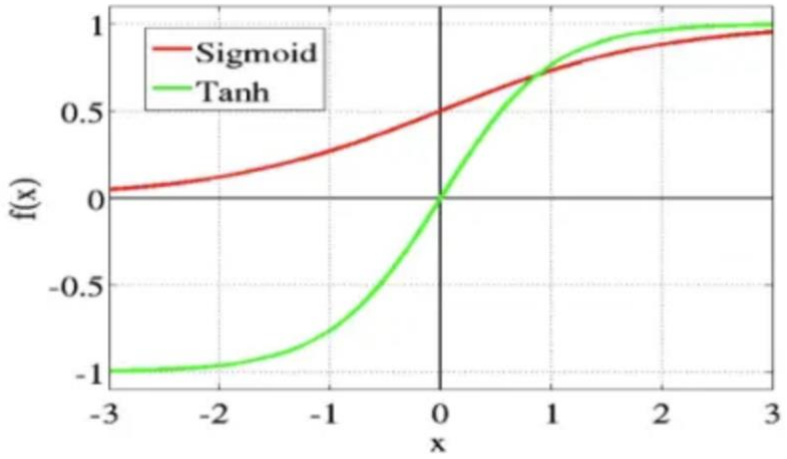
Şekilde görüldüğü gibi sigmoid fonksiyon eğrisi bir S şekline benzer. Sigmoid işlevini kullanmamızın ana nedeni, (0 ile 1) arasında var olmasıdır. Bu nedenle, özellikle olasılığı bir çıktı olarak tahmin etmemiz gereken modeller için kullanılır.

Herhangi bir şeyin olasılığı yalnızca 0 ile 1 aralığı arasında olduğundan, sigmoid doğru bir seçimdir. Fonksiyon türevlenebilirdir. Yani sigmoid eğrisinin eğimini herhangi iki noktada bulabiliriz.

Lojistik sigmoid işlevi, eğitim sırasında bir sinir ağının sıkışmasına neden olabilir.

Softmax fonksiyonu, çok sınıflı sınıflandırma için kullanılan daha genelleştirilmiş bir lojistik aktivasyon işlevidir.

Tanh or hyperbolic tangent aktivasyon fonksiyonu



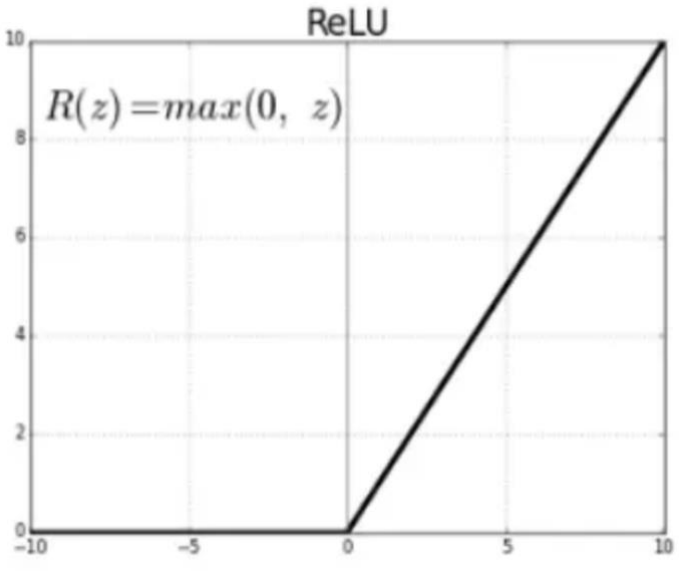
Tanh aslında lojistik sigmoid gibidir ancak ondan daha iyidir. Şekil 5'te görüldüğü gibi tanh fonksiyonunun aralığı (-1 ile 1) arasındadır. Bunun avantajı, negatif girişlerin güçlü bir şekilde negatif olarak eşleştirilmesi ve sıfır girişlerinin tanh grafiğinde sıfıra yakın bir şekilde eşlenmesidir.

Bu fonksiyon da türevlenebilir. Tanh işlevi esas olarak iki sınıf arasında sınıflandırmada kullanılır.

Hem tanh hem de lojistik sigmoid aktivasyon fonksiyonları ileri beslemeli ağlarda kullanılır.

ReLU (rectified linear unit) activation function

ReLU en çok kullanılan aktivasyon fonksiyonudur ve neredeyse tüm evrişimli sinir ağlarında veya derin öğrenmede kullanılmaktadır.



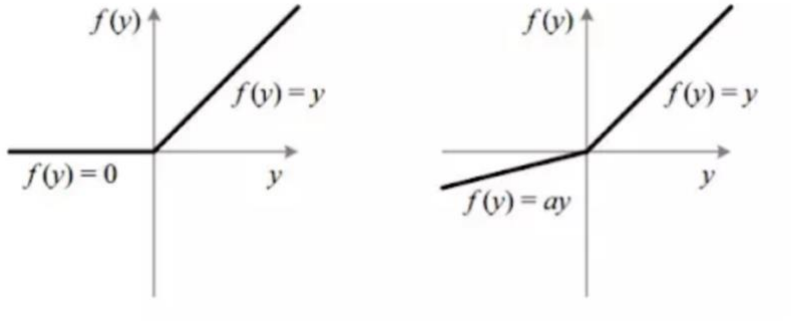
Şekilde görüldüğü üzere z sıfırdan küçük olduğunda f (z) sıfırdır ve z sıfırdan büyük veya eşit olduğunda f (z) z'ye eşittir. Aralık 0 dan sonsuzadır.

Fonksiyon ve türevi monotondur. Ancak sorun, tüm negatif değerlerin anında sıfırlanmasıdır ve bu da modelin verilere uygun şekilde uyması veya eğitilmesi yeteneğini azaltır. Bu, ReLU aktivasyon fonksiyonuna verilen herhangi bir negatif girdinin, değeri grafikte hemen sıfıra çevirdiği anlamına gelir, bu da negatif değerlerle eğitilmesini engeller.

Leakly ReLU

ReLU deki negatif değerlerin yok olması sorununu çözmek için geliştirilmiştir. Şekil 7'de görüldüğü üzere negatif bölgelerde türevin 0 olup öğrenmenin devam etmesini sağlar.

Ancak hız bakımından daha hızlı olmasını istiyorsak ReLU daha çok tercih edilmektedir. Genellikle ara katmanlarda ReLU tercih edilir.



Sinir ağları temel yapısı

En yaygın dört sinir ağı katmanı türü:

- **Tam bağlantılı (fully connected):** Tamamen bağlı katmanlar, bir katmandaki her nöronu bir sonraki katmandaki her nörona bağlar. Tamamen bağlı katmanlar, standart sinir ağlarından evrişimli sinir ağlarına (CNN) kadar tüm farklı sinir ağlarında bulunur.
- **Evrişimli (convolutional)**
- **Ters evrişim (deconvolutional)**
- **Tekrarlayan (recurrent)**