

★ Get unlimited access to the best of Medium for less than \$1/week. [Become a member](#)



# Yapay Sinir Ağlarında Aktivasyon Fonksiyonları Çeşitleri



Ayşegül Akbaş · [Follow](#)

Published in Kodcular

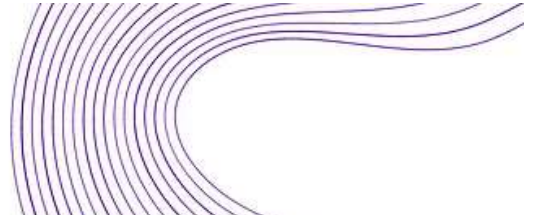
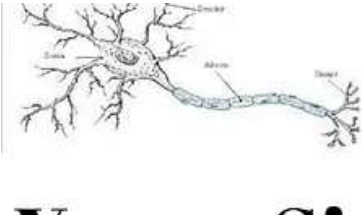
4 min read · Aug 16, 2021



Share



More



Open in app ↗



Search



## AKTİVASYON FONKSİYONU ÇEŞİTLERİ

Selamlar herkese,

Derin öğrenme algoritmalarıyla uğraşan kişiler yapay sinir ağları ve aktivasyon fonksiyonları gibi terimlere aşinalardır. Bu yazımda yapay sinir ağlarını eğitirken sıkça kullandığımız aktivasyon fonksiyonlarından ve çeşitlerinden bahsedeceğim. Umarım herkes için faydalı olur.

### Aktivasyon Fonksiyonu Nedir?

Yapay sinir ağlarına kompleks verileri öğretebilmemiz için aktivasyon fonksiyonları gereklidir. Aktivasyon fonksiyonlarının amacı weight(ağırlık) ve bias değerlerini ayarlamaktır.

$L=x.w+b$  işlemini yaptıktan sonra bu değeri bir aktivasyon fonksiyonundan geçiriyoruz. Bu fonksiyonlar tensorlere etki eder ve aslında bunu nöronların tetiklenmesi olarak da düşünebiliriz.

$$L = X \cdot W + b$$

weight

input

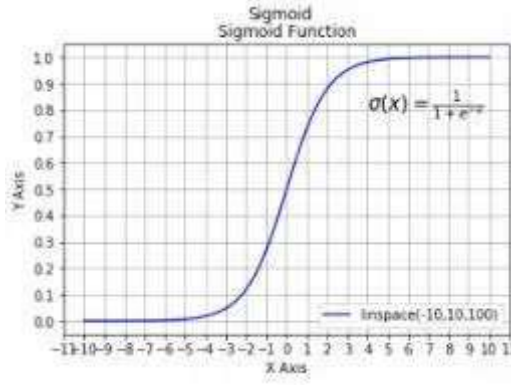
bias

### Aktivasyon Fonksiyonu Çeşitleri

Birden fazla aktivasyon fonksiyonu bulunmaktadır. Burada aktivasyon fonksiyonları çeşitlerinden bahsedeceğim. Dezavantaj ve avantajlarına göre hangisinin daha sık ya da daha az kullanıldığını hep birlikte görelim.

#### 1. Sigmoid Fonksiyonu

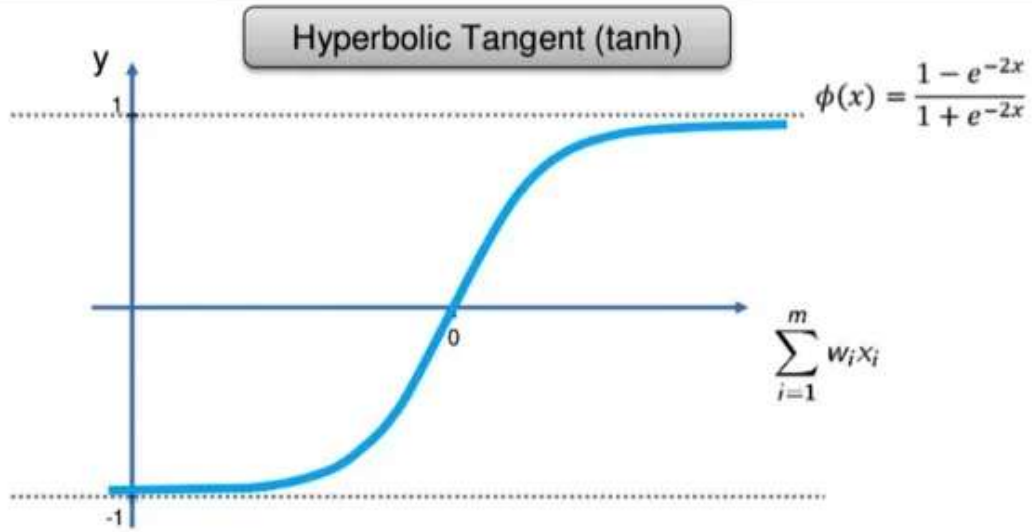
Geçmişte oldukça popüler olan bu fonksiyon aldığı değerleri 0 ve 1 arasına hapseder. Yüksek bir değer geldiğinde 1'e yakın olurken düşük bir değer geldiğinde 0'a yakın olur. Dolayısıyla 1'i ve 0'ı geçen herhangi bir değer gözlemlenmez. Sigmoid fonksiyonunun dezavantajları ise birden fazladır. Gradient ölümünün fazla olması, 0 odaklı bir fonksiyon olmaması ve  $\exp()$  hesaplamalarının oldukça yavaş olması bunlara örnek olarak gösterilebilir.



Sigmoid Fonksiyonu Grafiği

## 2. tanh (Hiperbolik Tanjant) Fonksiyonu

Hiperbolik tanjant kısaca “tanh”’n sigmoid’e benzerliği ile bilinir. İkisinde de sıkıştırma bulunur ancak tanh gelen değerleri  $[-1,1]$  arasına hapseder. Sigmoid fonksiyonuna göre daha iyi sonuçlar alınabilir çünkü 0 odaklıdır. Özellikle LSTM ve GRU gibi sinir ağlarında yaygın olarak kullanılan bir aktivasyon fonksiyonudur. Ancak gradient ölümleri hala devam etmektedir.

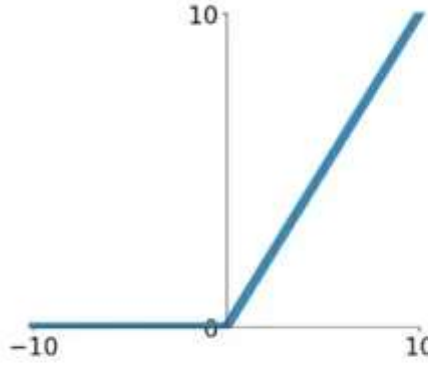


Tanh Fonksiyonu Grafiği

## 3. ReLu Fonksiyonu

Günümüzde en sık kullanılan ve popülerliğini koruyan aktivasyon fonksiyonudur. İlk olarak 2012 yılında kullanıldı ve günümüze kadar popülerliği giderek arttı. Ayrıca ReLu fonksiyonu biyolojik nörona benzerliğiyle bilinir. Çalışma mantığı olarak gelen değerlerin pozitif mi negatif mi olduğuna bakar sonrasında eğer gelen değer negatif bir değer ise işlem sonucunu 0 verir. Ancak gelen değer pozitifse herhangi bir sıkıştırma ya da değiştirme işlemi uygulamaz olduğu gibi geçer.

Bilgisayarlar tanh ve sigmoid'de karmaşık hesaplamalar yaparken ReLu'da yalnızca pozitiflik negatiflik durumuna bakar. Bu nedenle bilgisayar bu denklemi 6 kat daha hızlı hesaplar. Kısacası hızlıdır. Ancak muhteşem bir fonksiyon diyemeyiz çünkü bu fonksiyonda 0 odaklı olmadığından bazı nöronlar (%40'a kadar) ölebiliyor.

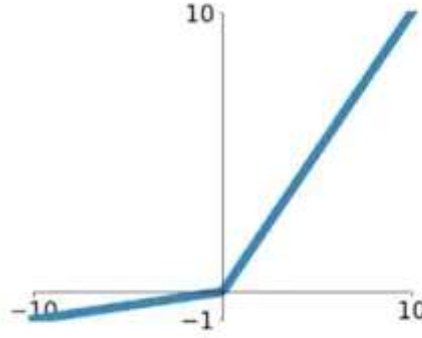


$$F(x) = \max(0, x)$$

ReLu Fonksiyonu Grafiği

#### 4. Leaky ReLu Fonksiyonu

ReLu fonksiyonlarında dezavantajlardan biri olarak nöronların ölümünden bahsetmiştik. Hatta bu ölümler bazen %40'a kadar çıkabilmektedir. İşte Leaky ReLu fonksiyonu tam burada devreye giriyor. Ölü nöron sonunu ortadan kaldırmak için geliştirilmiş bir fonksiyondur diyebiliriz. Negatif bir değer geldiğinde çok küçük bir sayı döndürür ve bu sayede nöronların ölmesinin önüne geçer. Hesaplama sayısının artmasına rağmen diğer fonksiyonlara göre oldukça hızlı çalışır. Son olarak ReLu'nun sahip olduğu tüm avantajları barındırır.

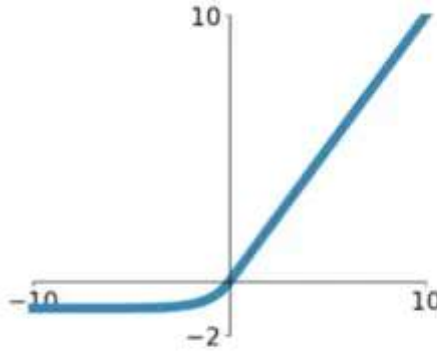


$$F(x) = \max(0.01x, x)$$

Leaky ReLu Fonksiyonu Grafiği

## 5. Exponential Linear Unit (ELU)

Bu fonksiyonda orta nokta 0 olduğundan ölü nöron sorunun önüne geçebilir. Exponential yani  $e^x$  cinsinden hesaplama yaptığı için oldukça yavaştır. Ancak ReLu'nun avantajlarına sahiptir.



$$f(\alpha, x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$$

Exponential Linear Unit (ELU) Grafiği

## 6. Maxout Fonksiyonu

Kullanılan bir diğer aktivasyon fonksiyonu ise Maxout fonksiyonudur. ReLu ve Leaky ReLu'yu genelleştiren bir fonksiyondur ancak parametre sayısı 2 katına çıktığından ReLu'ya göre oldukça yavaş çalışır. Avantajı ise ReLu'da olduğu gibi nöron ölümü

gerçekleşmez. Son olarak maxout fonksiyonunun herhangi bir grafiksel gösterimi yoktur.

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

Maxout Fonksiyonunun Formülü

### Hangi Aktivasyon Fonksiyonunu Kullanmalıyız?

Peki bu fonksiyonlardan ve özelliklerinden bahsettik, dezavantaj ve avantajlarını sıraladık. Hangisini kullanmalıyız sorusuna nasıl bir cevap verebiliriz?

Geçmiş dönemlerde son derece popüler olan “sigmoid fonksiyonu’nun” yerini artık ReLu Fonksiyonu almıştır. Dolayısıyla çalışmalarınızda ReLu fonksiyonu kullanmanızı şiddetle tavsiye ederim.

*Bu yazımda sizlere derin öğrenme aktivasyon fonksiyonlarının çeşitlerini, performanslarını ve hangi koşullarda hangi fonksiyonun etkili olduğunu ve özelliklerini açıklamaya çalıştım. Umarım derin öğrenme algoritmalarıyla ilgilenen kişiler ve diğer herkes için faydalı olabilmişimdir :)*

Bir daha ki yazımda görüşmek üzere..

### Referanslar

- Kızrak, A. (2020, January 7). *Derin Öğrenme için Aktivasyon Fonksiyonlarının Karşılaştırılması*. Medium. <https://ayyucekizrak.medium.com/derin-%C3%B6%C4%9Frenme-i%C3%A7in-aktivasyon-fonksiyonlar%C4%B1n%C4%B1n-kar%C5%9F%C4%B1la%C5%9Ft%C4%B1r%C4%B1lmas%C4%B1-cee17fd1d9cd>
- Kumawat, D. (2019, August 22). *7 types of activation functions in neural network*. Analytics Steps | Learn AI, Machine Learning, Business Analytics, Startups & other Technologies. <https://www.analyticssteps.com/blogs/7-types-activation-functions-neural-network>

Computer Science



Follow

## Written by Ayşegül Akbaş

203 Followers · Writer for Kodcular

Mobile and iOS developer || computer engineer || read. learn. share.

More from Ayşegül Akbaş and Kodcular



# Swift



Ayşegül Akbaş in Kodcular

## Swift: Nedir Bu Closures?

Closure Nedir?

3 min read · Feb 26, 2022



147



2



Tahsin Safa Elmalı in Kodcular

## Nedir Bu Big O Notation?

Günümüzde sahip olduğumuz en değerli şeylerden birinin zaman olduğunu biliyoruz. Bu yüzden bir kullanıcın uygulamamız üzerinden yapacağı...

8 min read · Oct 10, 2019



1.7K



6





```
function foodPrices(prices, addValue){
  var newPriceList = [];

  for(var i = 0; i<prices.length; i++){
    var finalPrice = prices[i] + addValue;
    newPriceList.push(finalPrice);
  }

  console.log(i) //3
  console.log(finalPrice) //20

  return newPriceList;
}

const result = foodPrices([10,20,30],10);
console.log(result);
```



Tahsin Safa Elmalı in Kodcular

## Javascriptte var-let-const nedir? Aralarında ki farklar nelerdir?

Merhaba arkadaşlar, bugün sizlere ECMAScript 2015 (ES6) ile gelen let ve const ile değişken oluşturmanın yollarından bahsetmek istiyorum...

6 min read · Feb 16, 2019



1.5K



6



```
inet6 fe80::a00:27ff:fef2:4edc prefixlen 64 scopeid 0x20<link>
ether 08:00:27:f2:4e:dc txqueuelen 1000 (Ethernet)
RX packets 582 bytes 790492 (771.9 KiB)
RX errors 0 dropped 0 overruns 0 frame 0
TX packets 442 bytes 28737 (28.0 KiB)
TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
inet 127.0.0.1 netmask 255.0.0.0
inet6 ::1 prefixlen 128 scopeid 0x10<host>
loop txqueuelen 1000 (Local Loopback)
RX packets 20 bytes 1116 (1.0 KiB)
RX errors 0 dropped 0 overruns 0 frame 0
TX packets 20 bytes 1116 (1.0 KiB)
TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

wlan0: flags=4099<UP,BROADCAST,MULTICAST> mtu 1500
ether 1e:76:49:00:a7:a0 txqueuelen 1000 (Ethernet)
RX packets 0 bytes 0 (0.0 B)
RX errors 0 dropped 0 overruns 0 frame 0
TX packets 0 bytes 0 (0.0 B)
```



Ayşegül Akbaş in Kodcular

## MAC adresi nedir? Kali Linux'da Mac Adresi Nasıl Değiştirilir?

Mac (Media Access Control) adresi, ağa bağlı bir Network Interface Card-NIC (Ethernet Kartı-Ağ Kartı) devre kartındaki Chip üzerinde...

2 min read · Apr 26, 2020



51



2



See all from Ayşegül Akbaş

See all from Kodcular

### Recommended from Medium



James Presbitero Jr. in Practice in Public

## These Words Make it Obvious That Your Text is Written By AI

These 7 words are painfully obvious. They make me cringe. They will make your reader cringe.

4 min read · Jan 1



32K



847



Sheila Teo in Towards Data Science

## How I Won Singapore's GPT-4 Prompt Engineering Competition

A deep dive into the strategies I learned for harnessing the power of Large Language Models (LLMs)



• 23 min read • Dec 29, 2023



10.7K



122



### Lists



#### AI Regulation

6 stories • 302 saves



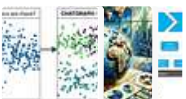
#### ChatGPT

21 stories • 435 saves



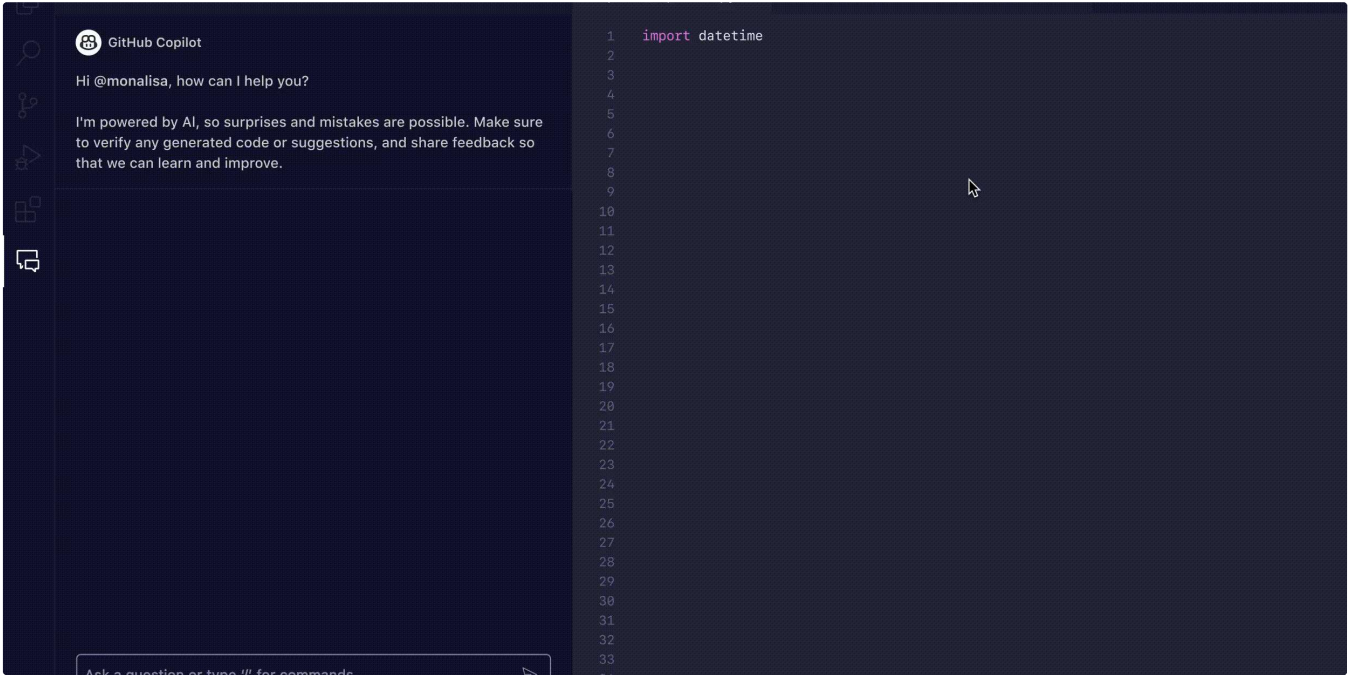
#### Natural Language Processing

1147 stories • 618 saves



#### ChatGPT prompts

36 stories • 1069 saves



Jacob Bennett in Level Up Coding

## The 5 paid subscriptions I actually use in 2024 as a software engineer

Tools I use that are cheaper than Netflix

🌟 · 5 min read · Jan 4

👏 5.8K    💬 69

🔖+    ⋮

Anmol Tomar in CodeX

# Say Goodbye to Loops in Python, and Welcome Vectorization!

Use Vectorization—a super-fast alternative to loops in Python

🌟 · 5 min read · Dec 28, 2023

👏 3.5K    💬 42

🔖  
...

AL Anany

# The ChatGPT Hype Is Over—Now Watch How Google Will Kill ChatGPT.

It never happens instantly. The business game is longer than you know.

🌟 · 6 min read · Sep 1, 2023

👏 21K    💬 643

🔖  
...



Unbecoming

## 10 Seconds That Ended My 20 Year Marriage

It's August in Northern Virginia, hot and humid. I still haven't showered from my morning trail run. I'm wearing my stay-at-home mom...

★ · 4 min read · Feb 16, 2022



74K



1051



See more recommendations