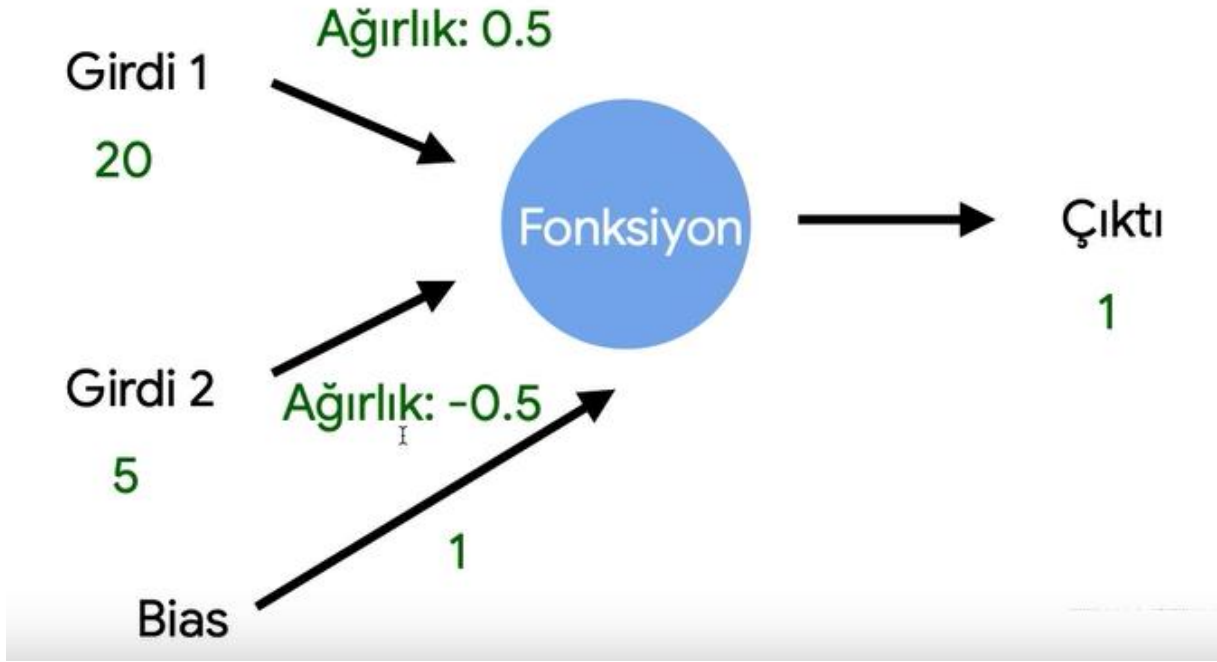


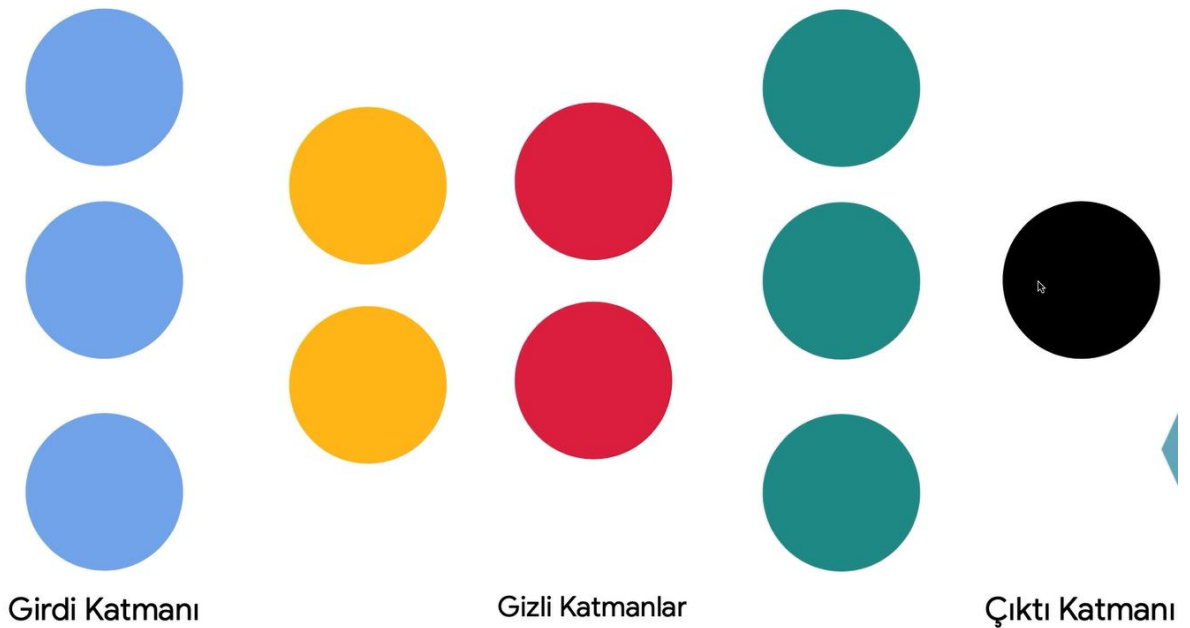
# TENSERFLOW

## Yapay Sinir Ağları (Artificial Neural Network)

Yapay sinir ağları, insan beyninin bilgi işleme tekniğinden esinlenerek geliştirilmiş bir bilgi işlem teknolojisidir. YSA ile basit biyolojik sinir sisteminin çalışma şekli taklit edilir. Yani biyolojik nöron hücrelerinin ve bu hücrelerin birbirleri ile arasında kurduğu sinaptik bağın dijital olarak modellenmesidir.



## Deep Network

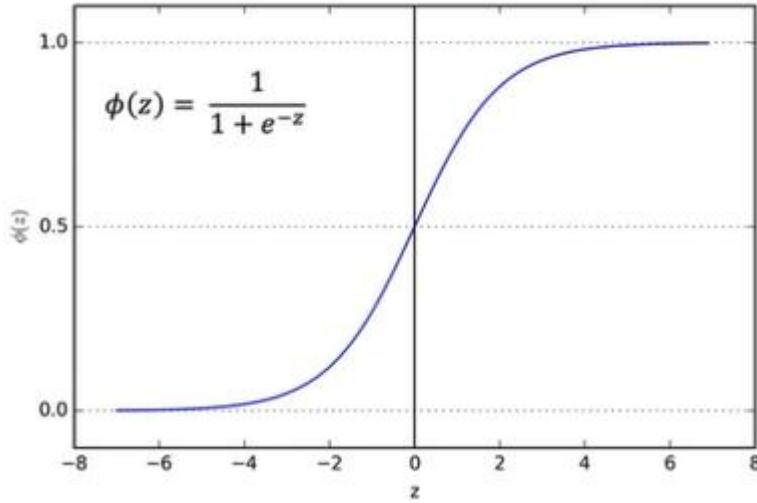


Veri ne kadar zor ve geniş kapsamlı ise o kadar fazla nörona ve katmana ihtiyaç duyulur. Gizli katman sayısı 2 den fazla olursa deep network (derin ağ) denir.

## Aktivasayon Fonsiyonları

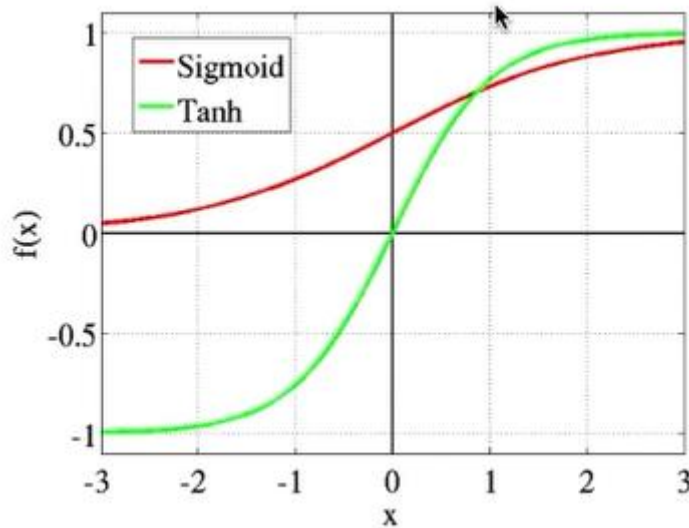
### 1. Sigmoid Fonsiyonu

- 0 ile 1 arasında değer alır.
- Genelde sınıflandırma problemlerinde işe yarar.



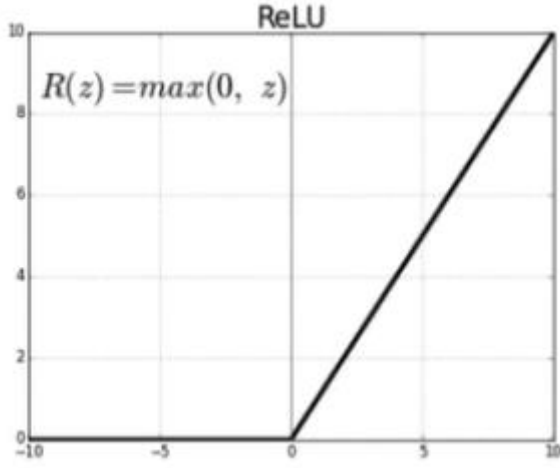
### 2. Tahn ( Hiperbolik Tanjant ) Fonsiyonu

- -1 ile 1 arasında değer alır.
- Negatif değerlerle daha geniş bir kapsam sağlar ve genelde sınıflandırma operasyonlarında kullanılır.



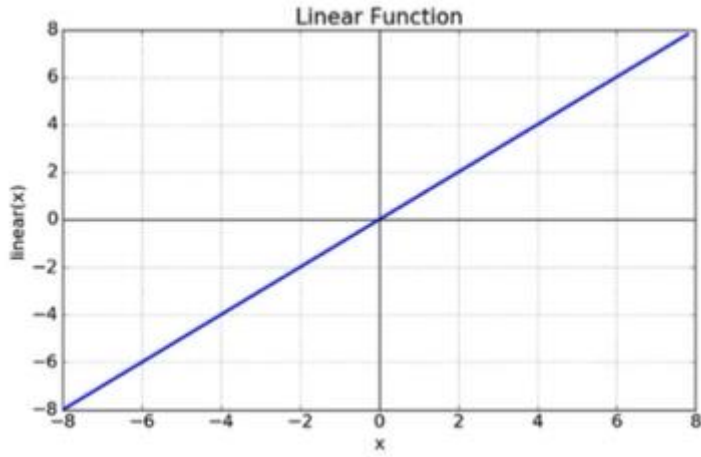
### 3. ReLU (Rectified Linear Unit)

- 0 ile sonsuz arasında değer alırlar.,
- Derin öğrenme alanında sıklıkla karşımıza çıkar.


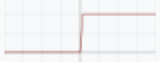


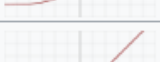

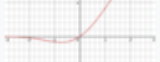


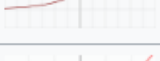


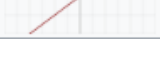


#### 4. Linear Fonksiyonlar

- $F(x)=x$
- Sonsuz değer alabilir fakat non-linear olmaması sebebiyle modellerde sorunlara yol açabilir.

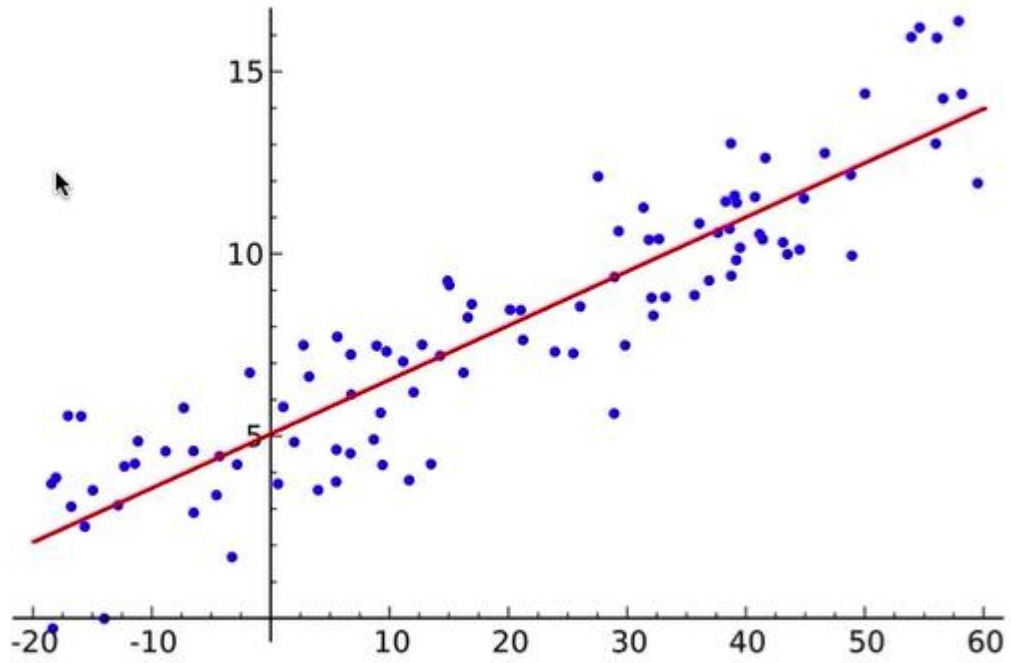


Bütün aktivasyon fonksiyonları aşağıda tabloda gösterilmektedir.

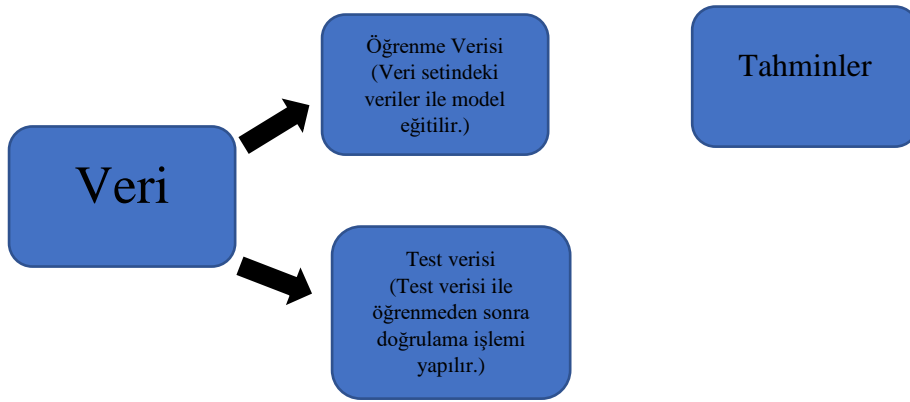
Name	Plot	Function, $g(x)$	Derivative of $g$ , $g'(x)$	Range	Order of continuity
Identity		$x$	1	$(-\infty, \infty)$	$C^\infty$
Binary step		$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$	0	$\{0, 1\}$	$C^{-1}$
Logistic, sigmoid, or soft step		$\sigma(x) \doteq \frac{1}{1 + e^{-x}}$	$g(x)(1 - g(x))$	$(0, 1)$	$C^\infty$
Hyperbolic tangent (tanh)		$\tanh(x) \doteq \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$1 - g(x)^2$	$(-1, 1)$	$C^\infty$
Rectified linear unit (ReLU) <sup>[8]</sup>		$(x)^+ \doteq \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$ $= \max(0, x) = x \mathbf{1}_{x>0}$	$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$	$[0, \infty)$	$C^0$
Gaussian Error Linear Unit (GELU) <sup>[2]</sup>		$\frac{1}{2}x \left(1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right)$ $= x\Phi(x)$	$\Phi(x) + x\phi(x)$	$(-0.17 \dots, \infty)$	$C^\infty$
Softplus <sup>[9]</sup>		$\ln(1 + e^x)$	$\frac{1}{1 + e^{-x}}$	$(0, \infty)$	$C^\infty$
Exponential linear unit (ELU) <sup>[10]</sup>		$\begin{cases} \alpha(e^x - 1) & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$ with parameter $\alpha$	$\begin{cases} \alpha e^x & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$	$(-\alpha, \infty)$	$\begin{cases} C^1 & \text{if } \alpha = 1 \\ C^0 & \text{otherwise} \end{cases}$
Scaled exponential linear unit (SELU) <sup>[11]</sup>		$\lambda \begin{cases} \alpha(e^x - 1) & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$ with parameters $\lambda = 1.0507$ and $\alpha = 1.67326$	$\lambda \begin{cases} \alpha e^x & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$	$(-\lambda\alpha, \infty)$	$C^0$
Leaky rectified linear unit (Leaky ReLU) <sup>[12]</sup>		$\begin{cases} 0.01x & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$	$\begin{cases} 0.01 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$	$(-\infty, \infty)$	$C^0$
Parametric rectified linear unit (PReLU) <sup>[13]</sup>		$\begin{cases} \alpha x & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$ with parameter $\alpha$	$\begin{cases} \alpha & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$	$(-\infty, \infty)$	$C^0$
Sigmoid linear unit (SiLU) <sup>[2]</sup> Sigmoid shrinkage, <sup>[14]</sup> SiL <sup>[15]</sup> or Swish-1 <sup>[16]</sup>		$\frac{x}{1 + e^{-x}}$	$\frac{1 + e^{-x} + xe^{-x}}{(1 + e^{-x})^2}$	$[-0.278 \dots, \infty)$	$C^\infty$
Gaussian		$e^{-x^2}$	$-2xe^{-x^2}$	$(0, 1]$	$C^\infty$

## REGRESYON

- Çocukların boyları ile babalarının boyları arasında bir ilişki olabilir mi ?(Francis Galyon)
- Çocukların boyları toplam veri setindeki ortalamaya yakın olma eğilimleri gösterebilir.
- $Y=a*x+b$



Kırmızı çizgi regresyon çizgisidir.

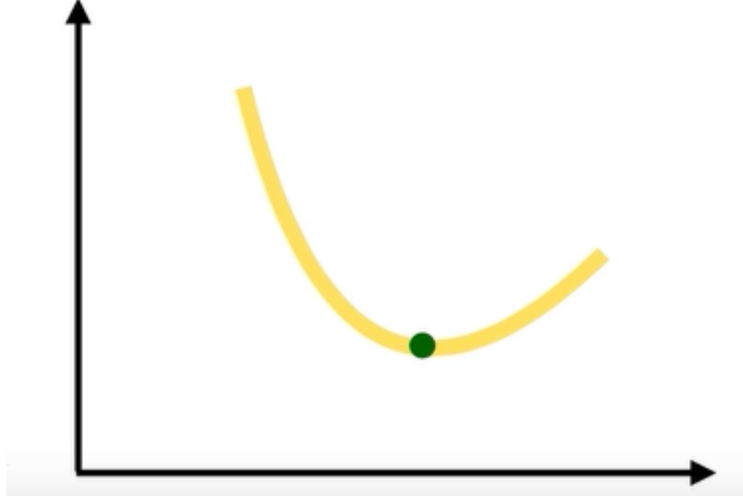


## Learning Rate

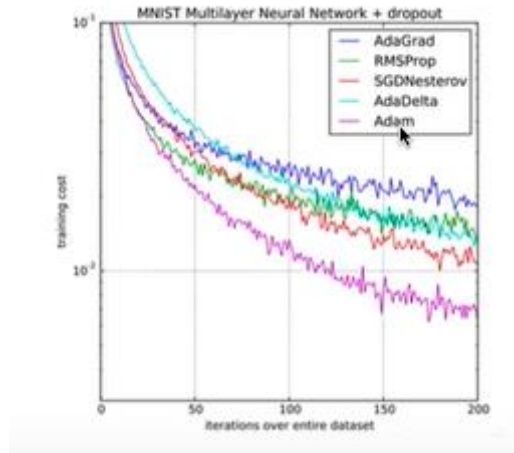
Öğrenme oranı da bizim optimize etmeye çalıştığımız değerlerdendir. Maliyet fonksiyonu için çok önemlidir. Öğrenme oranını çok büyük olursa minimum maliyet noktasını kaçırabiliriz çok küçük olursa da yapacağımız işlem çok uzun sürer bu yüzden optimize etmemiz gereken bir değerdir.

## Maliyet Fonksiyonu

Gradient Descent bir fonksiyonun minimumunu bulmak için kullandığımız optimizasyon fonksiyonudur. Maliyet fonksiyonunu minimize etmek için de kullanabiliriz.

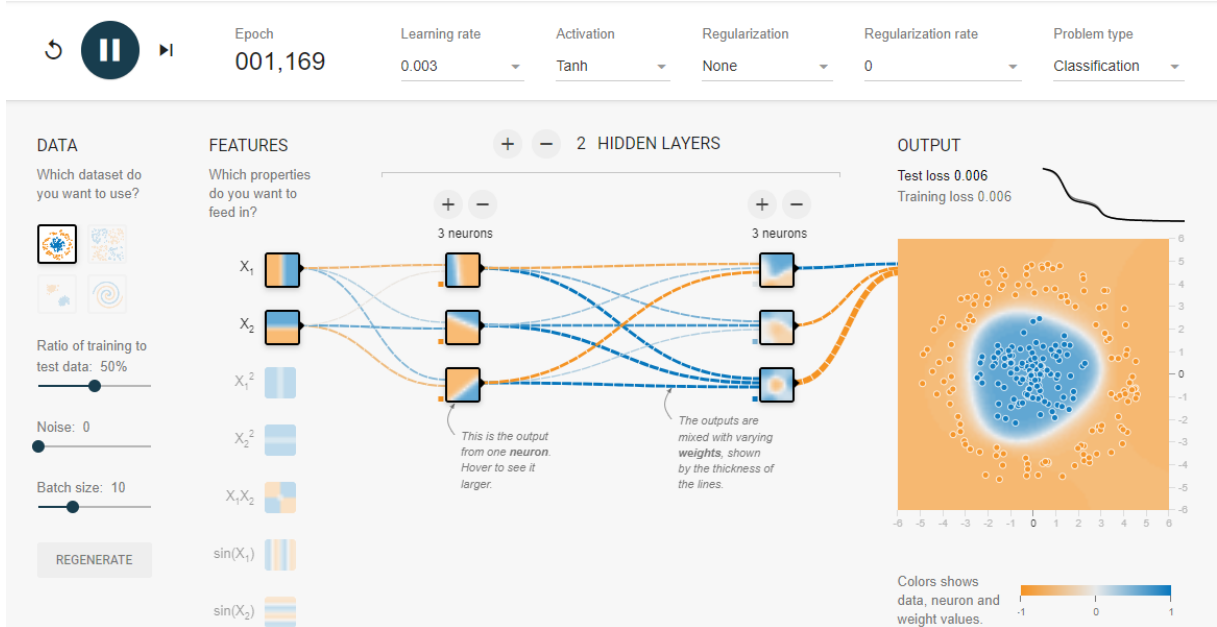
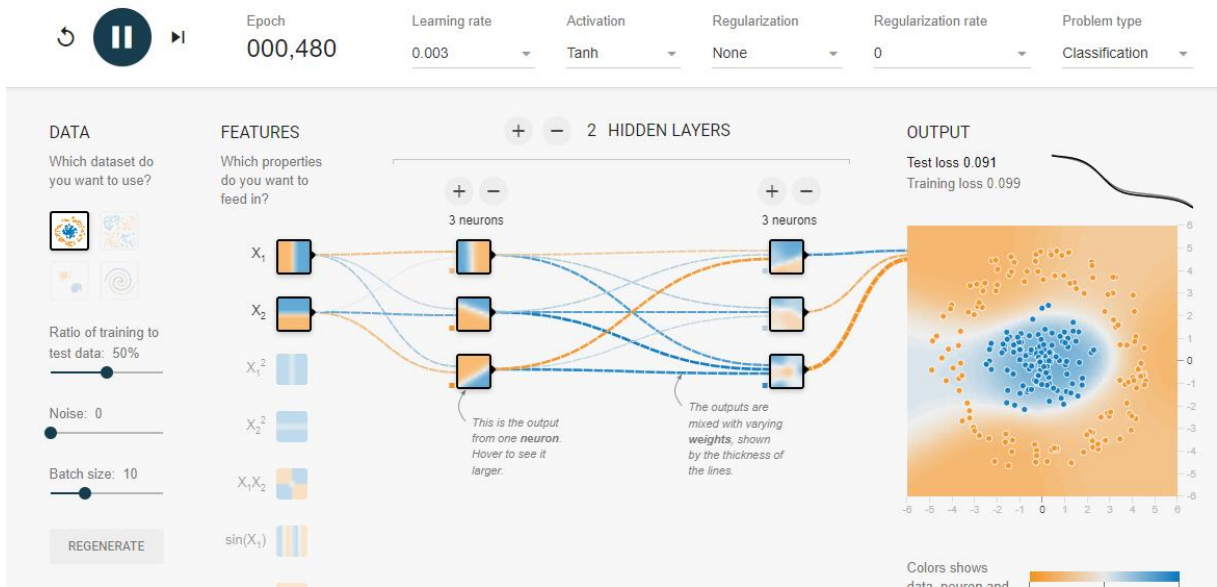


Aşağıdaki grafikte farklı optimizasyon algoritmalarının nasıl çalıştığını gösteren bir grafiktir.



Noise :Veri setindeki gürültüyü ifade eder.

Batch size : Veri setindeki verileri ne kadar vereceğimizi ayarlarız.



[ANN PLAYGROUND](#) sayfasından veri setindeki verileri yapay sinir ağı ile nasıl sınıflandırıldığını test edebilirsiniz.