BİL 475 Örüntü Tanıma

Hafta-8:

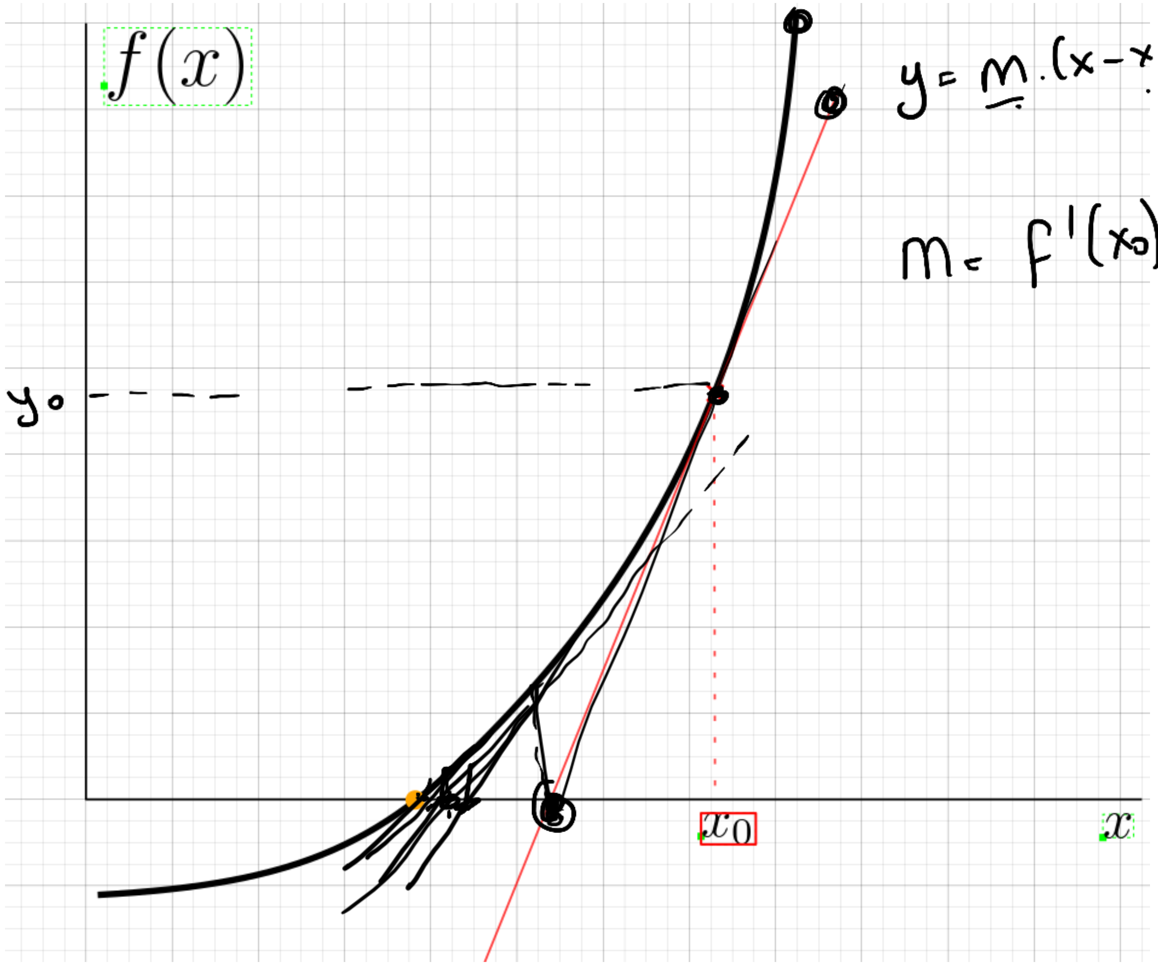
Yapay Sinir Ağları

Gradyan Azalma Algoritması

$$\boxed{w_{t+1} = w_t - \mu \frac{dE}{dw_t}}$$

Gradyan İnişi : Vekil Fonksiyon Kavramı

Kök Bulma Problemi !



$y = m.(x - x_0) + y_0$

$m = f'(x_0)$

$y = f'(x_0).(x - x_0) + y_0$

$0 = m.x - m.x_0 + y_0$

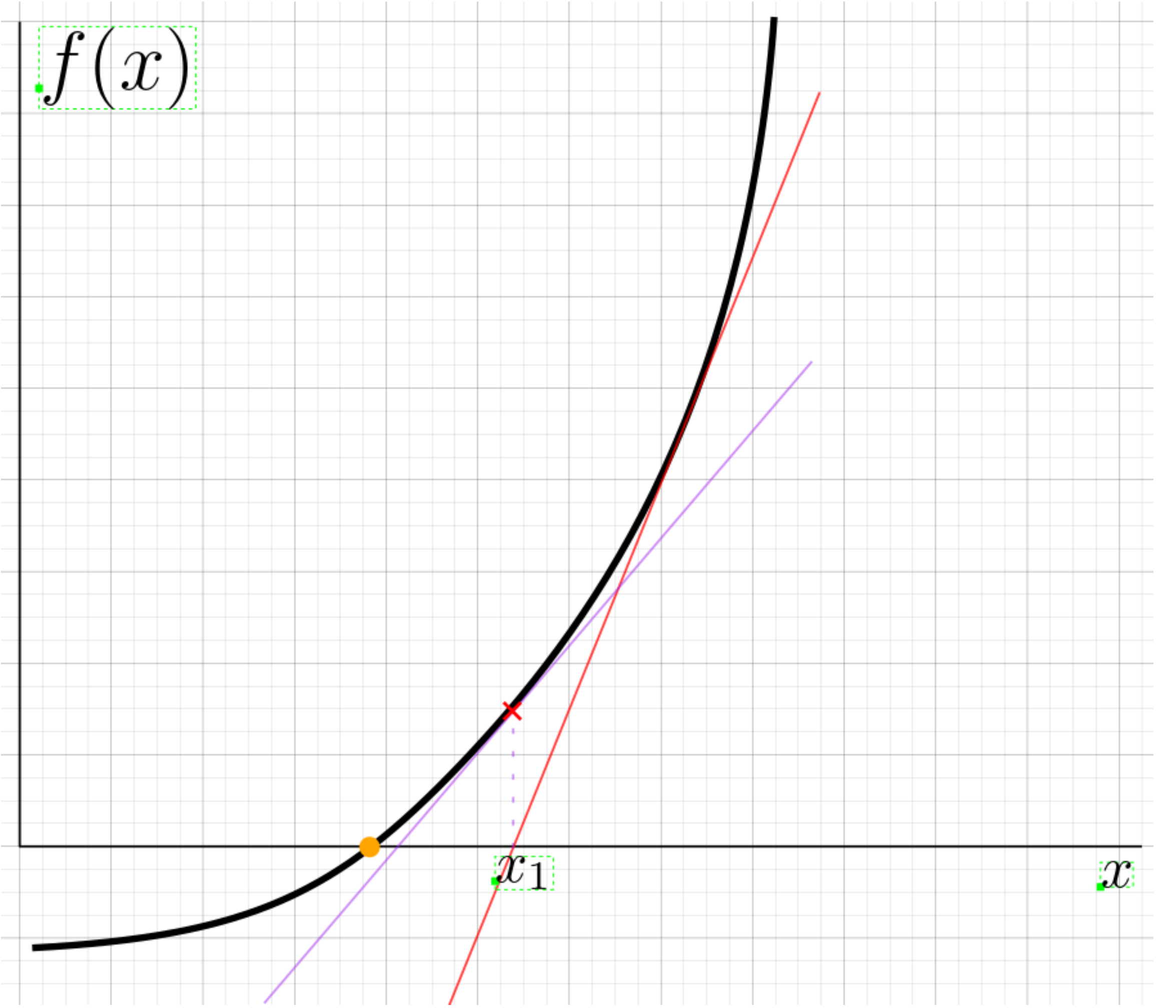
$f(x_0) = 0$

$m.x = m.x_0 - y_0$

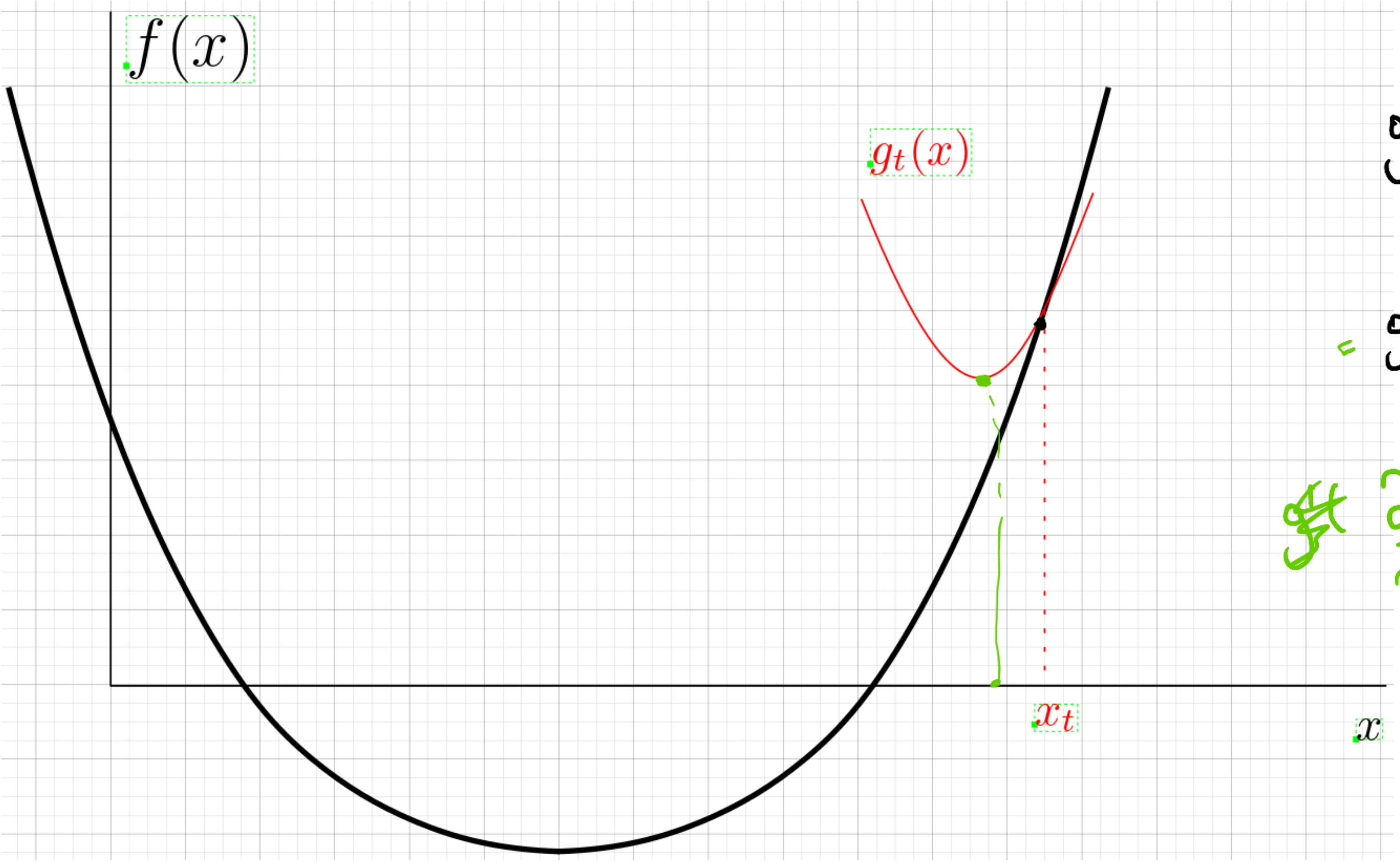
$$\boxed{x = x_0 - \frac{y_0}{m}}$$

$$\boxed{x = x_0 - \frac{f(x_0)}{f'(x_0)}}$$

Gradyan İnişi : Vekil Fonksiyon Kavramı



$$g_t(x) = \underline{f(x_t)} + \underline{f'(x_t)}(x - x_t)$$



$g(x_t) = f(x_t) + f'(x_t).(x - x_t) + \frac{1}{2} f''(x_t).(x-x$

$= g(x) = f(x_t) + f'(x_t).(x - x_t) + \frac{\alpha}{2}.(x - x_t)$

$\frac{\partial g(x)}{\partial x} = 0 + f'(x_t) + \alpha(x - x_t).1$

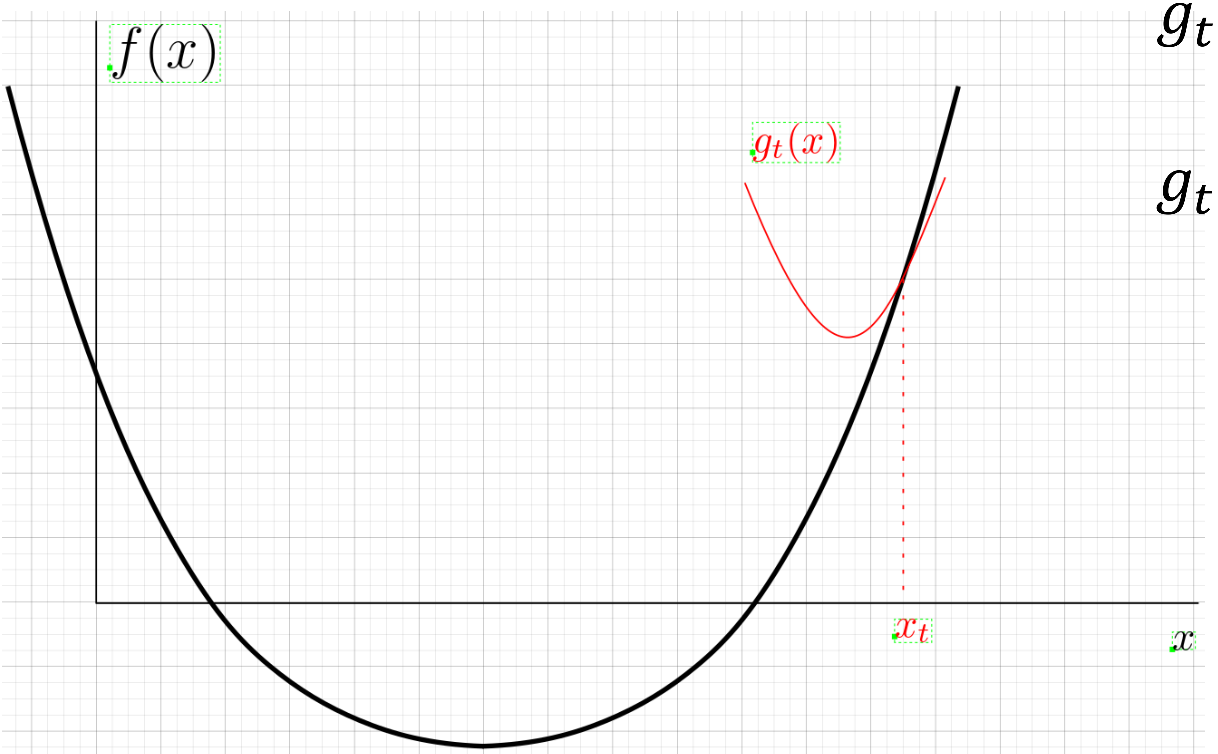
$0 = f'(x_t) + \alpha.x - \alpha.x_t$

$\alpha.x = \alpha.x_t - f'.(t_t)$

$x = x_t - \frac{f'(x_t)}{\alpha} \quad \mu = \frac{1}{\alpha}$

$x = x_t - \mu f'(x_t)$

Gradyan İnişi : Vekil Fonksiyon Kavramı



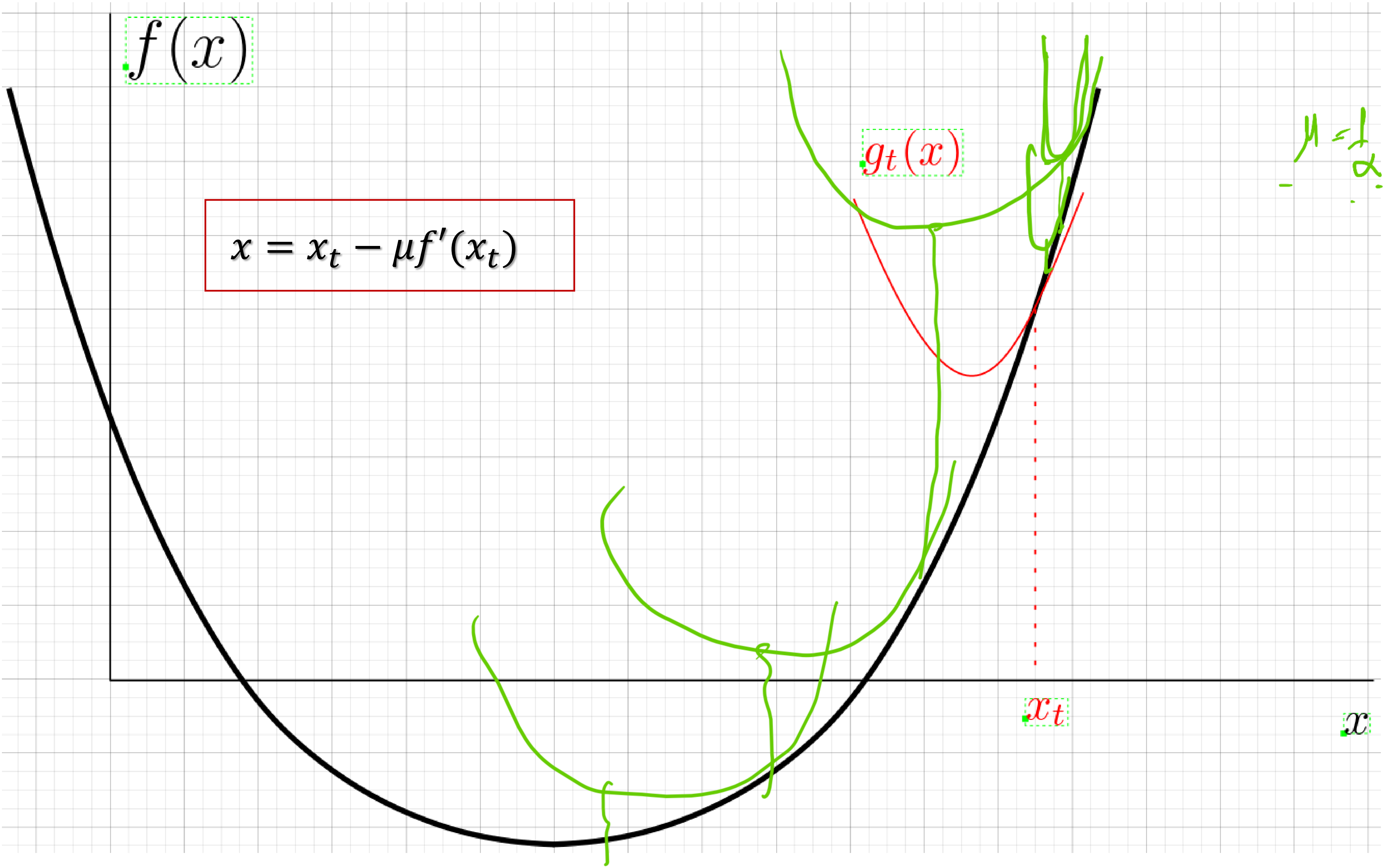
$$g_t(x) = f(x_t) + f'(x_t)(x - x_t) + f''(x_t)(x - x_t)^2$$

$$g_t(x) = f(x_t) + f'(x_t)(x - x_t) + \alpha(x - x_t)^2$$

$$g_t'(x) = 0$$

$$0 = f'(x_t) + 2\alpha(x - x_t)$$

$$\boxed{x = x_t - \mu f'(x_t)}$$



$f(x)$
$x = x_t - \mu f'(x_t)$
$g_t(x)$
x_t
x
$\mu = \frac{1}{\alpha}$

$f'(x)$ $\quad$ $f'(x_0)$

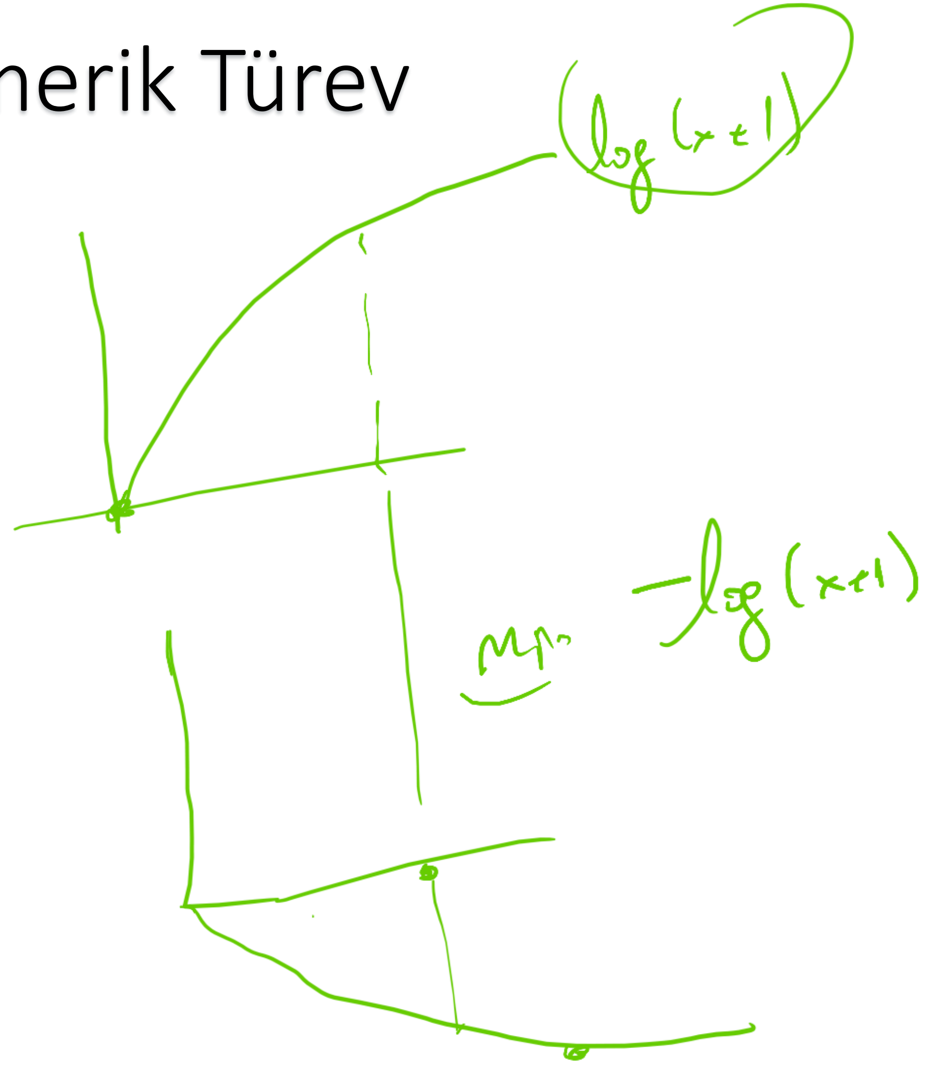
Gradyan İnişi – Nümerik Türev

$$\frac{df(x)}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

$$\frac{f(x_0+h) - f(x_0)}{h} \qquad h = 10^{-7}$$

$$\frac{f(x) - f(x-h)}{h}$$

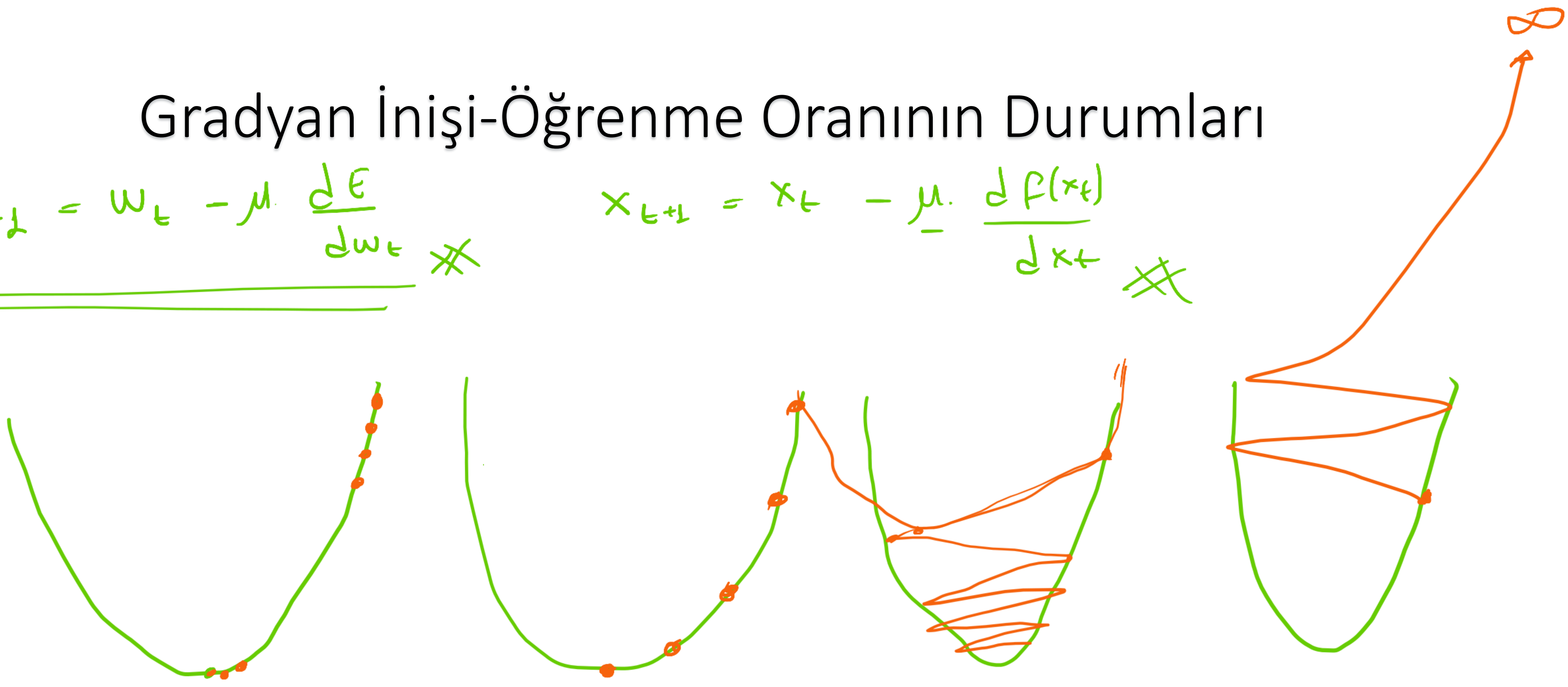
$$\frac{f(x+h) - f(x-h)}{2h}$$



Gradyan İnişi-Öğrenme Oranının Durumları

$$w_{t+1} = w_t - \mu \cdot \frac{dE}{dw_t}$$

$$x_{t+1} = x_t - \mu \cdot \frac{df(x_t)}{dx_t}$$



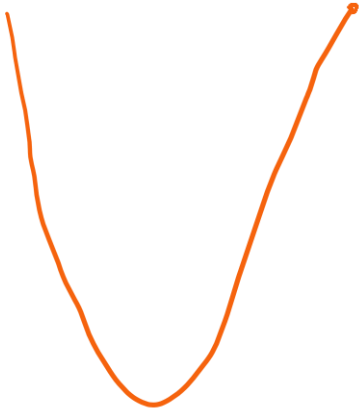
$\mu <<$ $\mu \sim$ $\mu >$ $\mu >>$

Gradyan İnişi

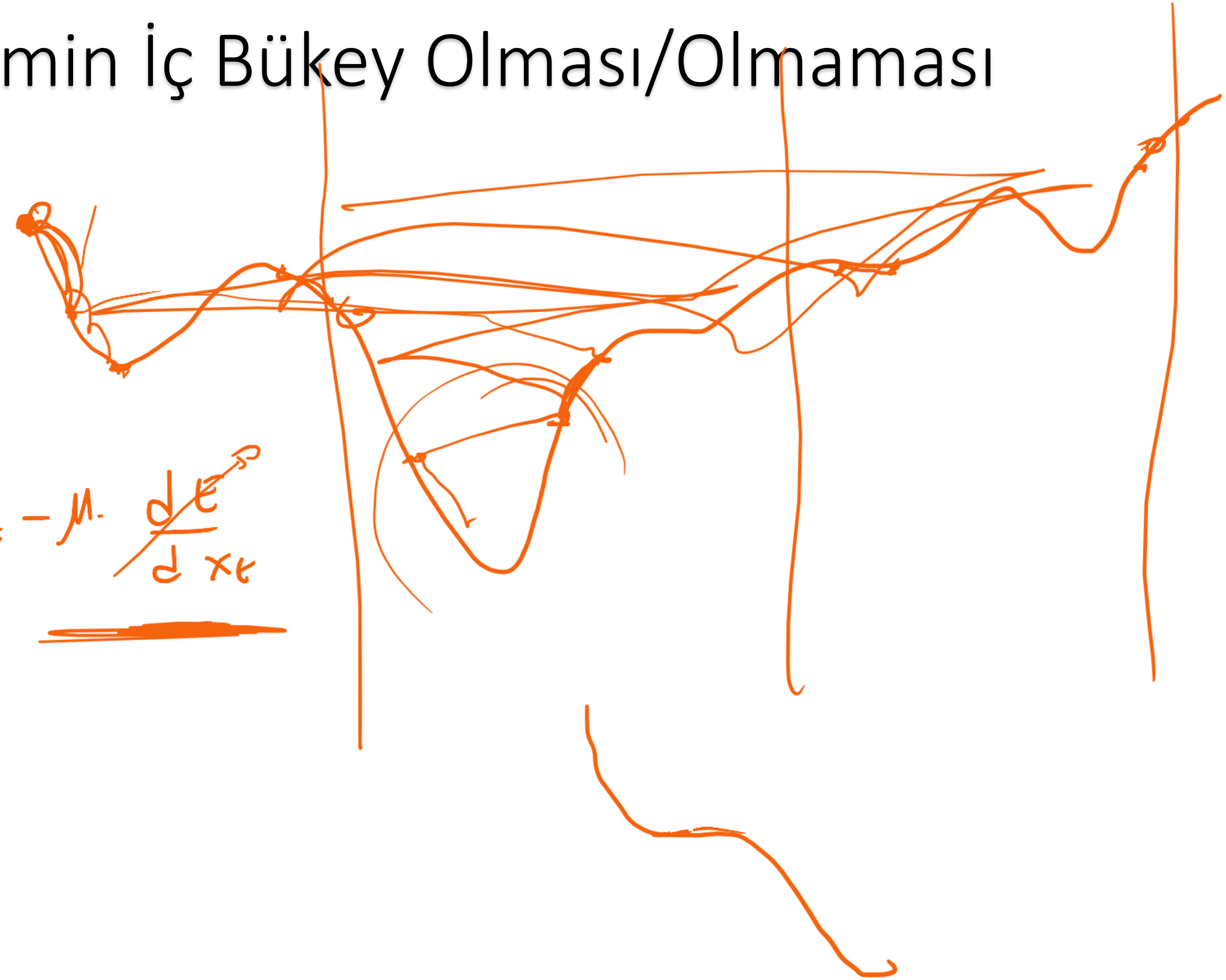
S3C (P.10) $f(x) = (x-2)^2$ kayıp fonksiyonunun $x_0 = 10$ noktasından başlayarak 3 adım boyunca gradyan inişini hesaplayınız. ($lr = 0.5$)

x	f(x)	f'(x)	x_{t+1}

Gradyan İnişi-Problemin İç Bükey Olması/Olmaması



$$x_{t+1} = x_t - \mu \cdot \frac{dE}{dx_t}$$



Yapay Sinir Ağları

Bulletin of Mathematical Biology Vol. 52, No. 1/2, pp. 99–115, 1990.
Printed in Great Britain.

0092–8240/90$3.00 + 0.00
Pergamon Press plc
Society for Mathematical Biology

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY*

■ WARREN S. MCCULLOCH AND WALTER PITTS
University of Illinois, College of Medicine,
Department of Psychiatry at the Illinois Neuropsychiatric Institute,
University of Chicago, Chicago, U.S.A.

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

Yapay Sinir Ağları

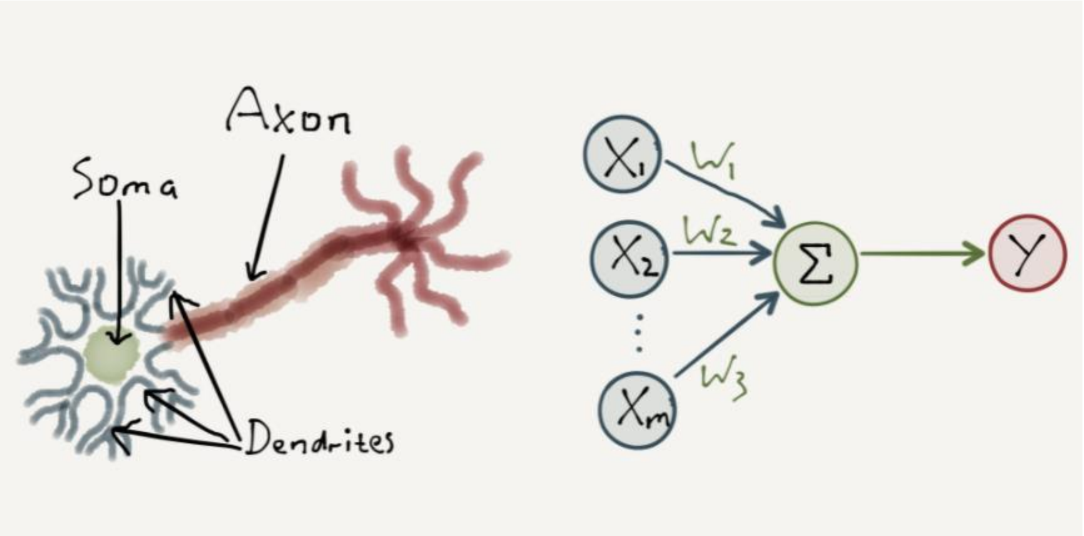
Bulletin of Mathematical Biology Vol. 52, No. 1/2, pp. 99–115, 1990.
Printed in Great Britain.

0092–8240/90$3.00 + 0.00
Pergamon Press plc
Society for Mathematical Biology

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY*

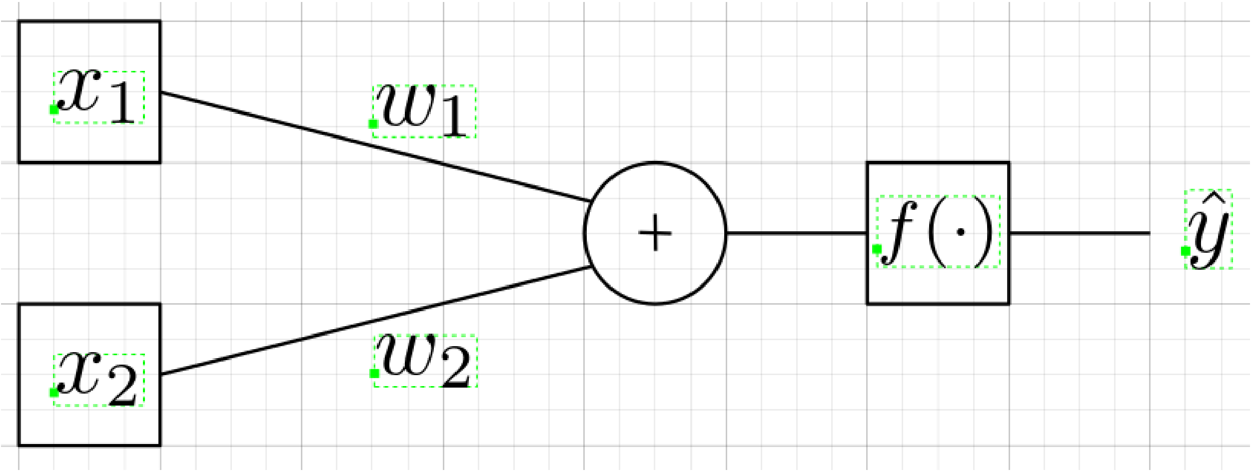
■ WARREN S. MCCULLOCH AND WALTER PITTS
University of Illinois, College of Medicine,
Department of Psychiatry at the Illinois Neuropsychiatric Institute,
University of Chicago, Chicago, U.S.A.

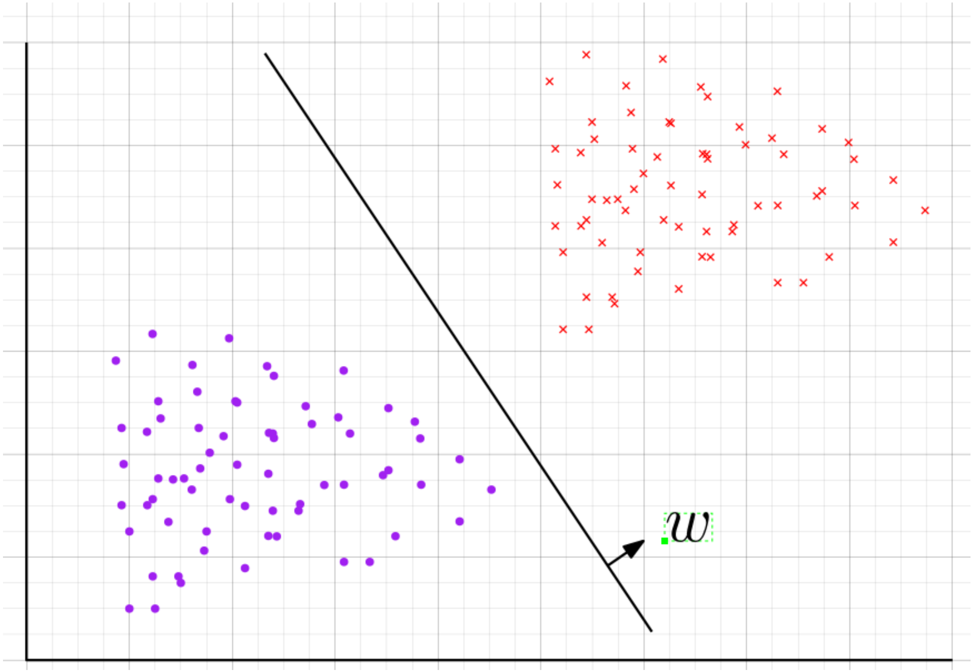
Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.



https://jontysinai.github.io/jekyll/update/2017/11/11/the-perceptron.html

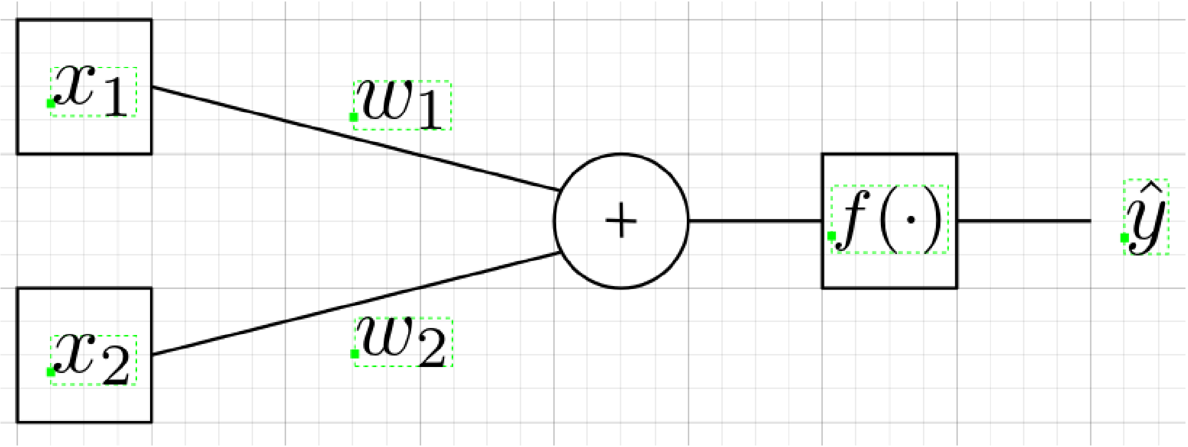
Yapay Sinir Ağları

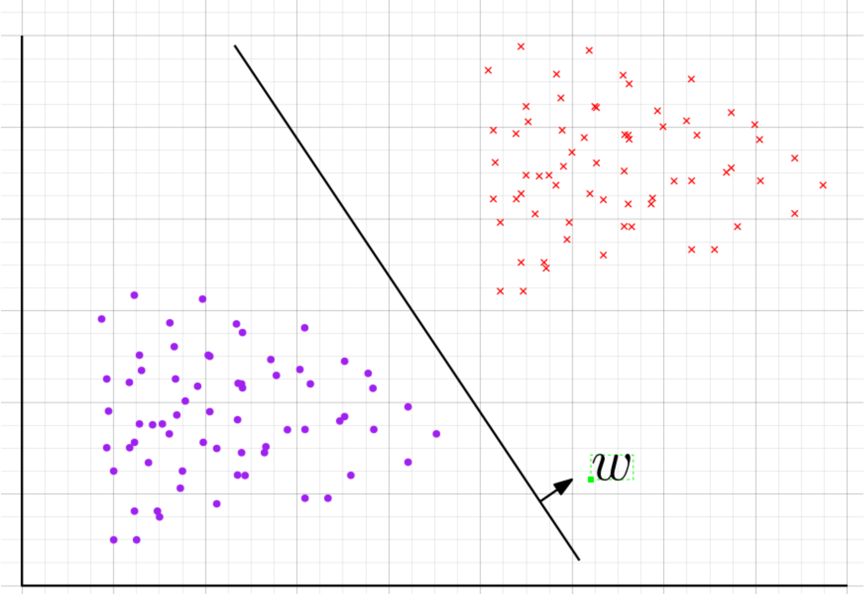




Yapay Sinir Ağları

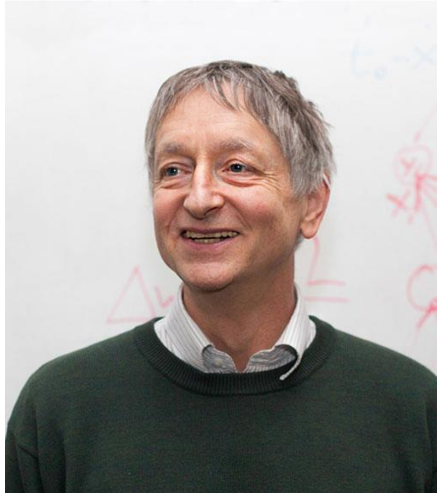
Yapay Sinir Ağları



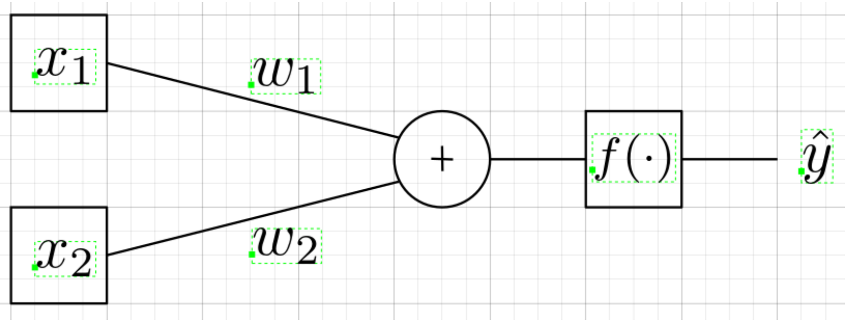


Hatanın Geriye Yayılımı

Geoffrey Hinton



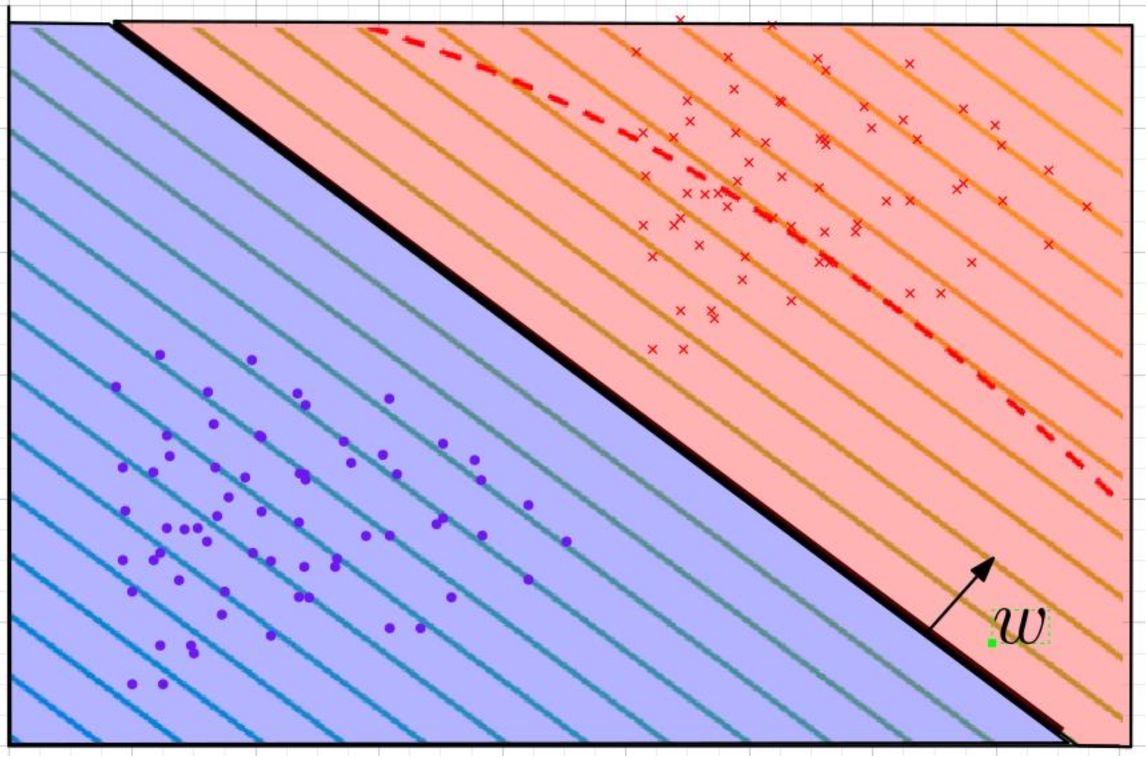
Yapay Sinir Ağları: Hatanın Geriye Yayılması



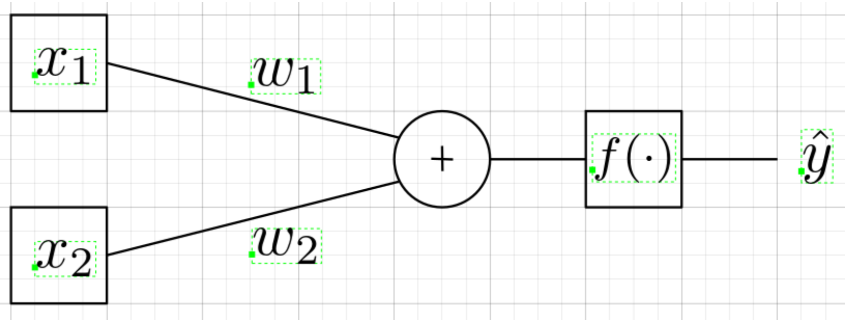
$$\hat{y}_i = f(\,\langle \boldsymbol{w}, \boldsymbol{x_i} \rangle\,)$$

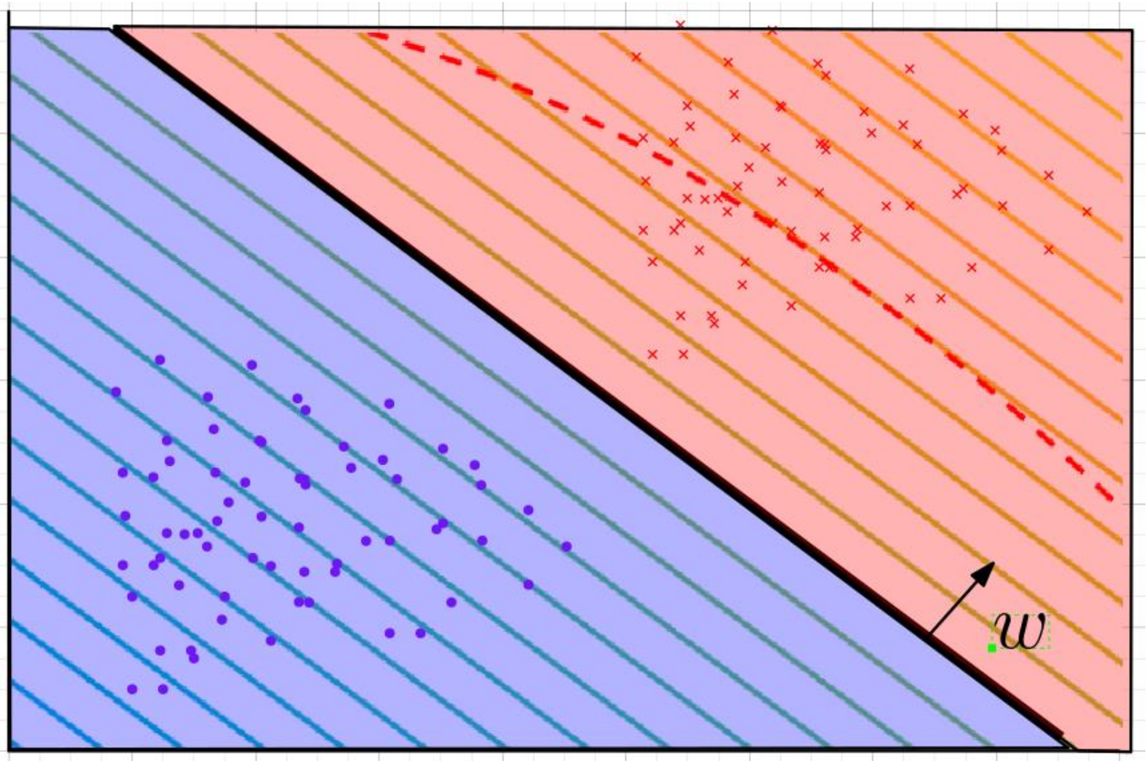
$$e_i = \frac{1}{2}(\hat{y}_i - y_i)^2$$

$$\frac{de_i}{d\boldsymbol{w}} = (\hat{y}_i - y_i)\frac{d\hat{y}_i}{d\boldsymbol{w}}$$



Yapay Sinir Ağları: Hatanın Geriye Yayılması





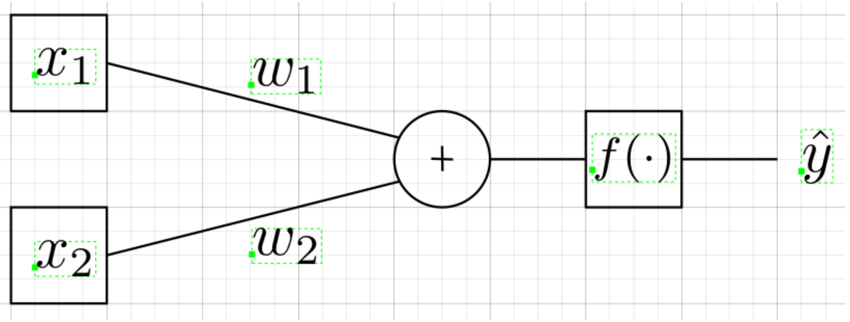
$$\hat{y}_i = f(\langle \boldsymbol{w}, \boldsymbol{x_i} \rangle)$$

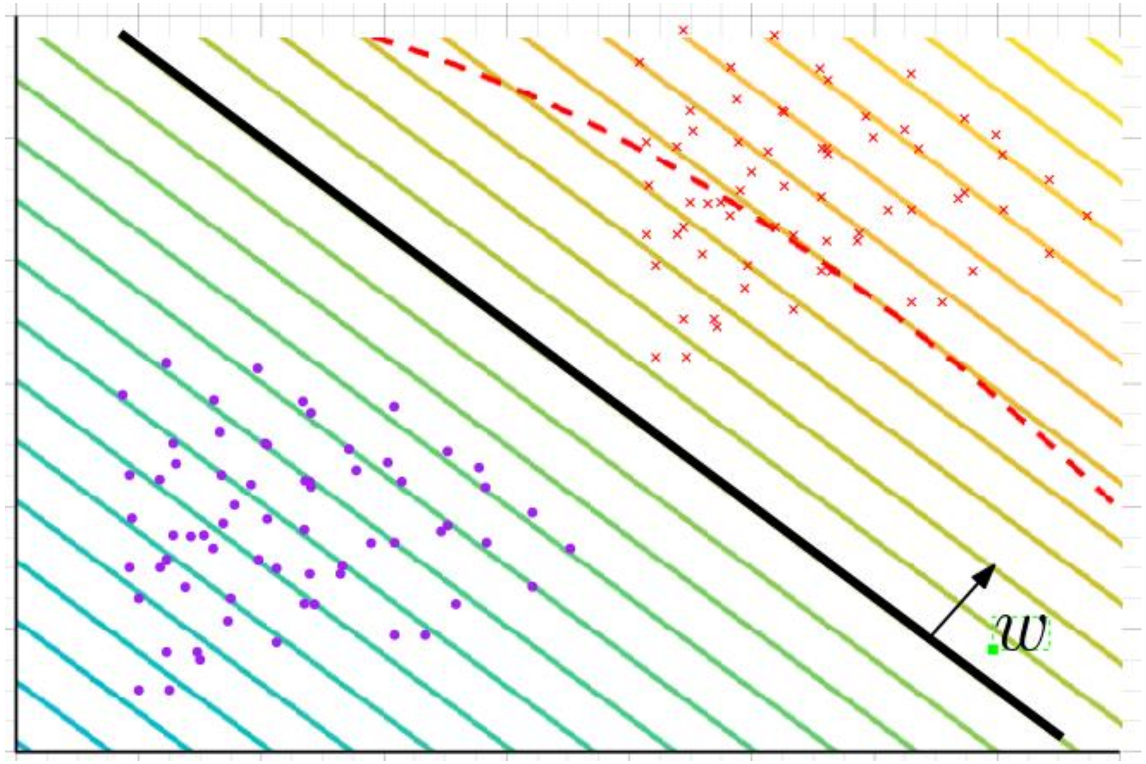
$$e_i = \frac{1}{2}(\hat{y}_i - y_i)^2$$

$$\frac{de_i}{d\boldsymbol{w}} = (\hat{y}_i - y_i)\frac{d\hat{y}_i}{d\boldsymbol{w}}$$

$$\frac{d\,f(x)}{dx}$$

Yapay Sinir Ağları: Hatanın Geriye Yayılması





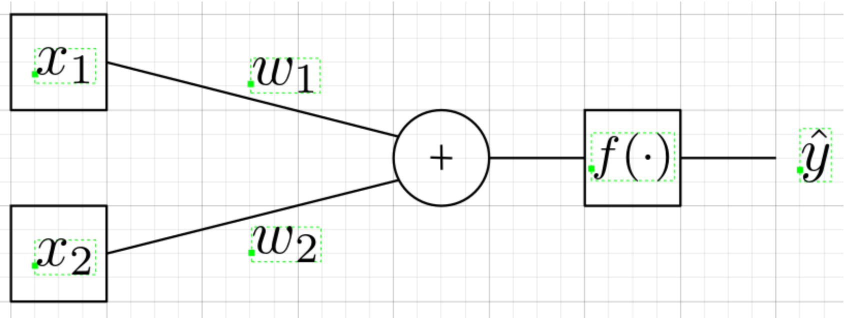
$$\hat{y}_i = f(\langle \boldsymbol{w}, \boldsymbol{x_i} \rangle)$$

$$e_i = \frac{1}{2}(\hat{y}_i - y_i)^2$$

$$\frac{de_i}{d\boldsymbol{w}} = (\hat{y}_i - y_i)\frac{d\hat{y}_i}{d\boldsymbol{w}}$$

$$\frac{d\,f(x)}{dx}$$

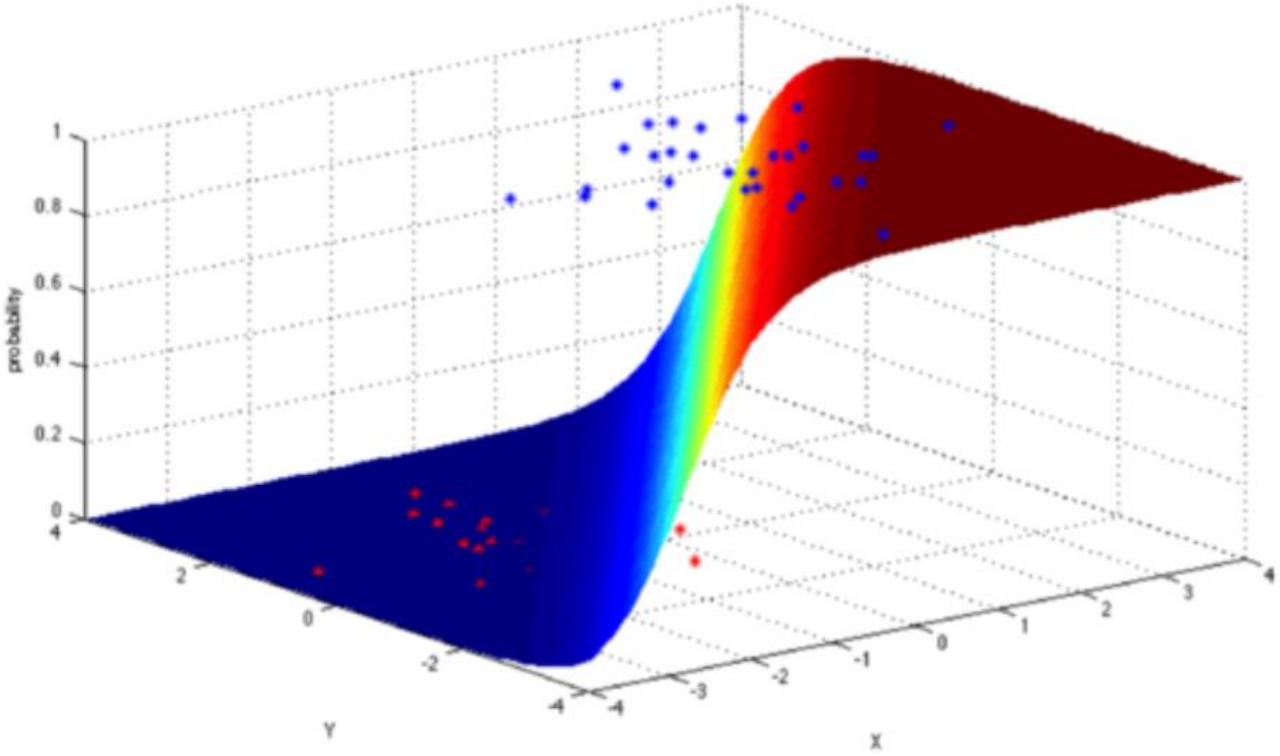
$$f(x) = \frac{1}{1 + e^{-x}}$$

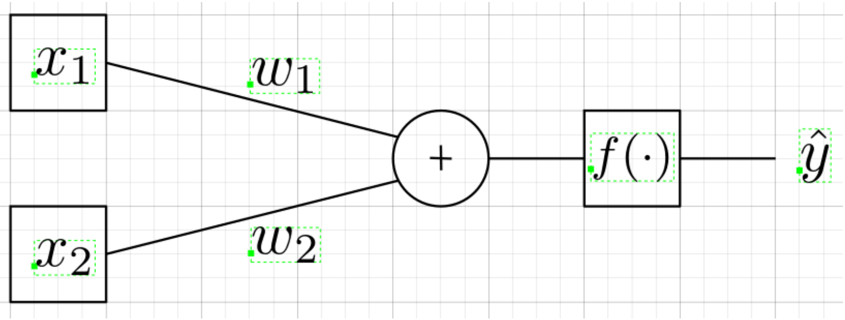


$\mathcal{L}\{\ f(\boldsymbol{x}_i;\boldsymbol{w})\ \ ,y_i\ \}$ $\qquad y_i \in \{0,1\}$

Lojistik Regresyon

$$P\{y=1|\boldsymbol{x}\} = \frac{1}{1+e^{-\langle \boldsymbol{w},\boldsymbol{x}\rangle}}$$





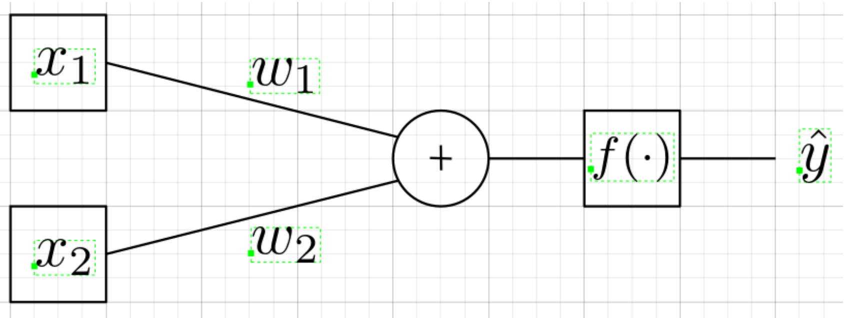
$$\mathcal{L}\{\ f(\boldsymbol{x}_i;\boldsymbol{w})\ \ ,y_i\ \} \qquad y_i \in \{0,1\}$$

Lojistik Regresyon

$$P\{y=1|\boldsymbol{x}\} = \frac{1}{1+e^{-\langle \boldsymbol{w},\boldsymbol{x}\rangle}}$$

$$o_i = \langle \boldsymbol{w}, \boldsymbol{x_i} \rangle \qquad \hat{y}_i = \sigma(\ o_i\ ) \qquad e_i = \frac{1}{2}(\hat{y}_i - y_i)^2$$

$$\frac{de_i}{d\boldsymbol{w}} = \frac{de_i}{d\hat{y}_i}\ \frac{d\hat{y}_i}{do_i}\ \frac{do_i}{d\boldsymbol{w}}$$



Lojistik Regresyon

$$P\{y = 1|\boldsymbol{x}\} = \frac{1}{1 + e^{-\langle \boldsymbol{w},\boldsymbol{x}\rangle}}$$

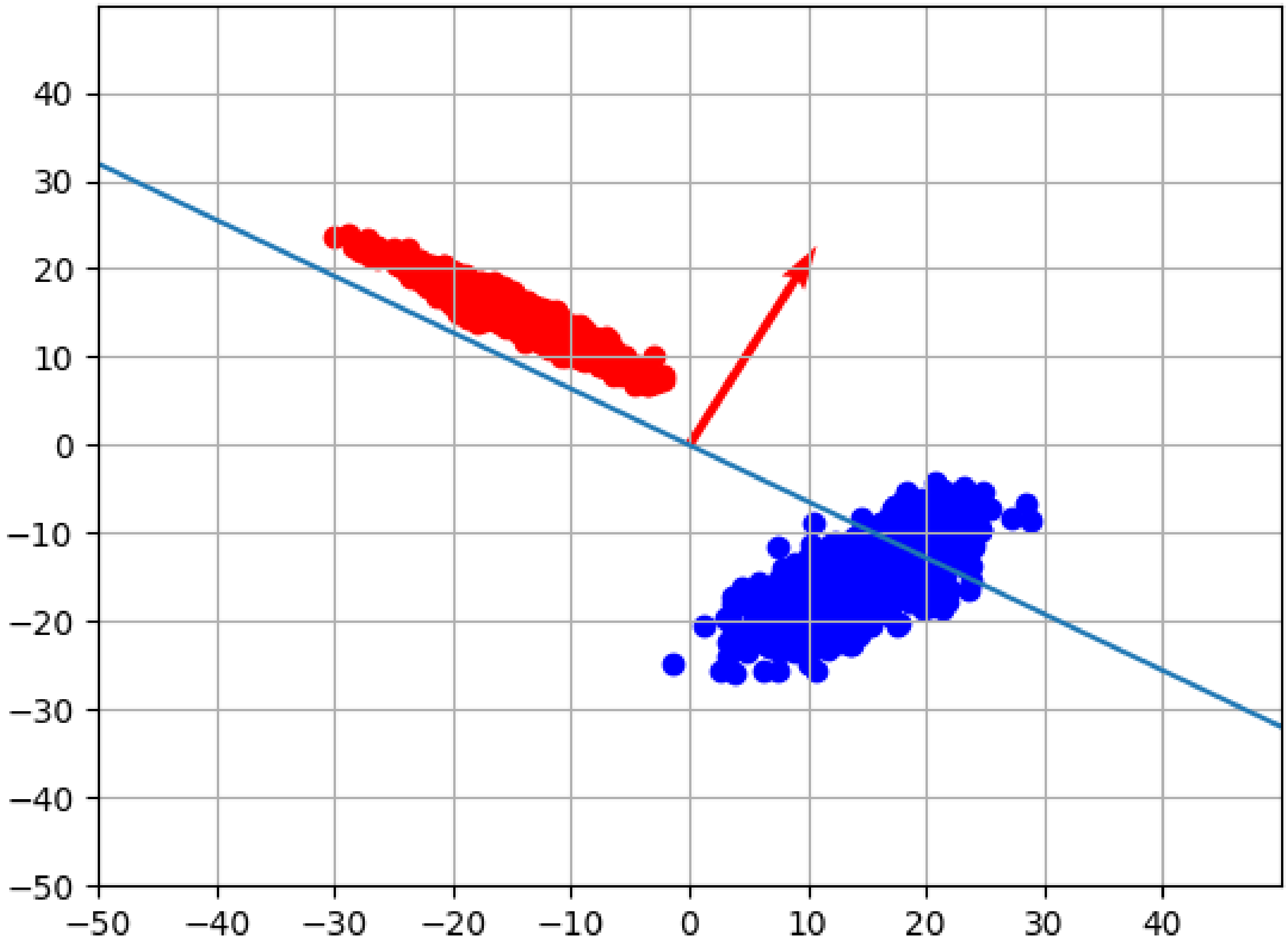
$$\mathcal{L}\{\ f(\boldsymbol{x}_i; \boldsymbol{w})\ \ , y_i\ \} \qquad y_i \in \{0,1\}$$

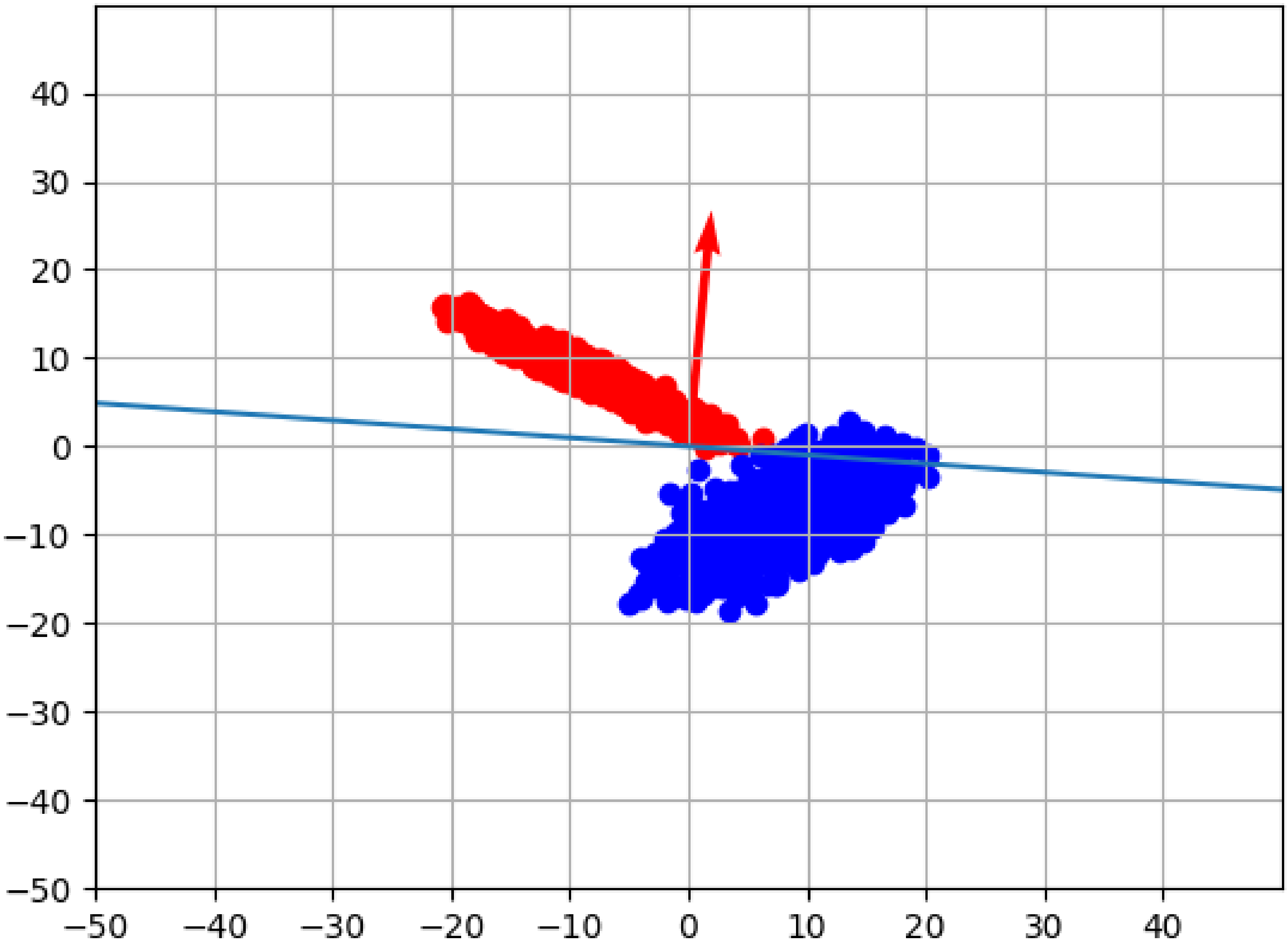
$$o_i = \langle \boldsymbol{w}, \boldsymbol{x_i} \rangle \qquad \hat{y}_i = \sigma(\ o_i\ ) \qquad e_i = \frac{1}{2}(\hat{y}_i - y_i)^2$$

$$\frac{de_i}{d\boldsymbol{w}} = \frac{de_i}{d\hat{y}_i}\ \frac{d\hat{y}_i}{do_i}\ \frac{do_i}{d\boldsymbol{w}}$$

$$\frac{de_i}{d\boldsymbol{w}} = (\hat{y}_i - y_i)\ \hat{y}_i(1 - \hat{y}_i)\ \boldsymbol{x_i}$$

Yapay Sinir Ağları: Hatanın Geriye Yayılması

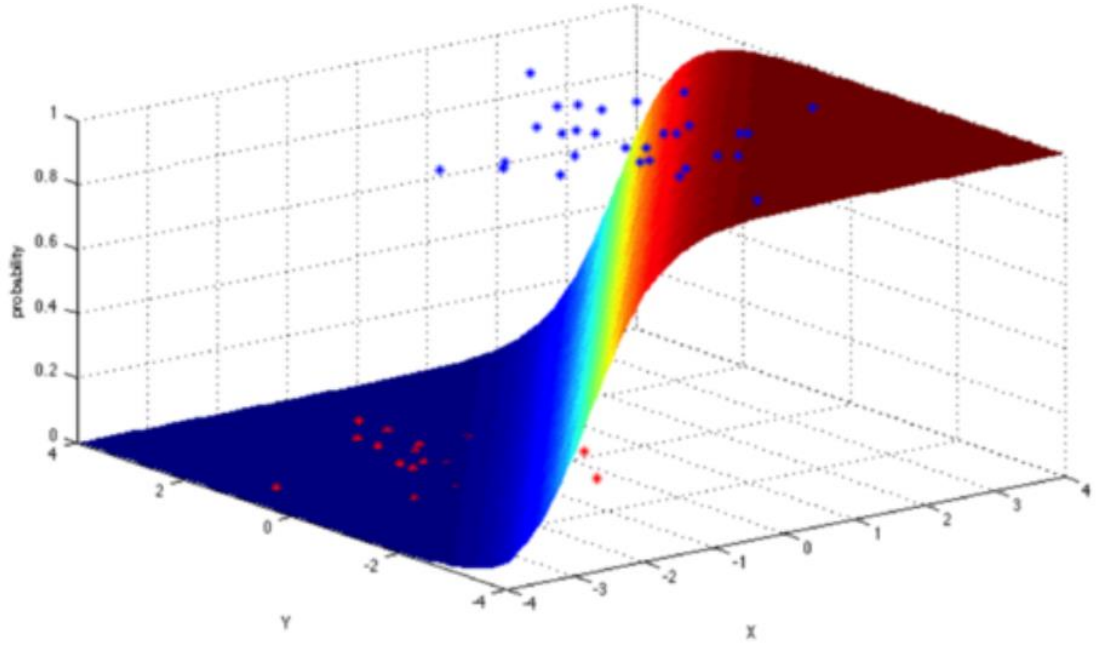


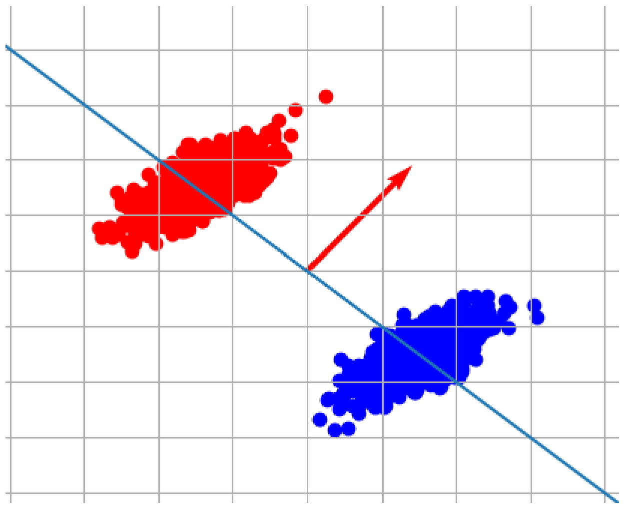


yılması

Gradyanların Sönümlenmesi

$$\frac{de_i}{d\boldsymbol{w}} = (\hat{y}_i - y_i)\hat{\boldsymbol{y}}_i(\boldsymbol{1} - \hat{\boldsymbol{y}}_i)\,\boldsymbol{x}_i$$

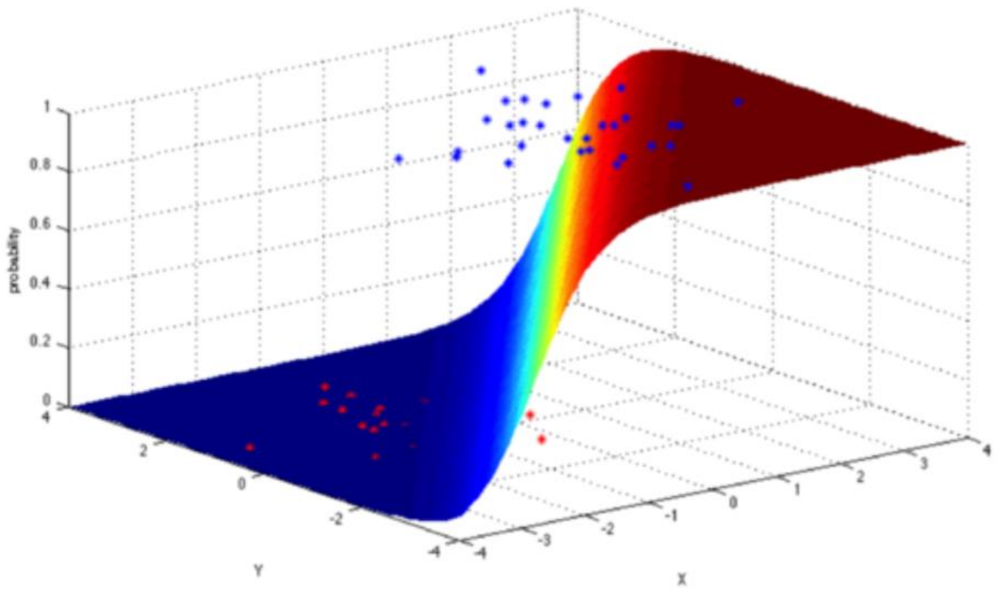




https://towardsdatascience.com/nothing-but-numpy-understanding-creating-binary-classification-neural-networks-with-e746423c8d5c

Gradyanların Sönümlenmesi

$$\frac{de_i}{d\boldsymbol{w}} = (p_i - y_i)\, \boldsymbol{p_i(1 - p_i)}\, \boldsymbol{x_i}$$

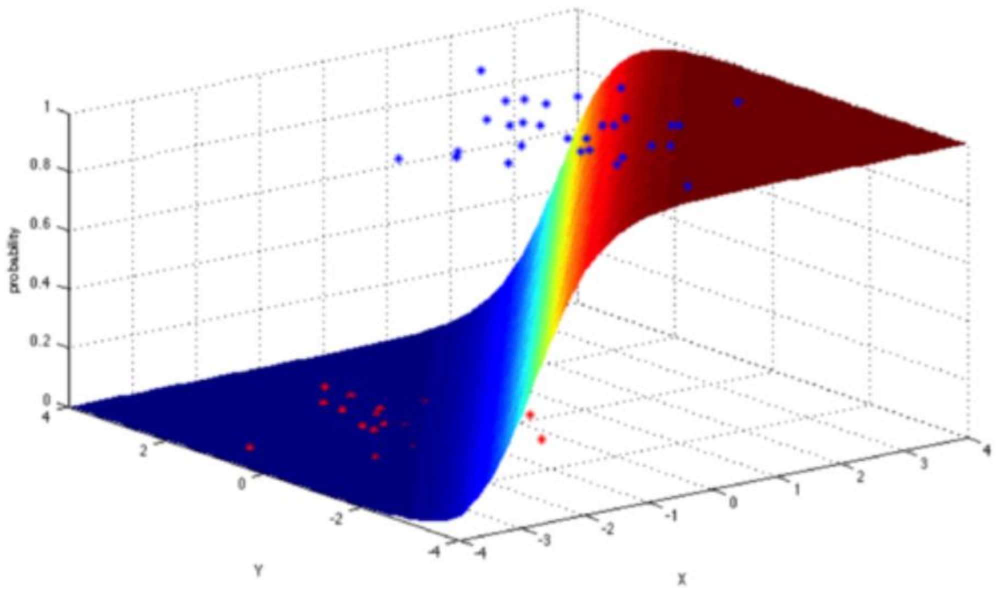


Cross Entropy Loss

$$\mathcal{L}_i(p_i, y_i) = -(y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

Gradyanların Sönümlenmesi

$$\frac{de_i}{d\boldsymbol{w}} = (p_i - y_i)\, \boldsymbol{p_i(1 - p_i)}\, \boldsymbol{x_i}$$



Cross Entropy Loss

$$\mathcal{L}_i(p_i, y_i) = -(y_i \log(p_i) + (1 - y_i)\log(1 - p_i))$$

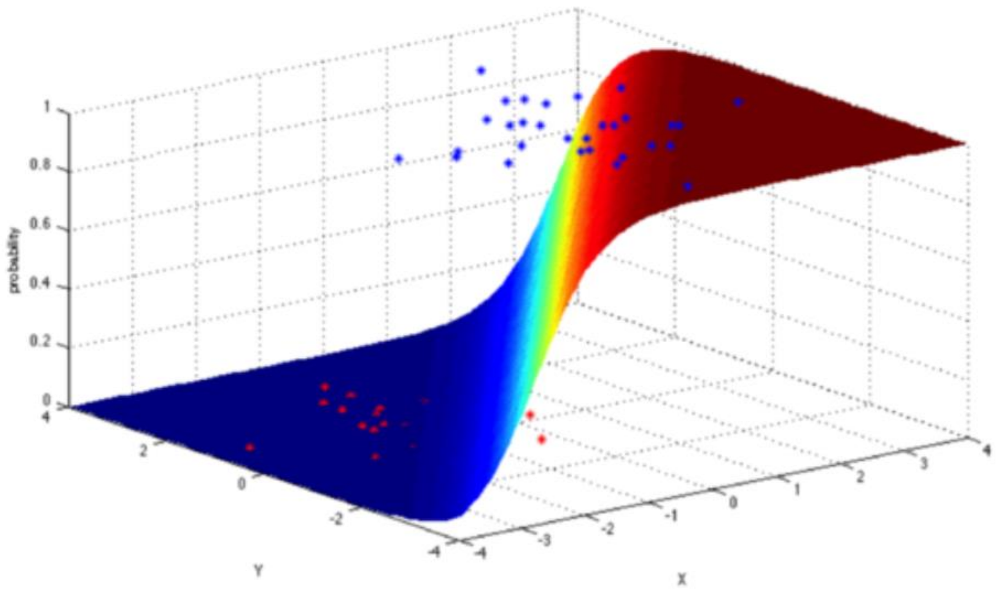
$$\frac{d\mathcal{L}_i(p_i, y_i)}{dp_i} = -\left(\frac{y_i}{p_i} - \frac{1 - y_i}{1 - p_i}\right) = \left(-\frac{y_i}{p_i} + \frac{1 - y_i}{1 - p_i}\right)$$

Gradyanların Sönümlenmesi

Cross Entropy Loss

$$\frac{de_i}{d\boldsymbol{w}} = (p_i - y_i)\,\boldsymbol{p_i(1 - p_i)}\,\boldsymbol{x_i}$$

$$\mathcal{L}_i(p_i, y_i) = -(y_i \log(p_i) + (1 - y_i)\log(1 - p_i))$$



$$\frac{d\mathcal{L}_i(p_i, y_i)}{dp_i} = -\left(\frac{y_i}{p_i} - \frac{1 - y_i}{1 - p_i}\right) = \left(-\frac{y_i}{p_i} + \frac{1 - y_i}{1 - p_i}\right)$$

$$\frac{de_i}{d\boldsymbol{w}} = \boldsymbol{(p_i - y_i)}\,\boldsymbol{p_i(1 - p_i)}\,\boldsymbol{x_i}$$

$$\frac{de_i}{do_i} = \begin{cases} -(1 - p_i) & y_i = 1 \\ p_i & y_i = 0 \end{cases}$$

1 Nöron YSA Özeti

- N boyutlu uzayda bir üst düzlem
- Aktivasyon fonksiyonu doğrusal olmamalıdır
- GD ile eğitim yapılır
- Sınıflandırma için Cross Entropy Loss kullanılır.