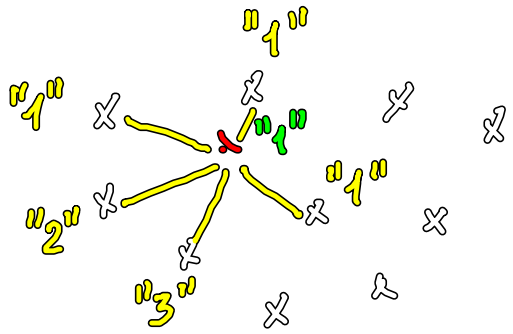
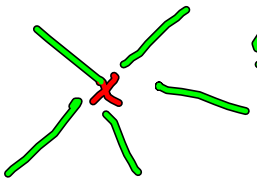
Klassifikatoren

K - NN Klassifiziert



55,000 Ziffern

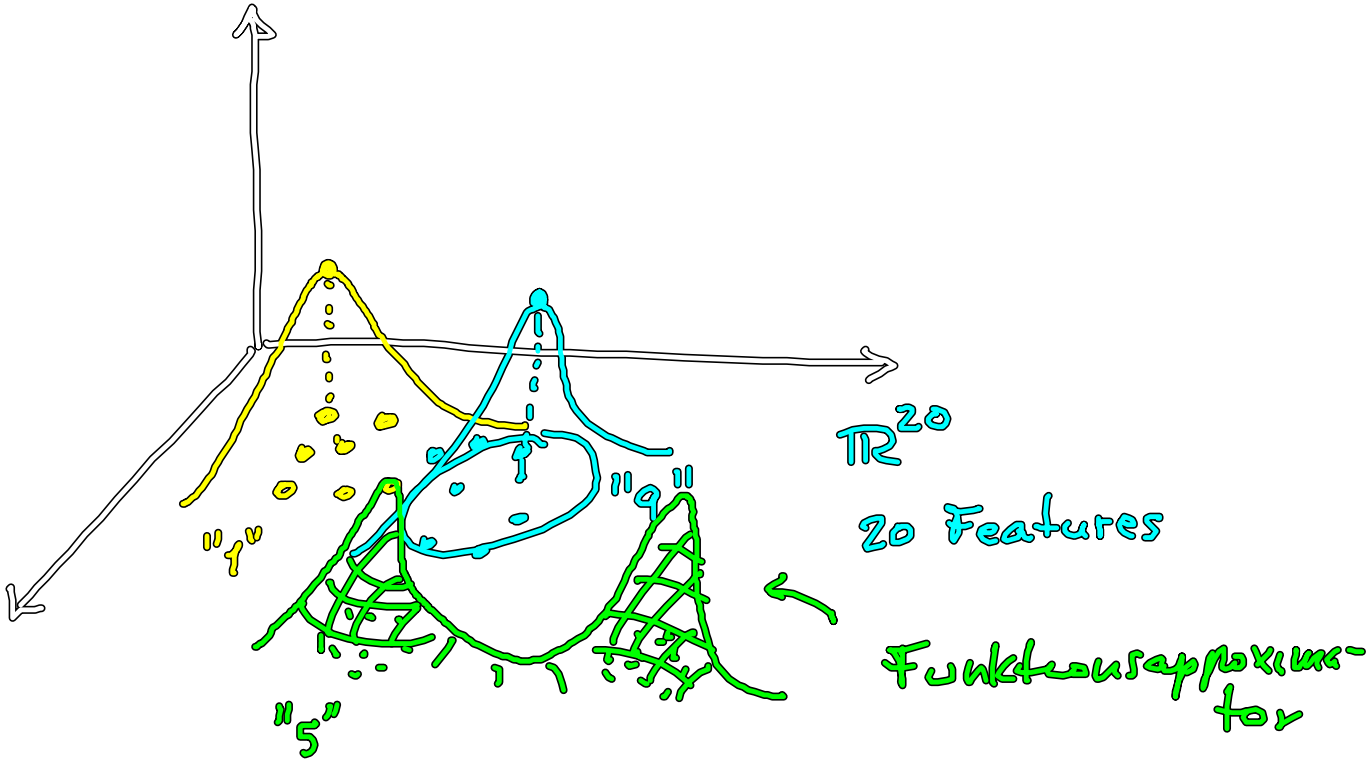


55,000

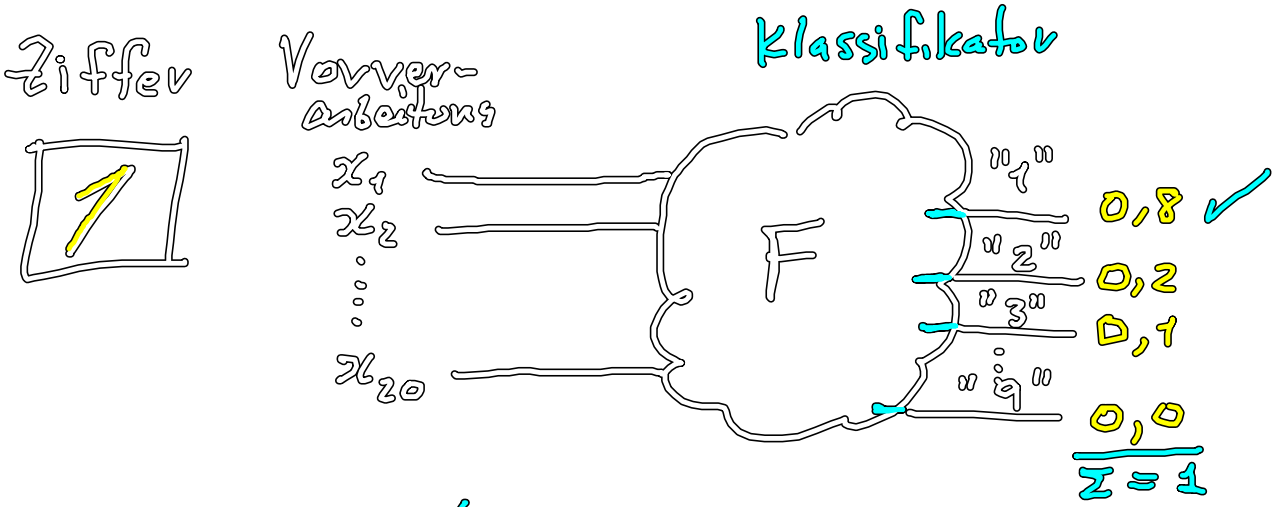
$$\sqrt{(x_1 - y_1)^2 + \cdots (x_n - y_n)^2}$$

55,000 x 20 ≈ 10^6

100,000 Ziffern



Allgemeiner Klassifikator

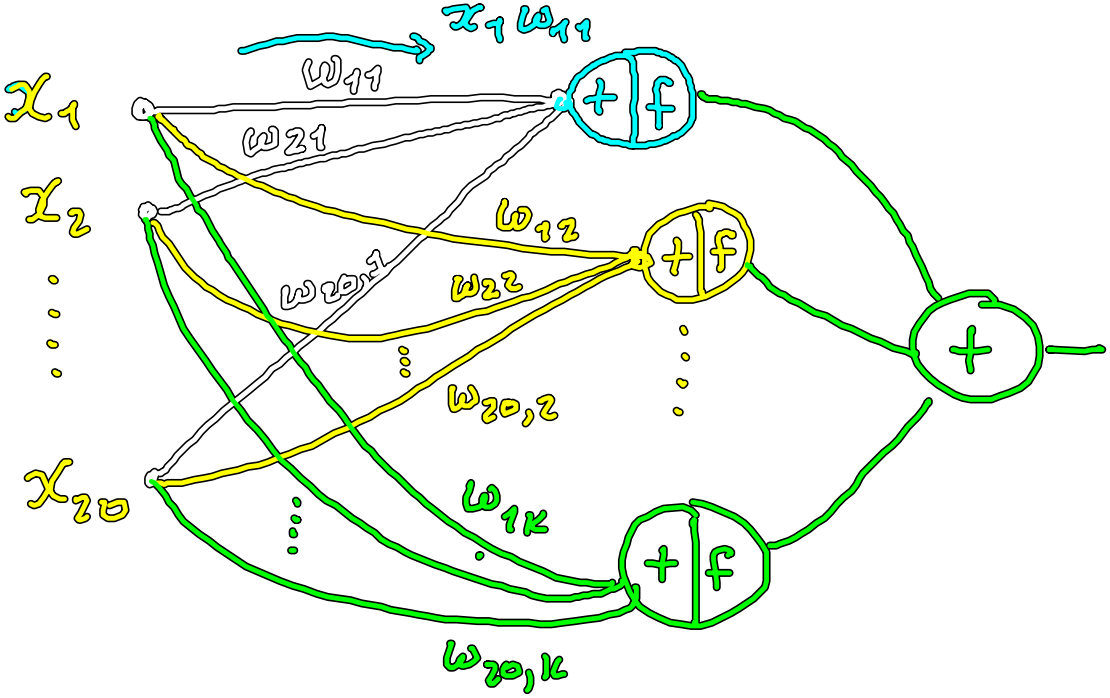


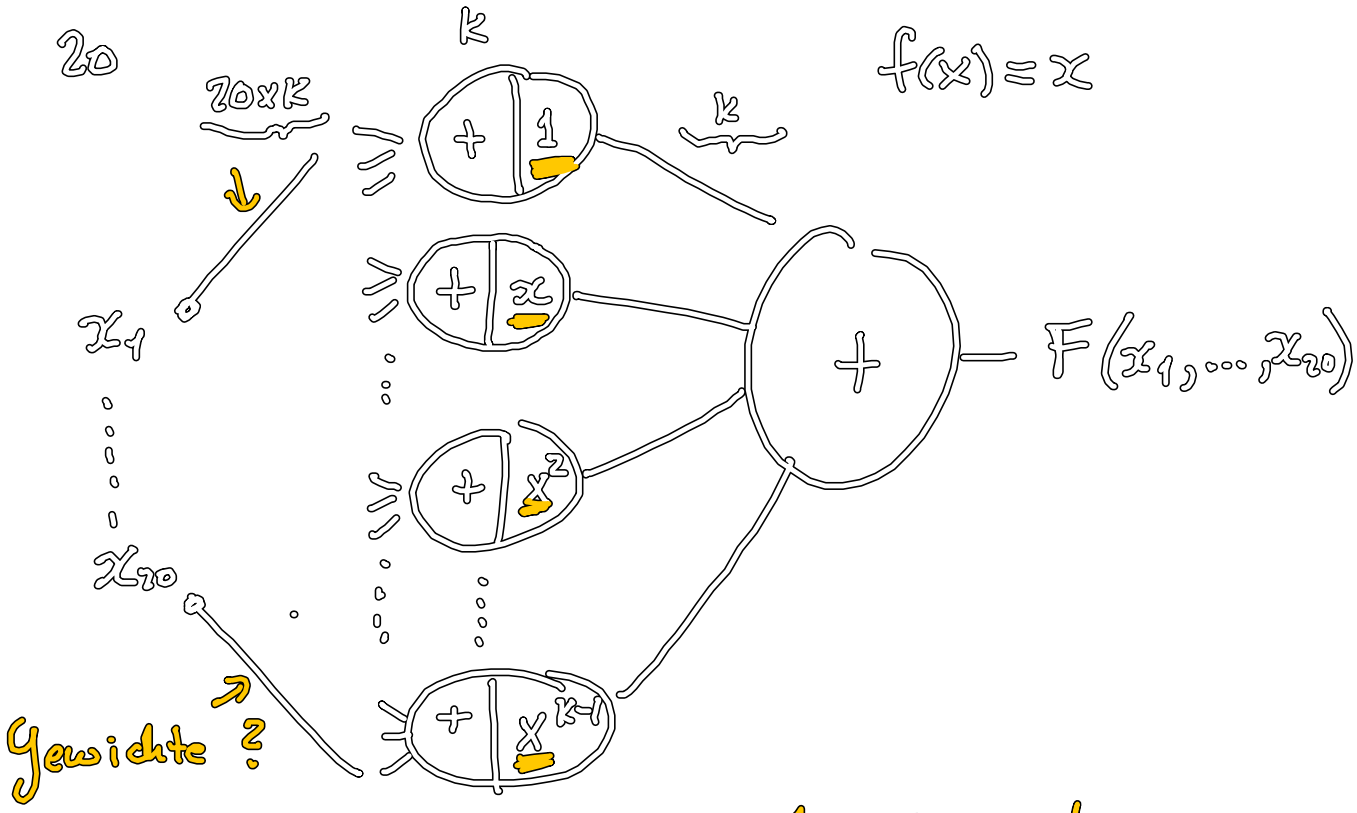
$F_1(x_1, x_2, \dots x_{20}) = 0,8$

$F_2(x_1, x_2, \dots x_{20}) = 0,2$

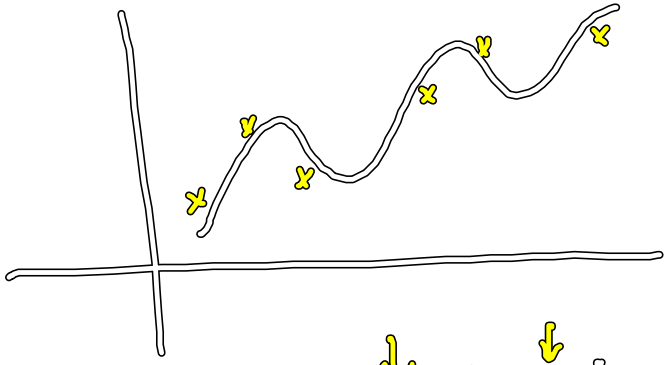
$\vdots$

Funktionennetz = Neuronales Netz

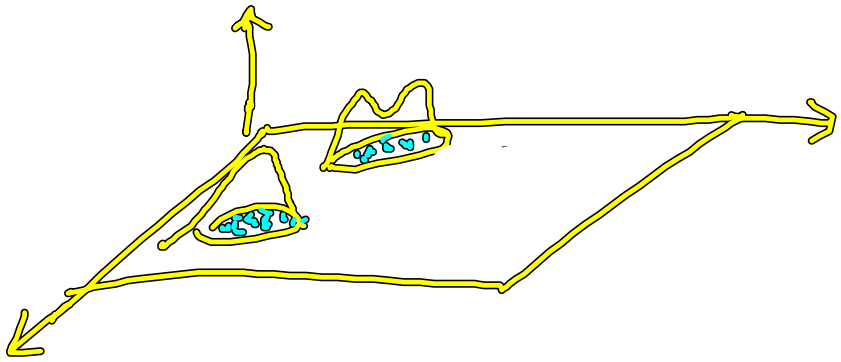


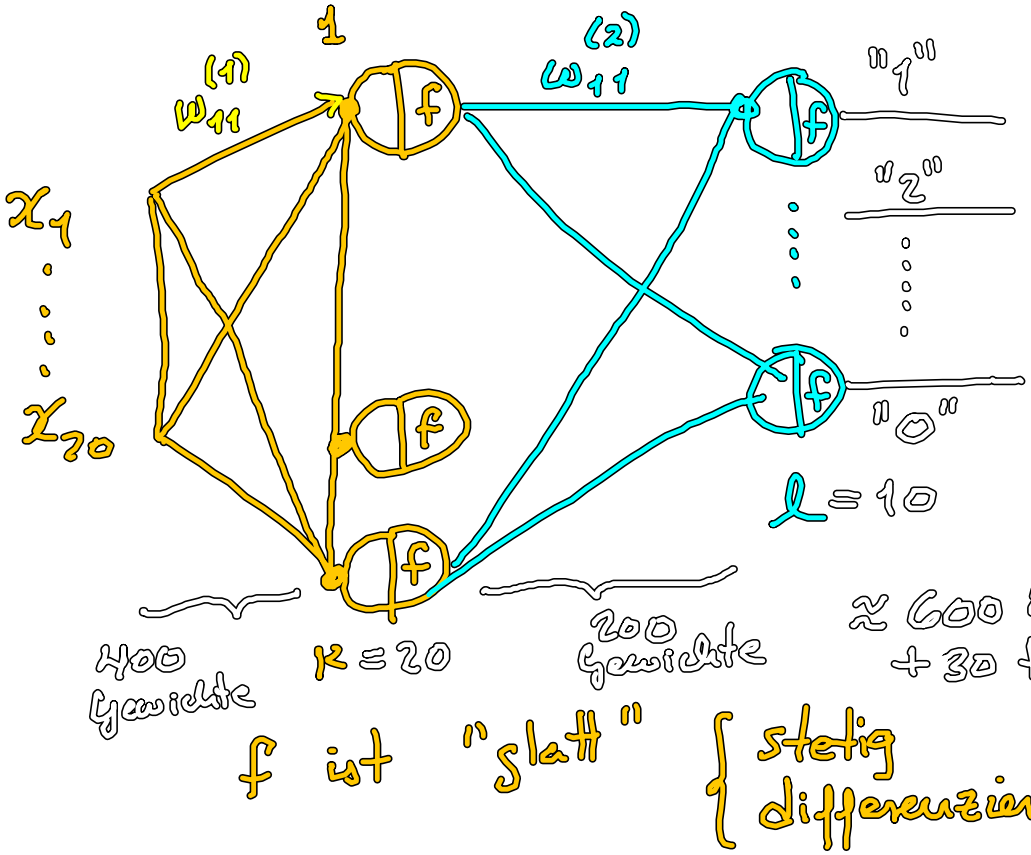


$$F(x_1, \dots, x_{20}) = a_0 + a_1 x_1 + a_2 x_2 \dots$$
$$\dots + a_{20} x_{20} + a_{21} x_1^2 + \dots$$
$$\dots + a_{40} x_{20}^2 + \dots \; x_1 x_2 \dots$$
$$\dots \dots x_1^{K-1} + \dots$$



$$\downarrow 1x^2 + \downarrow 12x^3 + \downarrow 4x^4 \ldots$$





f ist "glatt" { stetig, differenzierbar

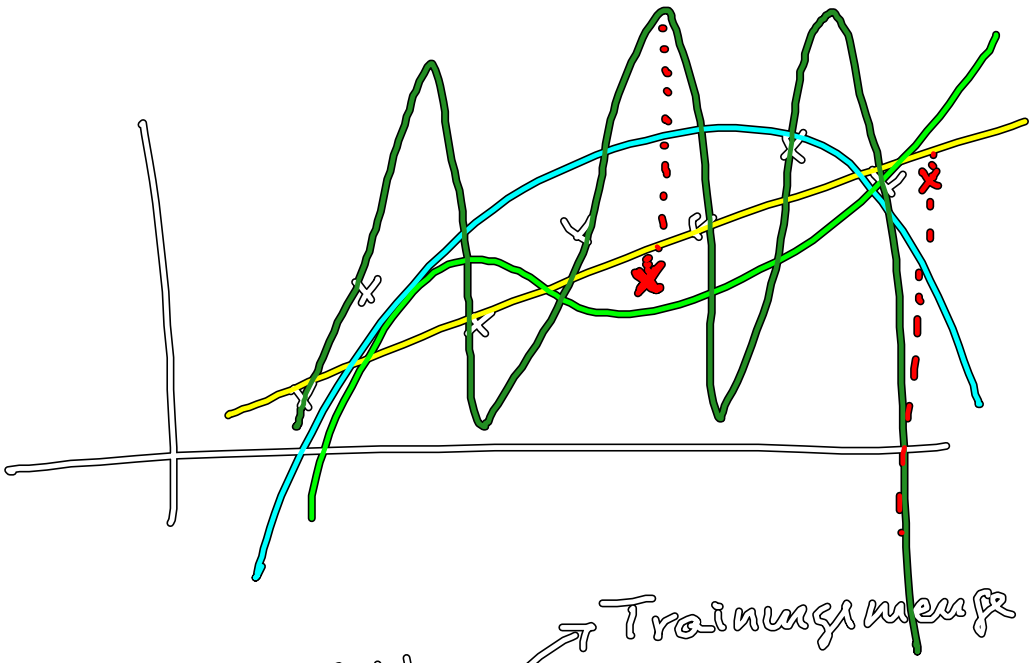
10^6 100,000 differ

vs. 600

K-NN vs. NN

NN schneller

genauer genau



Fehler → Trainingsmenge

Fehler → Recall (neue Eingaben)

Bestes Modell + Einfachstes

Trainingsmenge

100,000 Ziffern

↓

Merkmale

↓

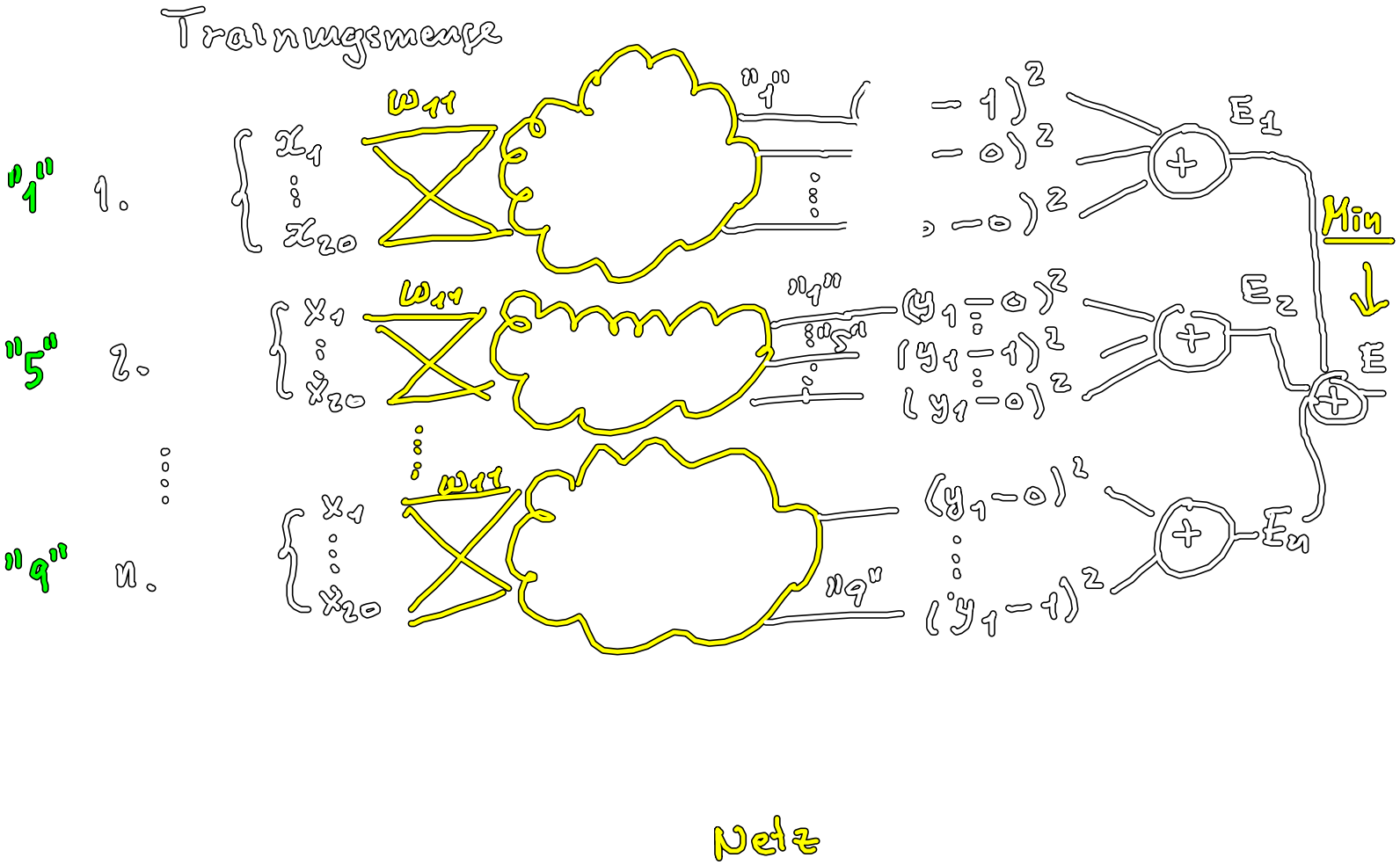
Gewichte

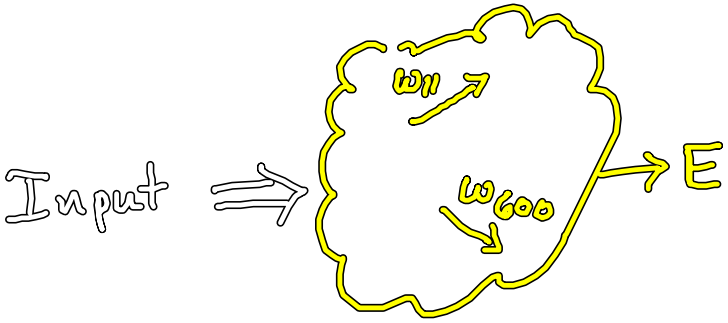
↓

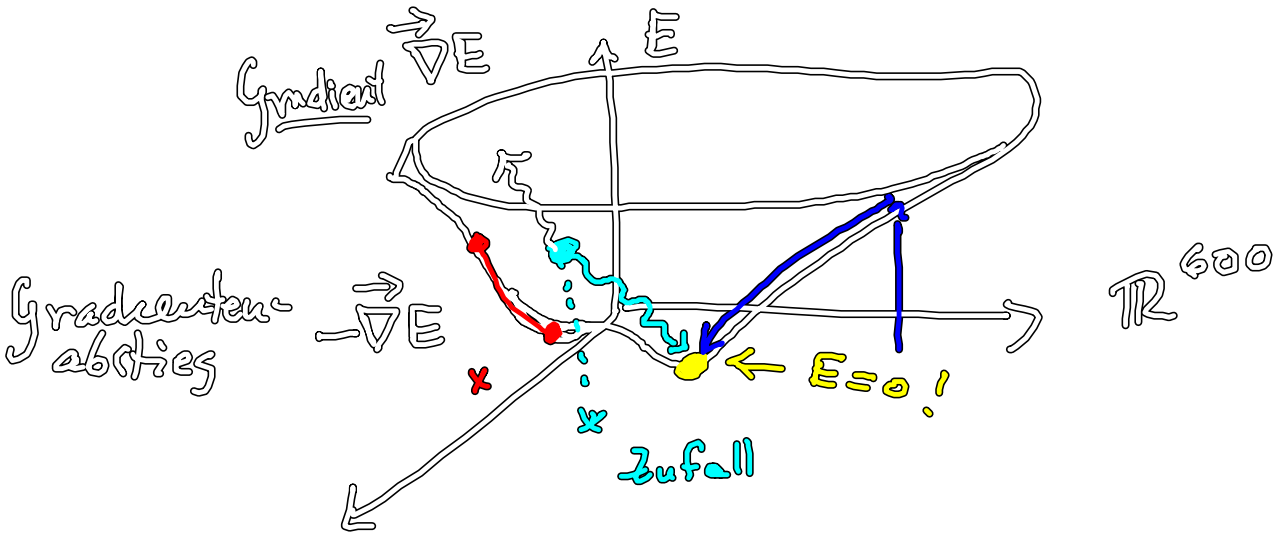
Recall

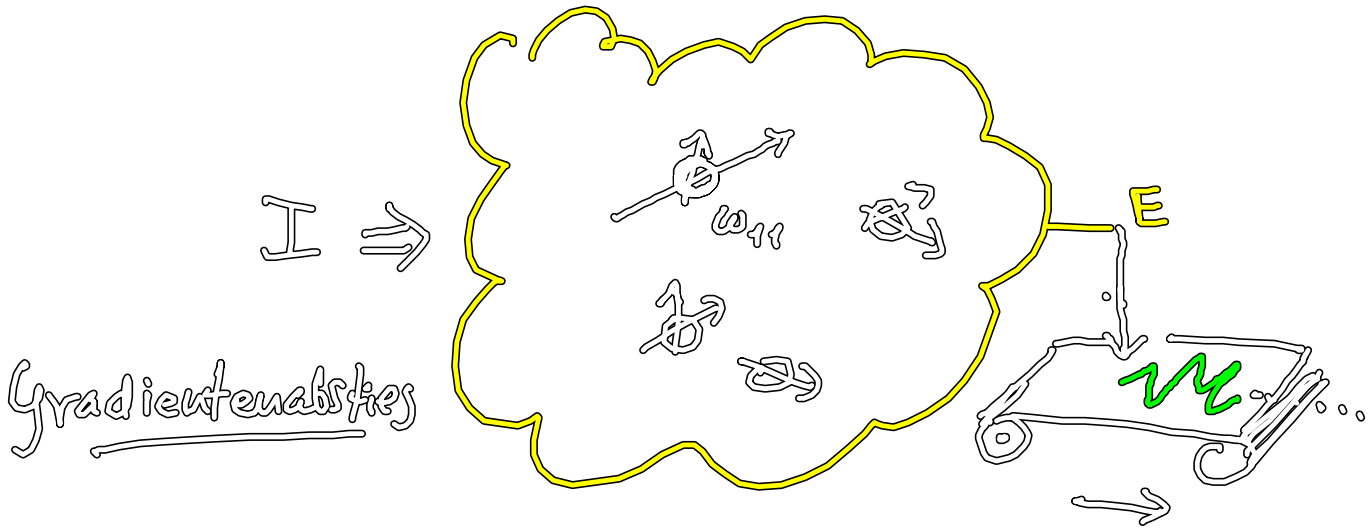
- Zufällige Gewichte
- testen
- Fehler der Trainingsmenge
- Fehler > ε
- korrigieren

(bis Fehler < ε)





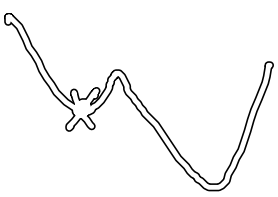




$$\frac{\partial E}{\partial w_{11}} \dots$$

$\Delta w_{11} \uparrow \quad \Delta E \uparrow$

$\Delta w_{11} \downarrow \quad \Delta E \downarrow$



$\frac{\partial E}{\partial w_{11}}$ .... $\frac{dE}{dw_{11}}$

$E(w_{11}, w_{12}, \dots w_{600})$ $\quad$ $E(w_{11}, \overbrace{w_{12} \dots w_{600}}^{\text{constant}})$

Backpropagation

Systematisch

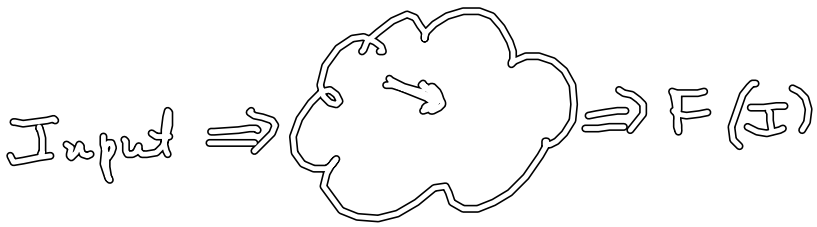
Vorwärtsschritt

$x \xrightarrow{w} wx$

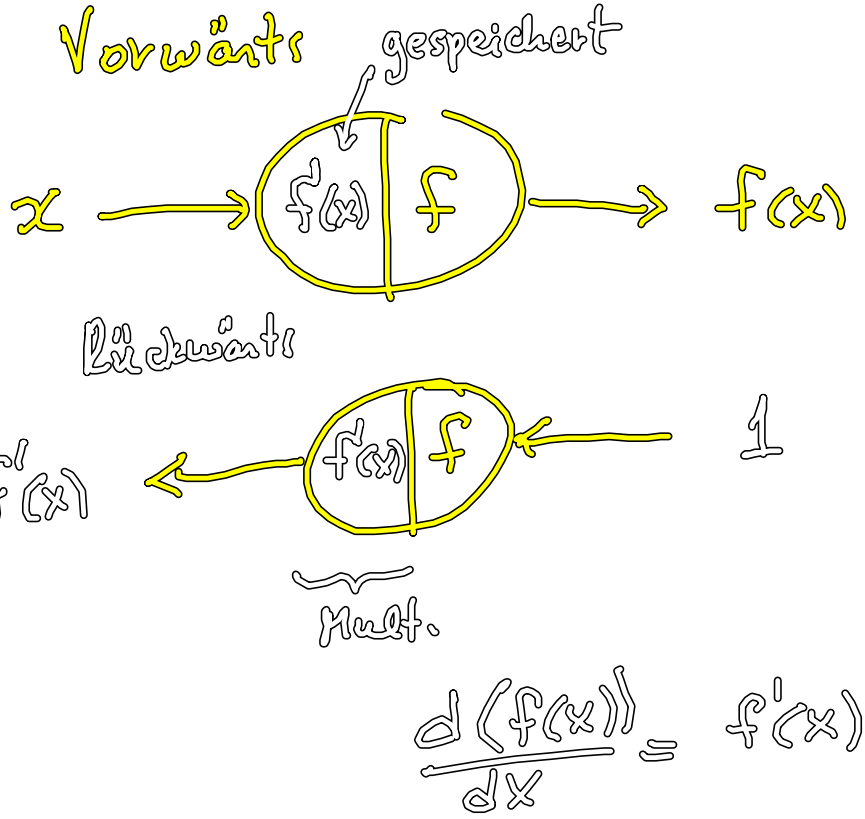
Backprop-Schritt

$w \xleftarrow{w} 1$ $\quad$ Traversierungsweg

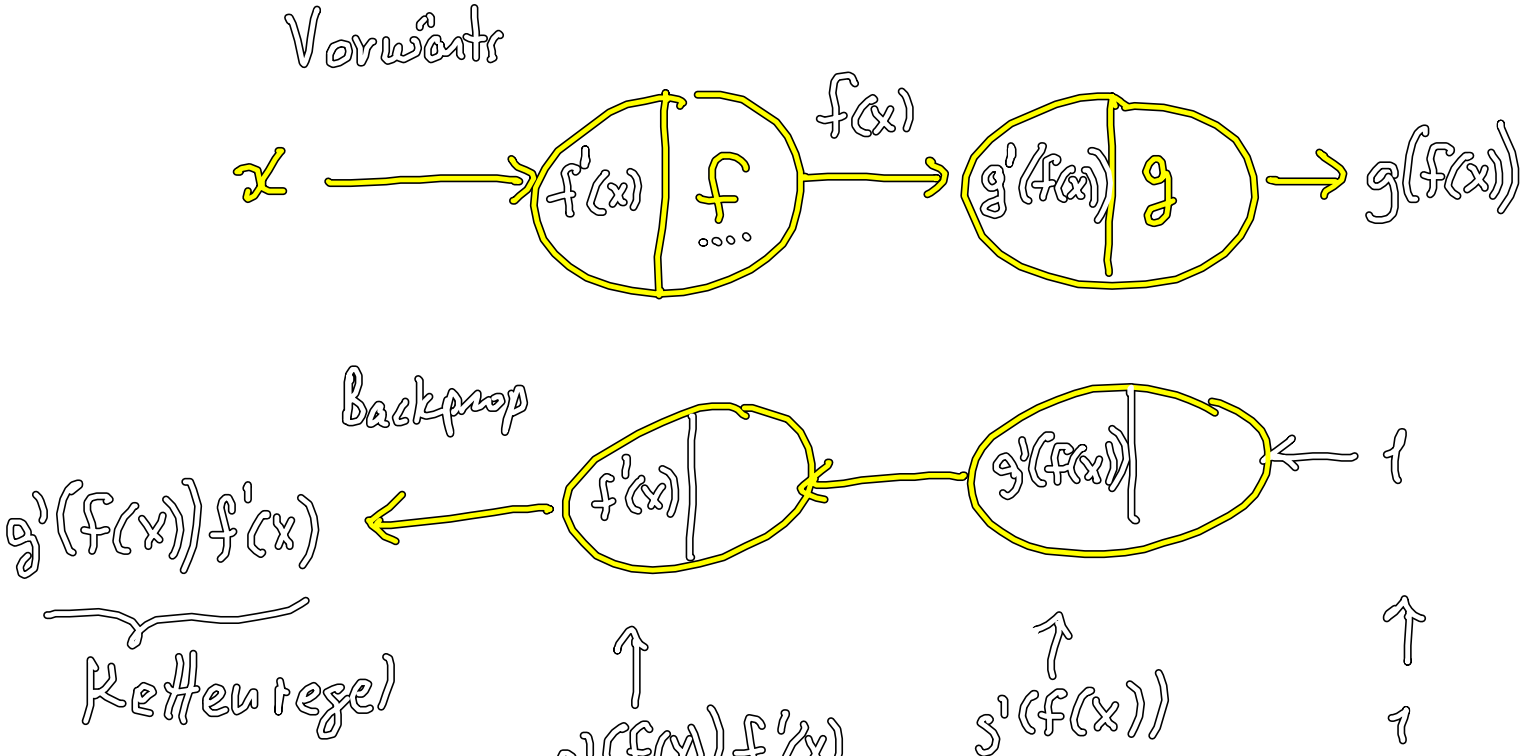
$\frac{d}{dx}(wx) = w$

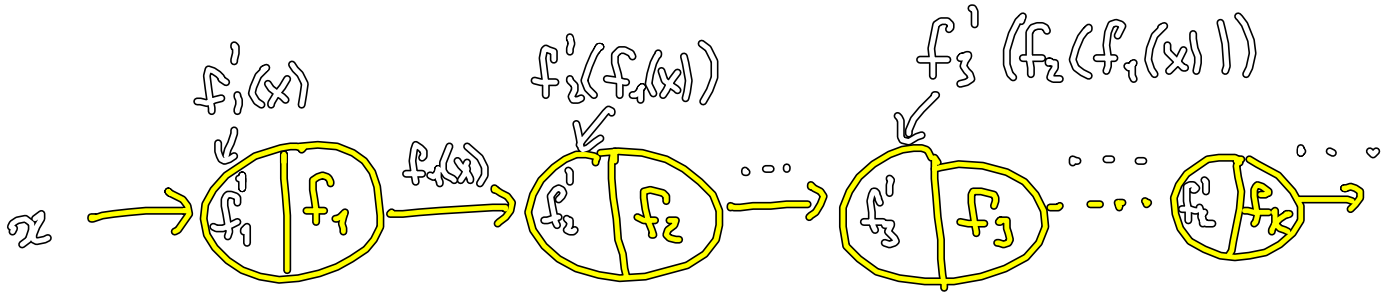


$\frac{dF}{dI}$

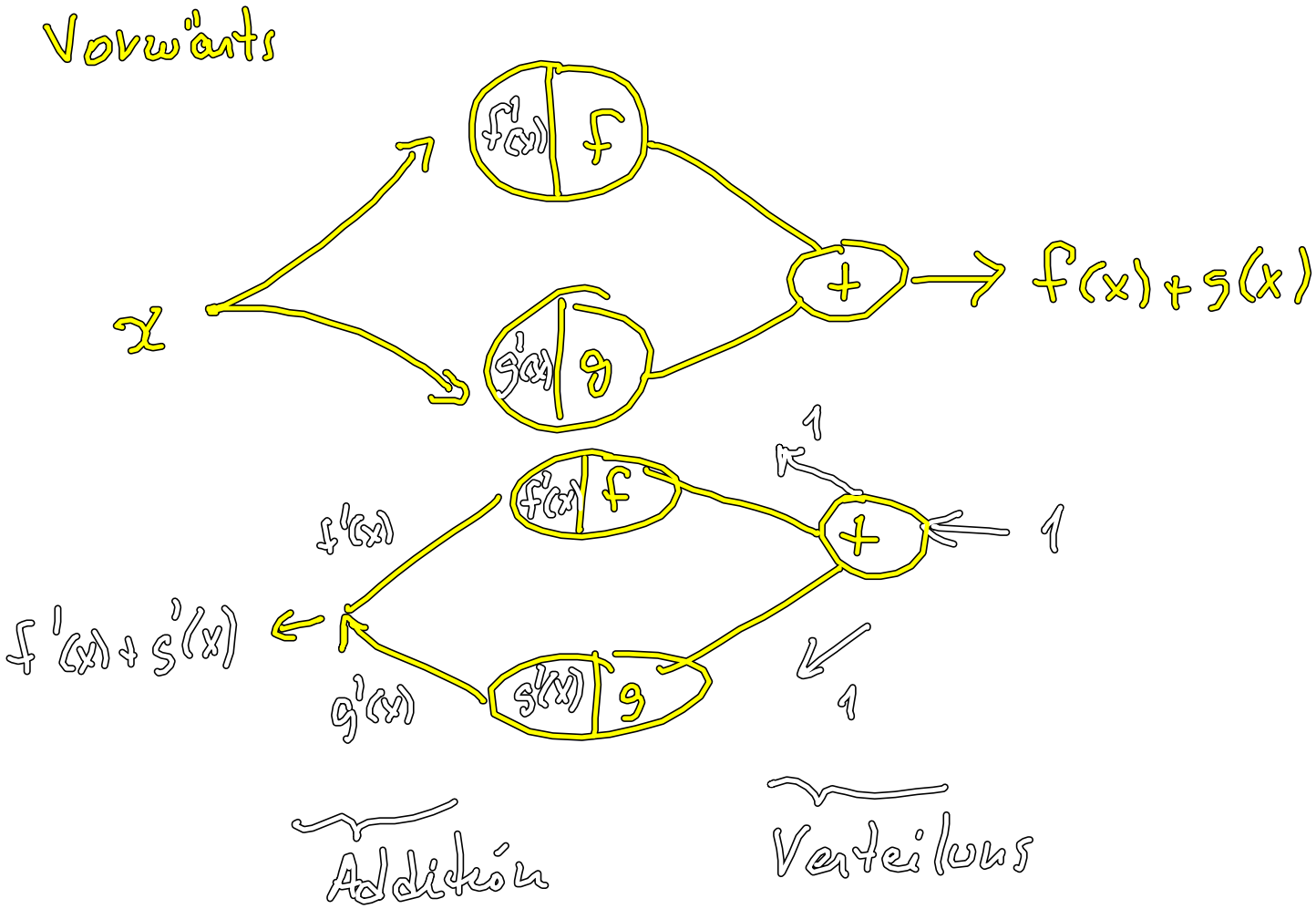


$$\frac{d(f(x))}{dx} = f'(x)$$



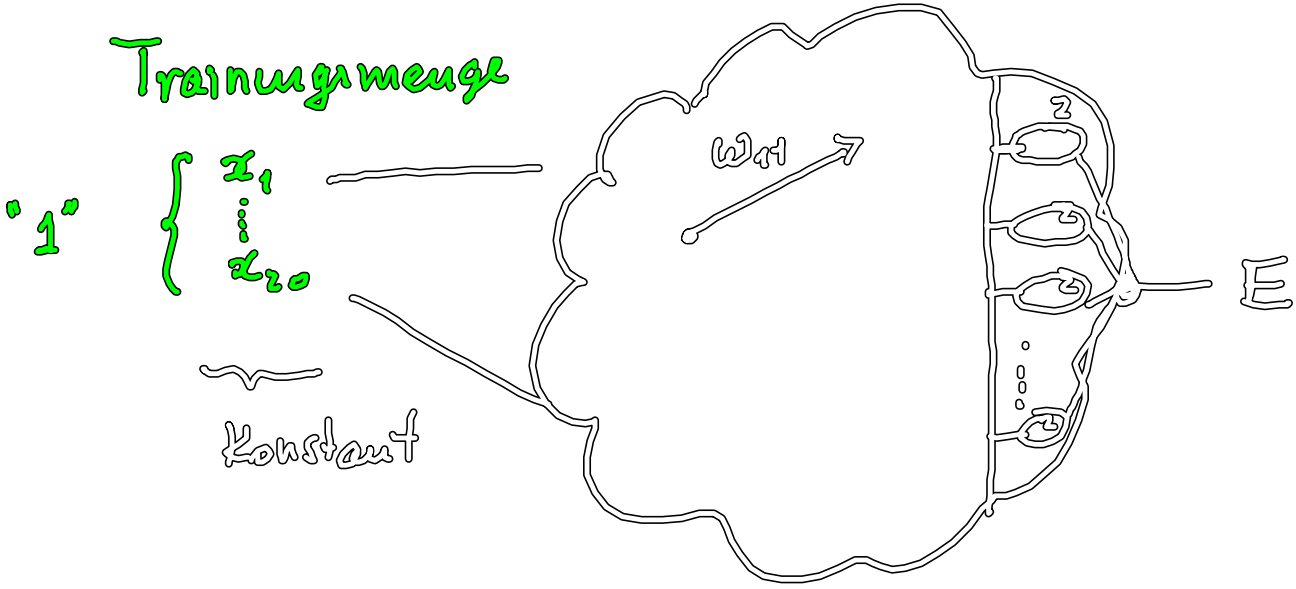


per Induktion beweisen....



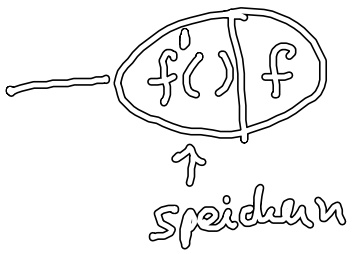
$$\frac{d}{dx}\left(f(x)+g(x)\right) = f'(x)+g'(x)$$

Jetzt an den Zweigen..

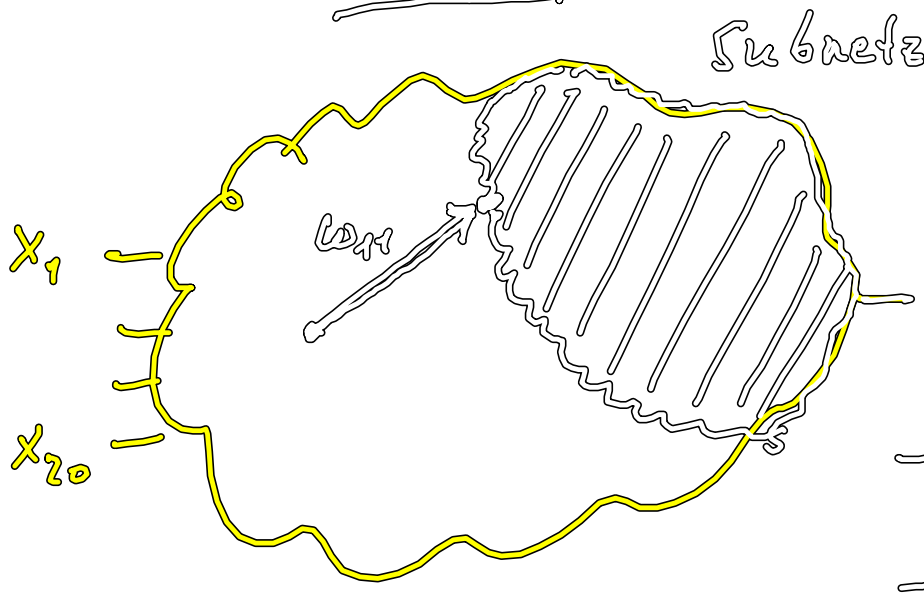


$$\vec{\nabla} E = \left(\frac{\partial E}{\partial w_{11}}, \frac{\partial E}{\partial w_{12}}, \dots \frac{\partial E}{\partial w_{\dots}} \right)$$

1. Schritt: Vorwärtslauf



2. Schritt: Backpropagation



- benutze gewichte
- multiplizieren Traversierung wert mit linken Seiten
- Verteilung
- Addition der Pfade

- keine Zyklen
- meins hat keine

Ergebnis = $\frac{\partial E}{\partial w_{11}}$

$$\vec{\nabla} E = \left(\frac{\partial E}{\partial w_{11}}, \frac{\partial E}{\partial w_{12}}, \frac{\partial E}{\partial w_{13}} \dots \frac{\partial E}{\partial w} \dots \right)$$

für alle Gewichte

3. Korrektur:

Konstante (Schrittweite)

$$w_{11} := w_{11} - \delta \frac{\partial E}{\partial w_{11}}$$

$$w_{12} := w_{12} - \delta \frac{\partial E}{\partial w_{12}}$$

$$\vdots$$

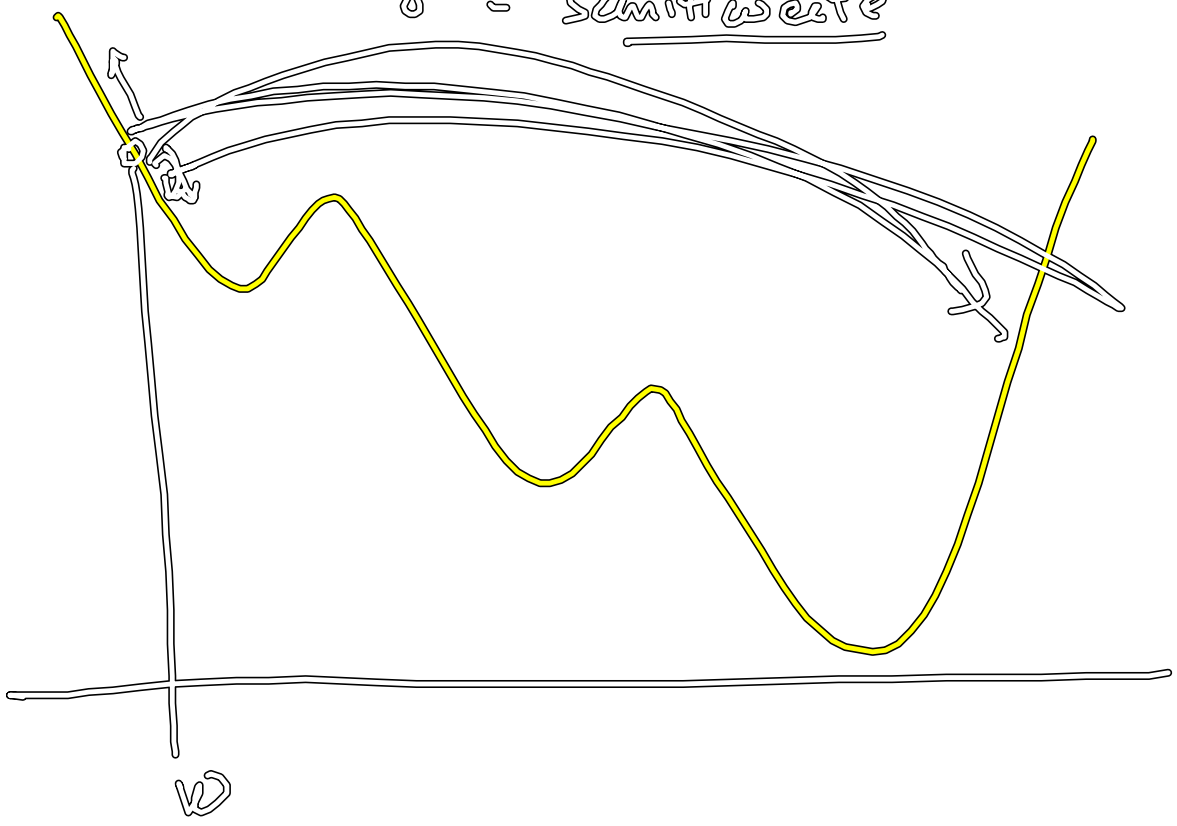
$$w_{600} := w_{600} - \delta \frac{\partial E}{\partial w_{600}}$$

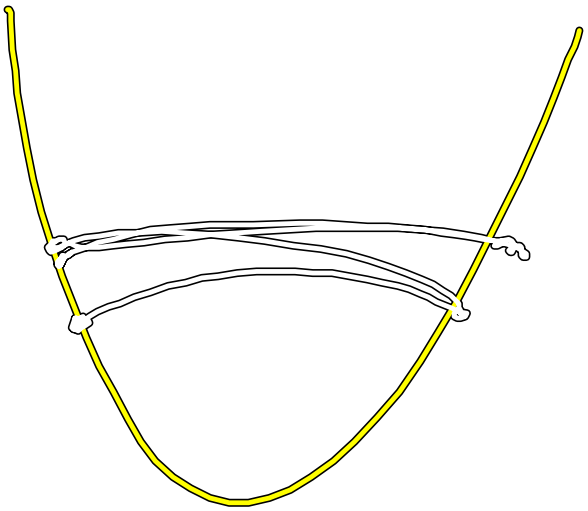
4.- Teste Fehler

groß → weiter bei (1)

klein genug → Stopp.

δ = Schrittweite





Zusammenfassung

→ Wir modellieren die Funktion

$$P \begin{cases} P_0(x_1 \dots x_{20}) \\ P_1(x_1 \dots x_{20}) \\ P_2(x_1 \dots x_{20}) \\ \vdots \\ P_q(x_1 \dots x_{20}) \end{cases}$$

→ Gradienten abstieg für die Justierung der Gewichte

→ Es gibt keine Garantie für globales Min.

↳ on-line (Randomized) G.A.

↳ off-line