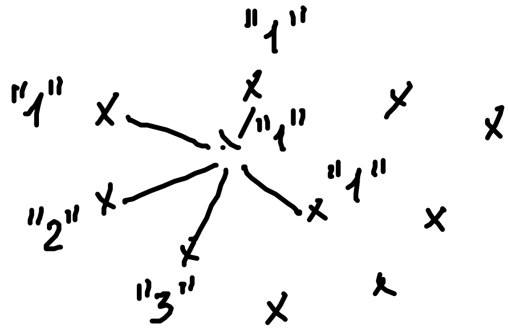
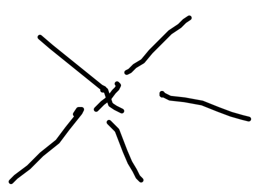
Klassifikatoren

K-NN Klassifiziert



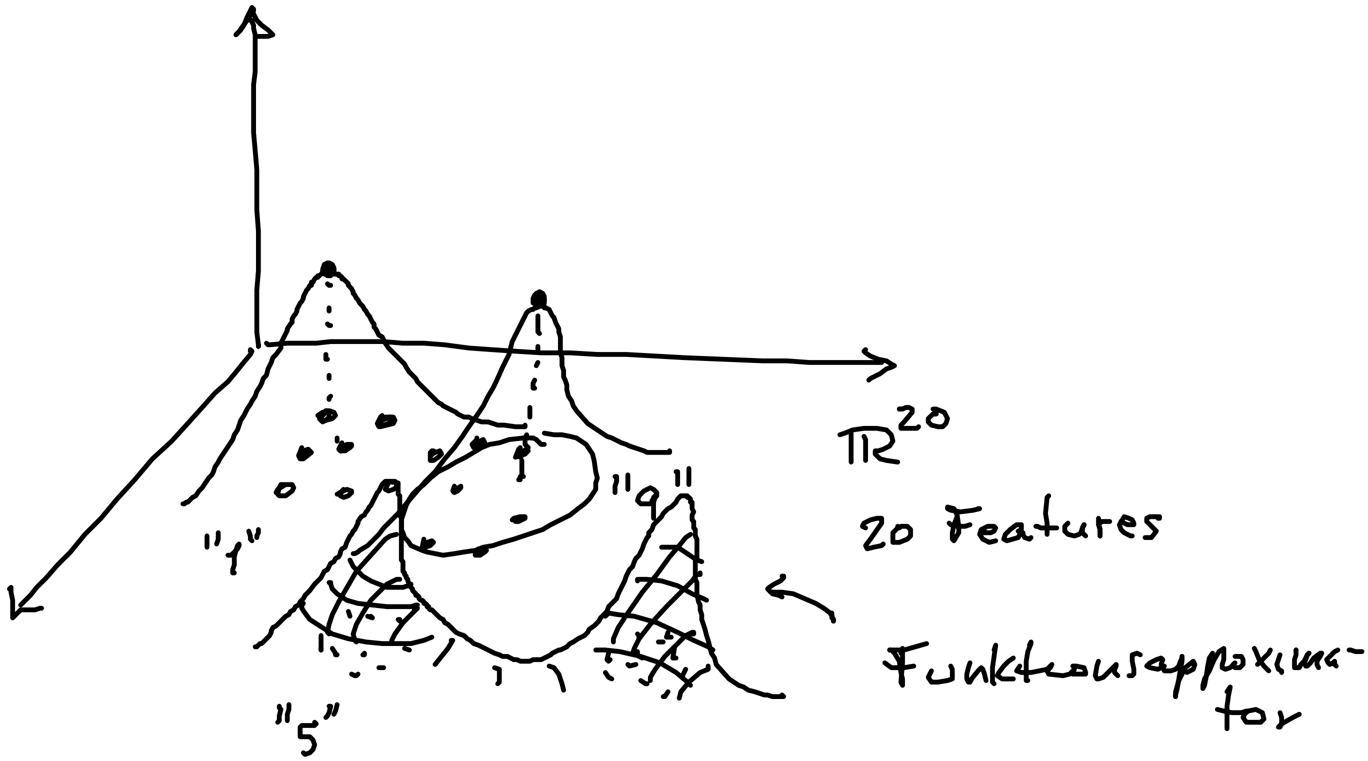
55,000 Ziffern



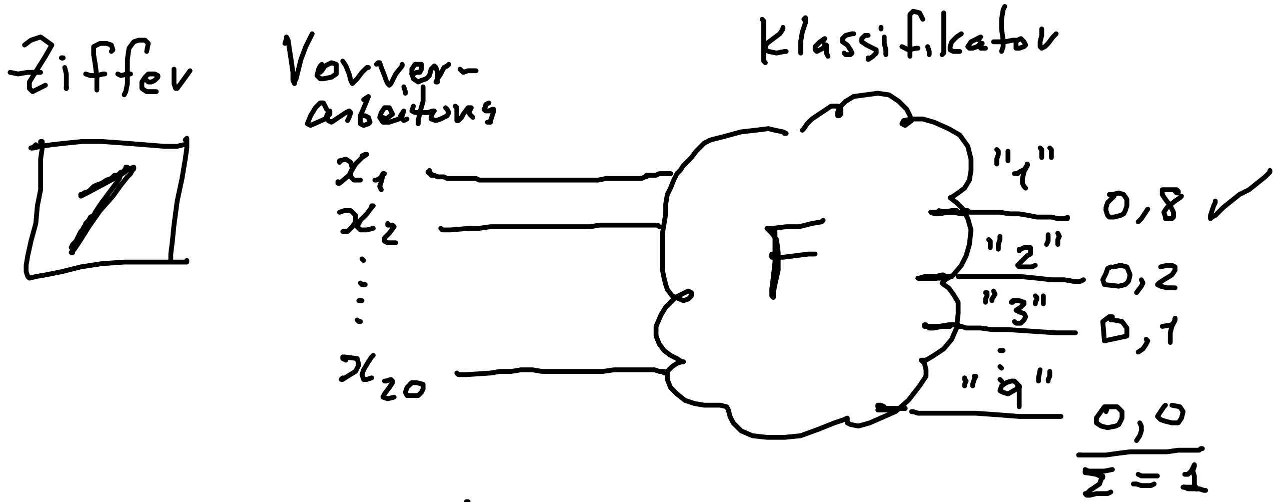
55,000 $\sqrt{\underbrace{(x_1-y_1)^2}+\cdots(x_n-y_n)^2}$

$55{,}000 \times 20 \approx 10^6$

100,000 Ziffern



Allgemeiner Klassifikator

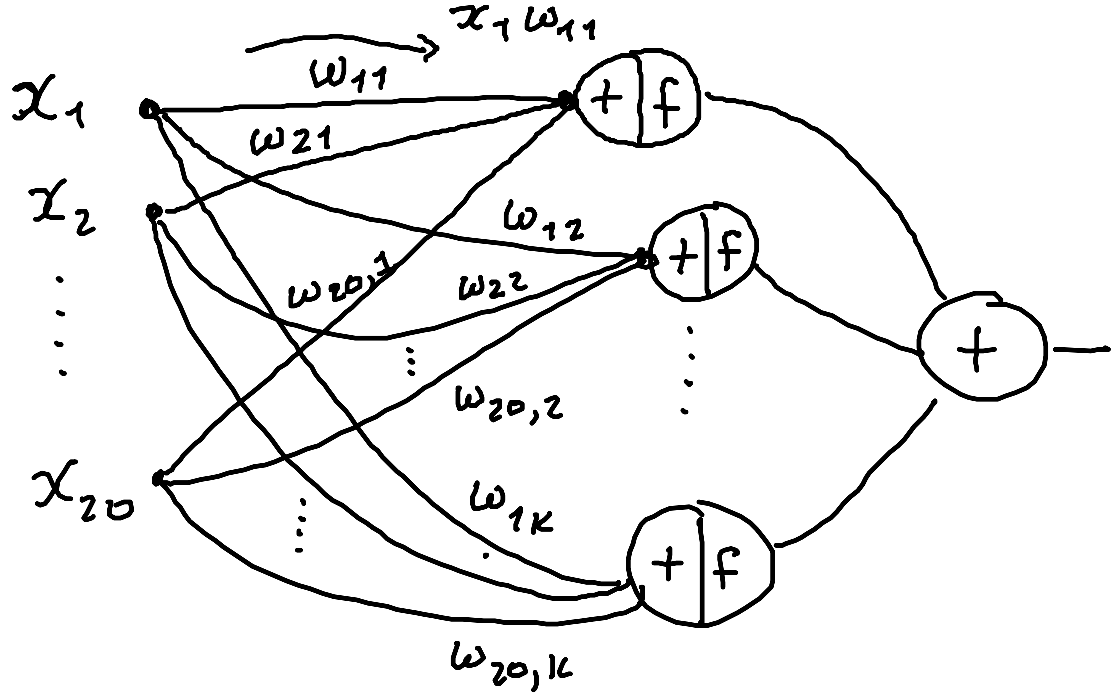


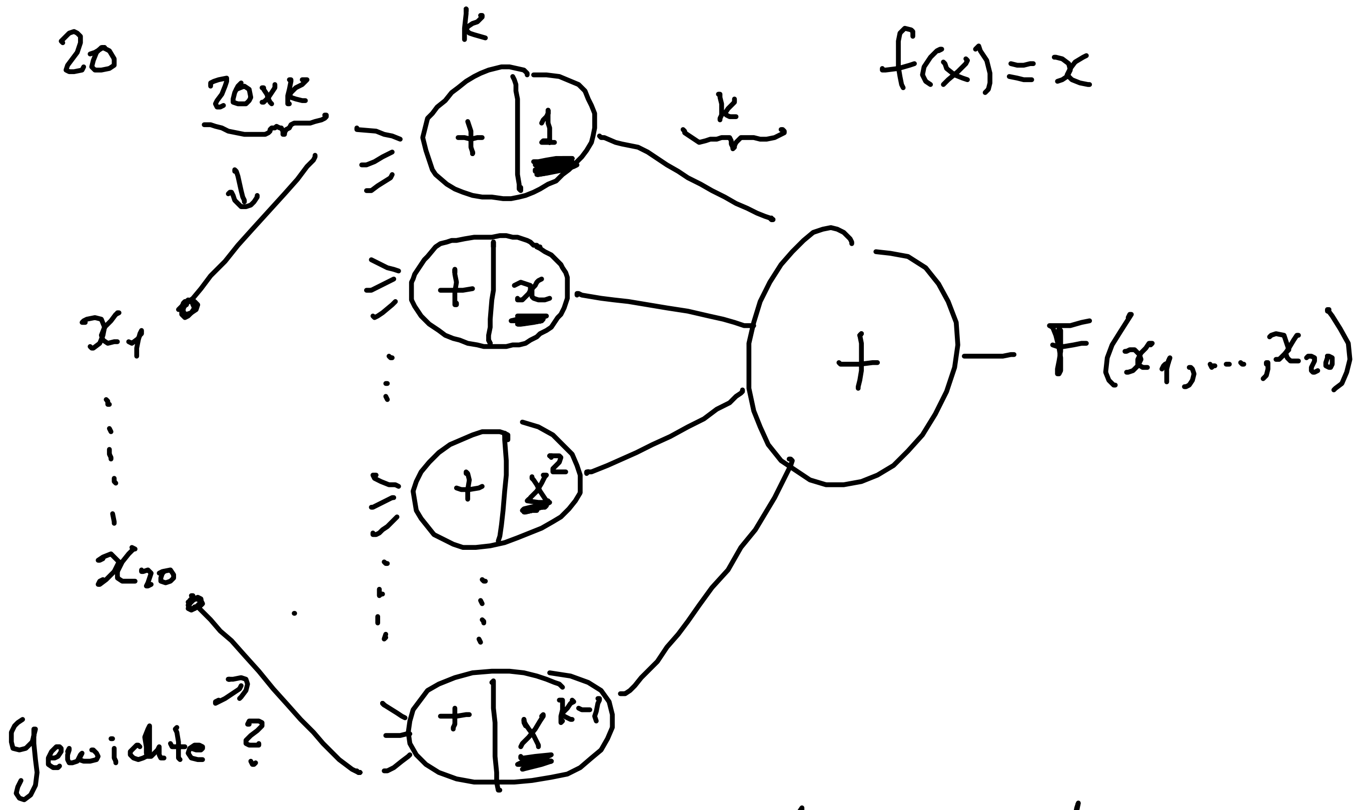
$$F_1(x_1, x_2, \dots x_{20}) = 0{,}8$$

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

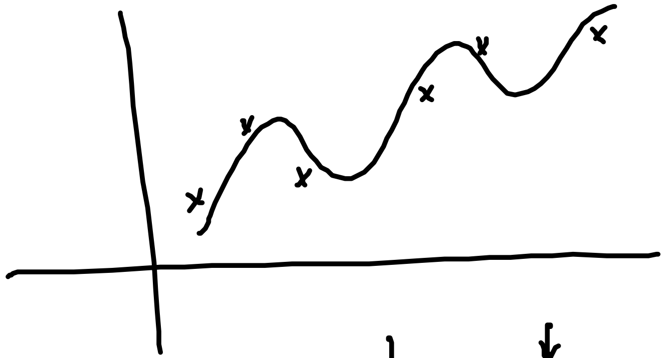
$$\vdots$$

Funktionennetz = Neuronales Netz

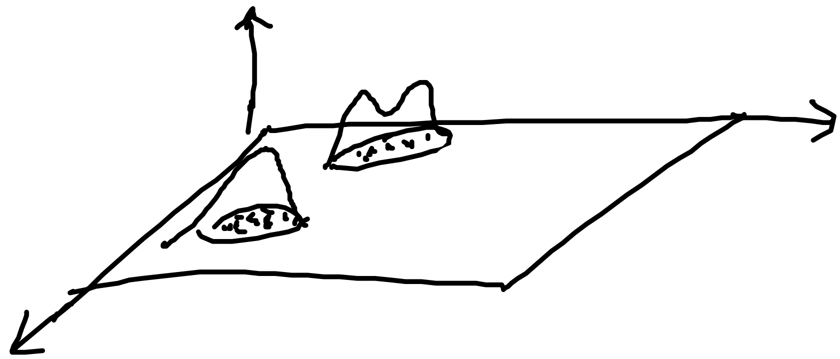


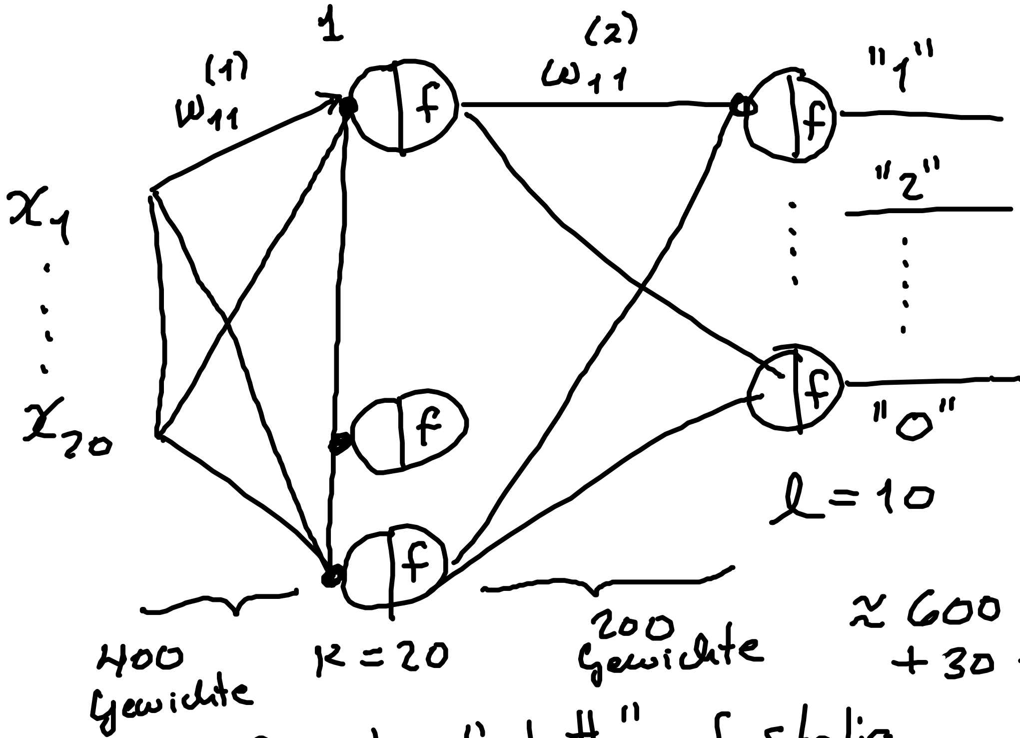


$$F(x_1, \ldots, x_{20}) = a_0 + a_1 x_1 + a_2 x_2 \cdots$$
$$\cdots + a_{20} x_{20} + a_{21} x_1^2 + \cdots$$
$$\cdots + a_{40} x_{20}^2 + \ldots x_1 x_2 \cdots$$
$$\cdots x_1^{k-1} + \cdots$$



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





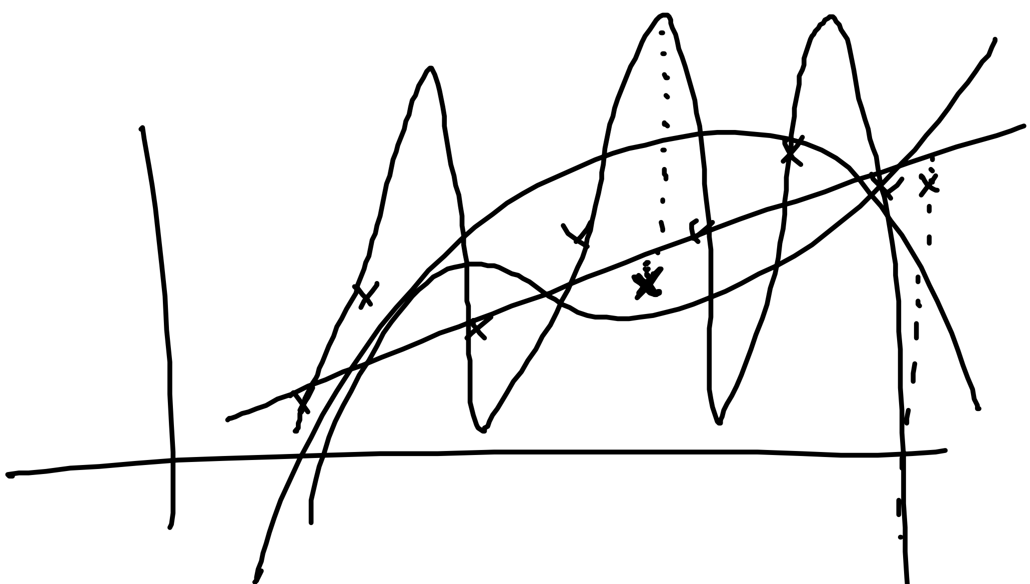
≈ 600 Mult. $+ 30$ f's

f ist "glatt" $\begin{cases} \text{stetig} \\ \text{differenzierbar} \end{cases}$

$\underline{10^6}$ 100,000 Ziffer

vs. 600

$\underbrace{\text{K-NN}}_{\text{genauer}}$ vs. $\underbrace{\underbrace{\text{NN}}_{\text{schneller}}}_{\text{genau}}$



Fehler → Trainingsmenge

Fehler → Recall (neue Eingaben)

<u>Bes</u>tes Modell + <u>Ei</u>nfachstes

Trainingsmenge

100,000 Ziffern

↓

Merkmale

↓

Gewichte

↓

Recall

- Zufällige Gewichte
- testen

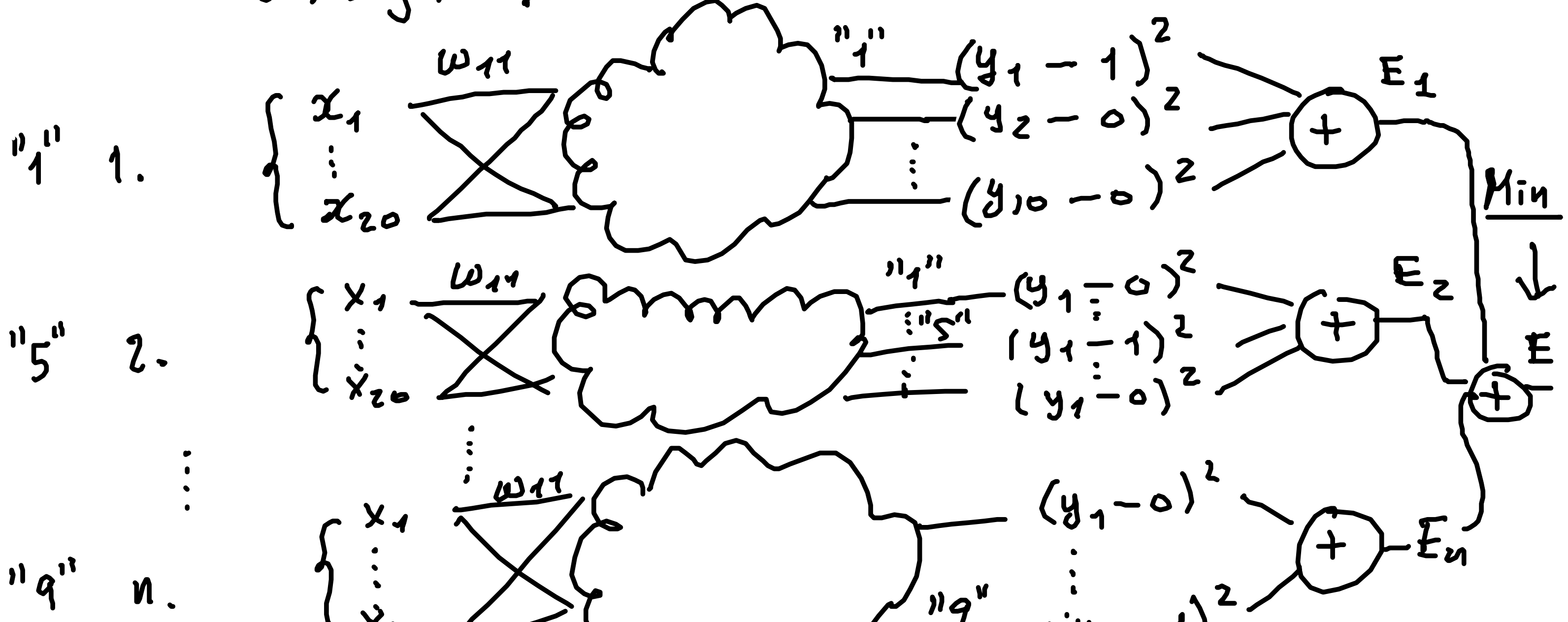
→ Fehler der Trainingsmenge

- Fehler > ε

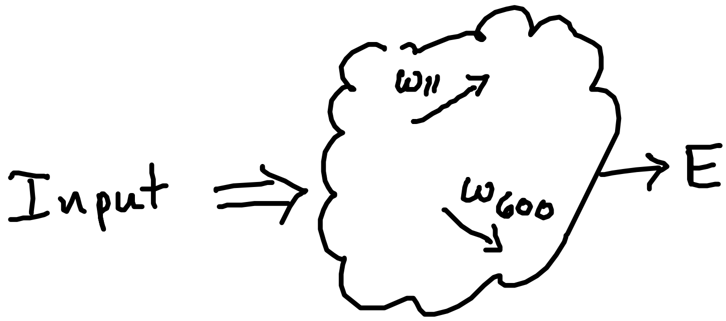
korrigieren

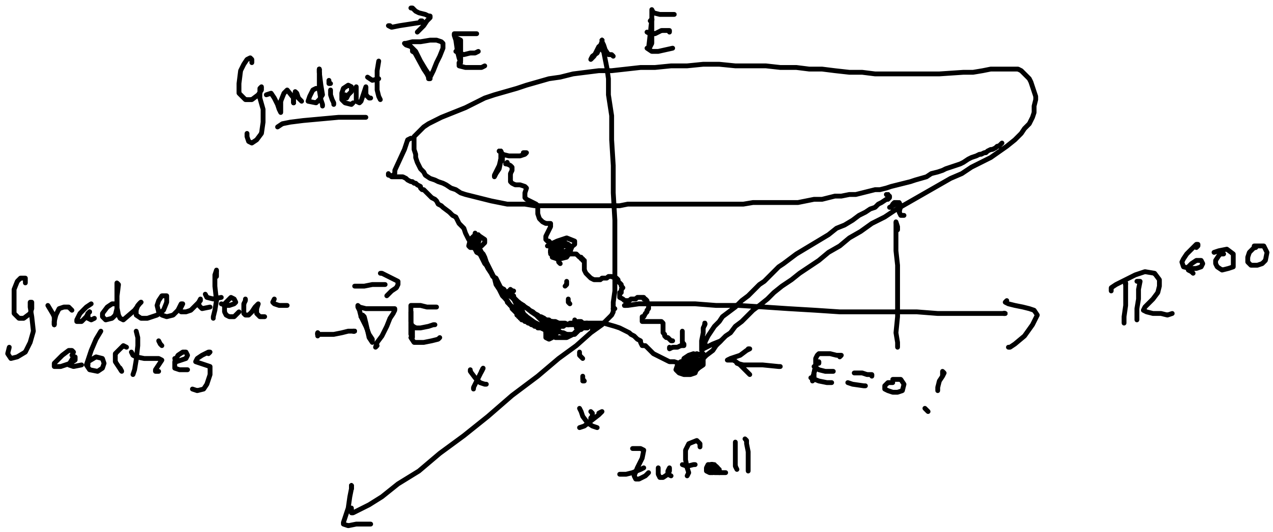
(bis Fehler < ε)

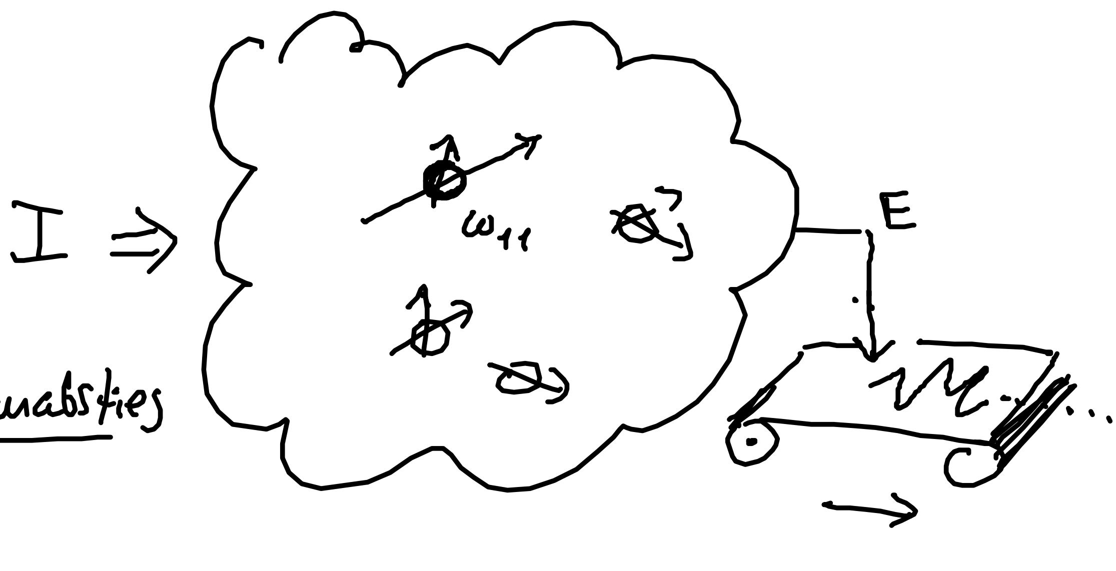
Trainingsmenge



Netz





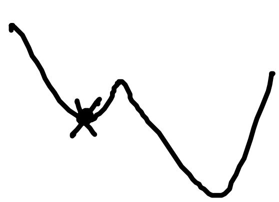


Gradientenabstieg

$$\frac{\partial E}{\partial \omega_{11}} \cdots$$

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

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



$$\frac{\partial E}{\partial w_{11}} \quad \dots \quad \frac{dE}{dw_{11}}$$

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

Backpropagation

Systematisch

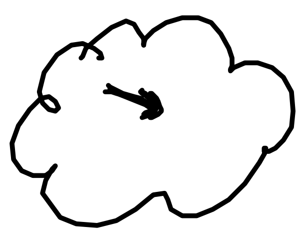
Vorwärtsschitt

$$x \xrightarrow{w} wx$$

Backprop-Schritt

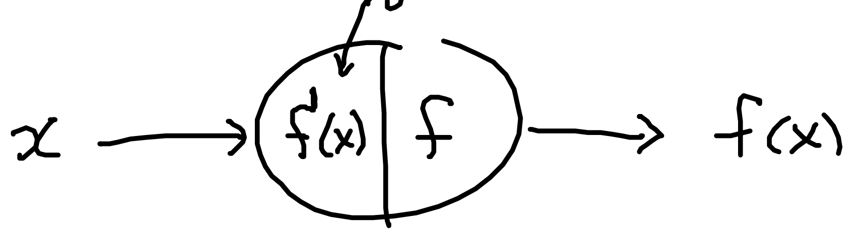
$$w \xleftarrow{w} 1$$ Traversierungswert

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

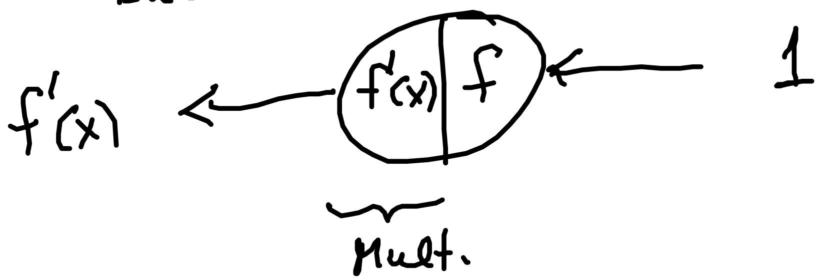
Input $\Rightarrow$  $\Rightarrow$ F(I)

$$\frac{dF}{dI}$$

Vorwärts

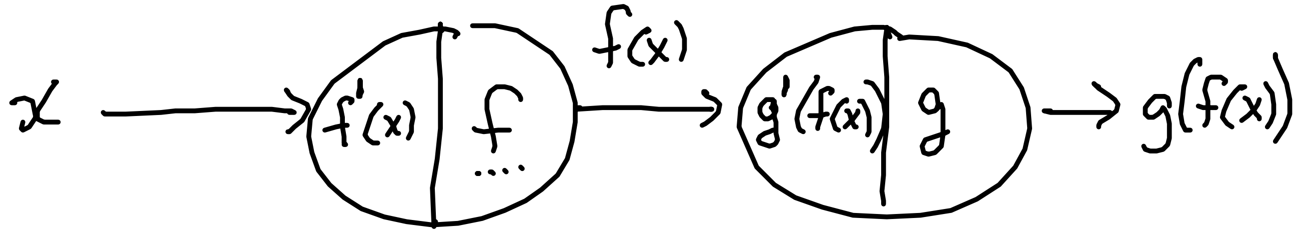


Rückwärts

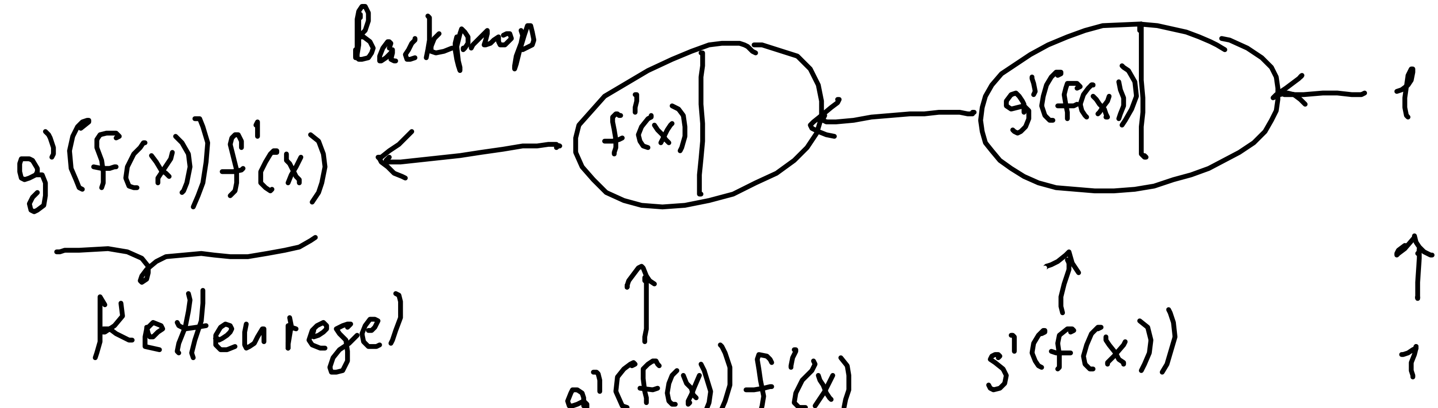


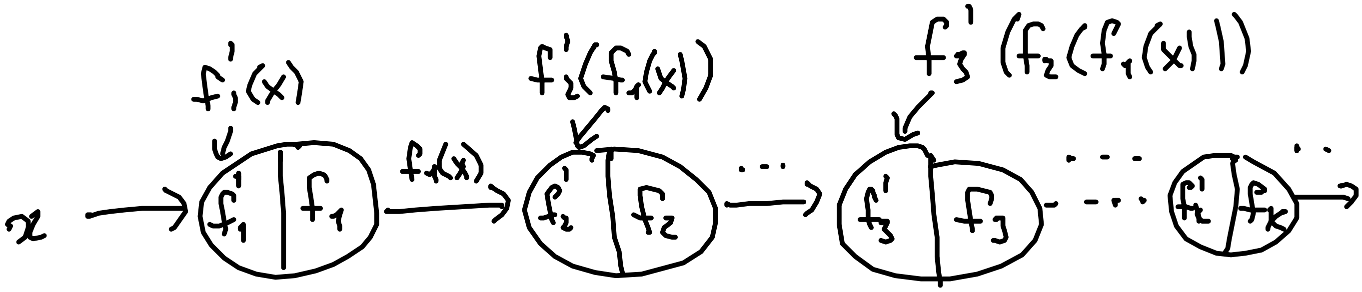
$$\frac{d(f(x))}{dx} = f'(x)$$

Vorwärts



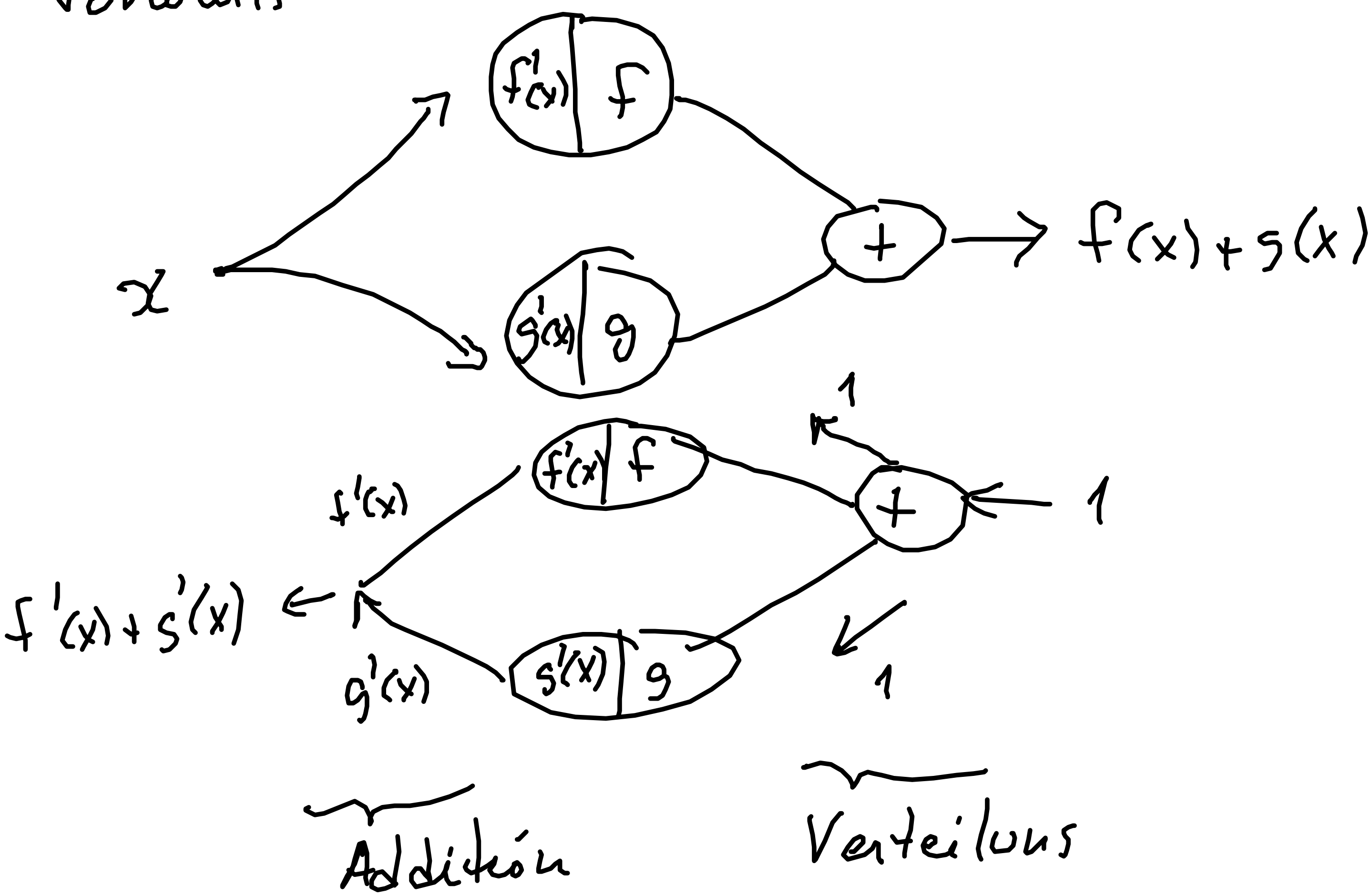
Backprop





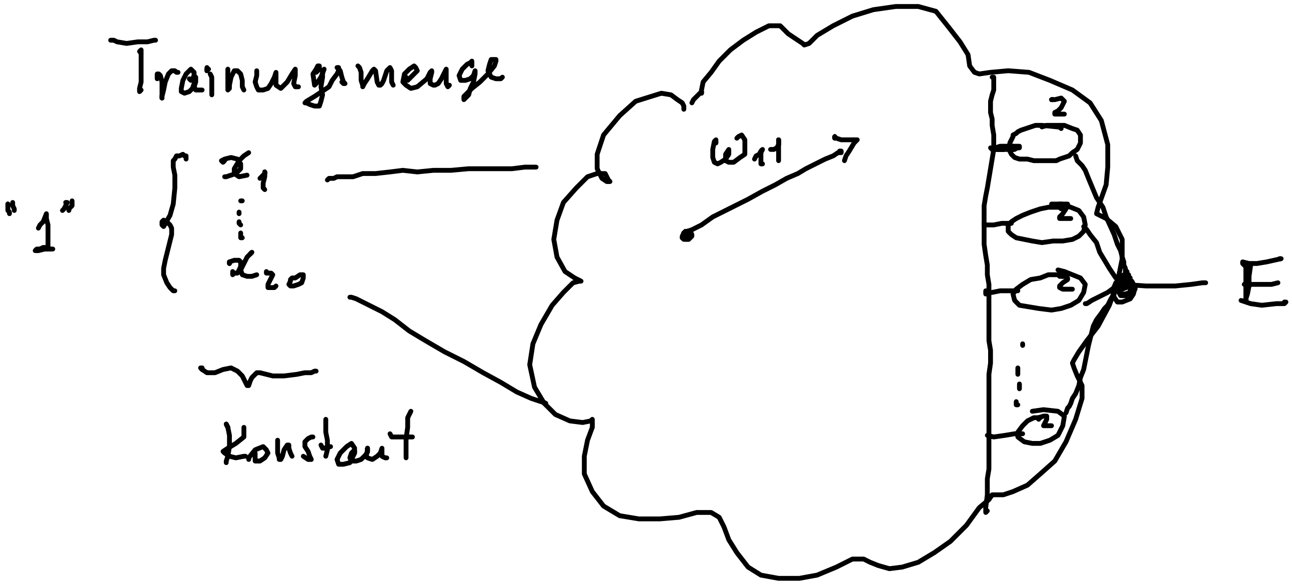
per Induktion beweisen....

Vorwärts



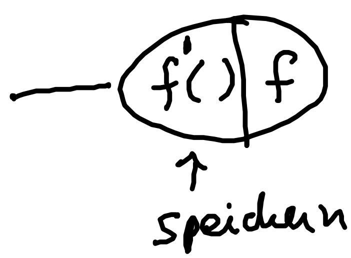
$$\frac{d}{dx}\left(f(x)+g(x)\right) = f'(x) + g'(x)$$

Jetz zu den Gewichten..

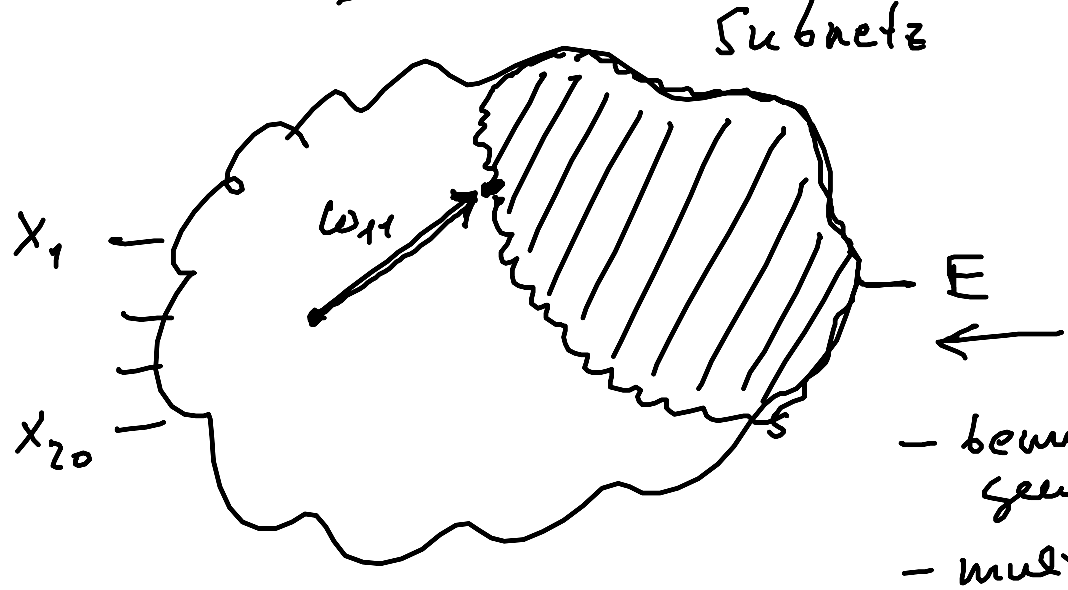


$$\vec{\nabla} E = \left(\frac{\partial E}{\partial w_{11}}, \frac{\partial E}{\partial w_{12}}, \dots \frac{\partial E}{\partial w_{\dots}} \right)$$

1. Schritt : Vorwärtslauf



2. Schritt: Backprop.



- benutze Gewichte
- multipliziere Traversierungswert 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)

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

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

$$\vdots$$

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

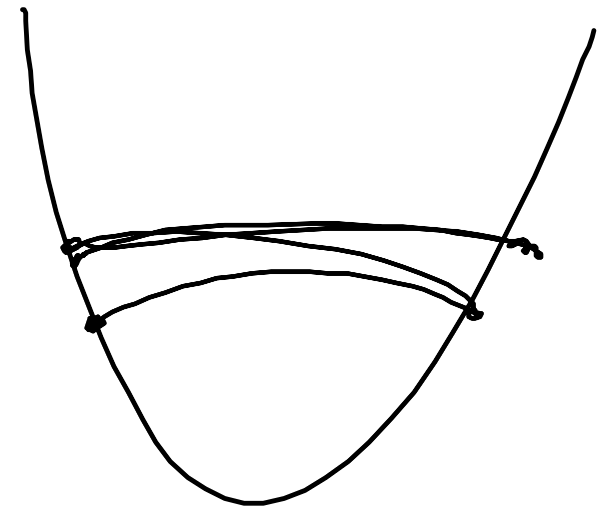
4.- Teste Fehler

groß → weiter bei (1)

Klein genug → Stopp.

δ = Schrittweite

δ ....



Zusammenfassung

→ Wir modellieren die Funktion

$$P \begin{cases} P_0(x_1 \ldots x_{20}) \\ P_1(x_1 \ldots x_{20}) \\ P_2(x_1 \ldots x_{20}) \\ \vdots \\ P_9(x_1 \ldots x_{20} \end{cases}$$

→ Gradientenabstieg für die Justierung der Gewichte

→ Es gibt keine Garantie für globales Min.

- ↳ on-line (G.A. Randomisiert)
- ↳ off-line