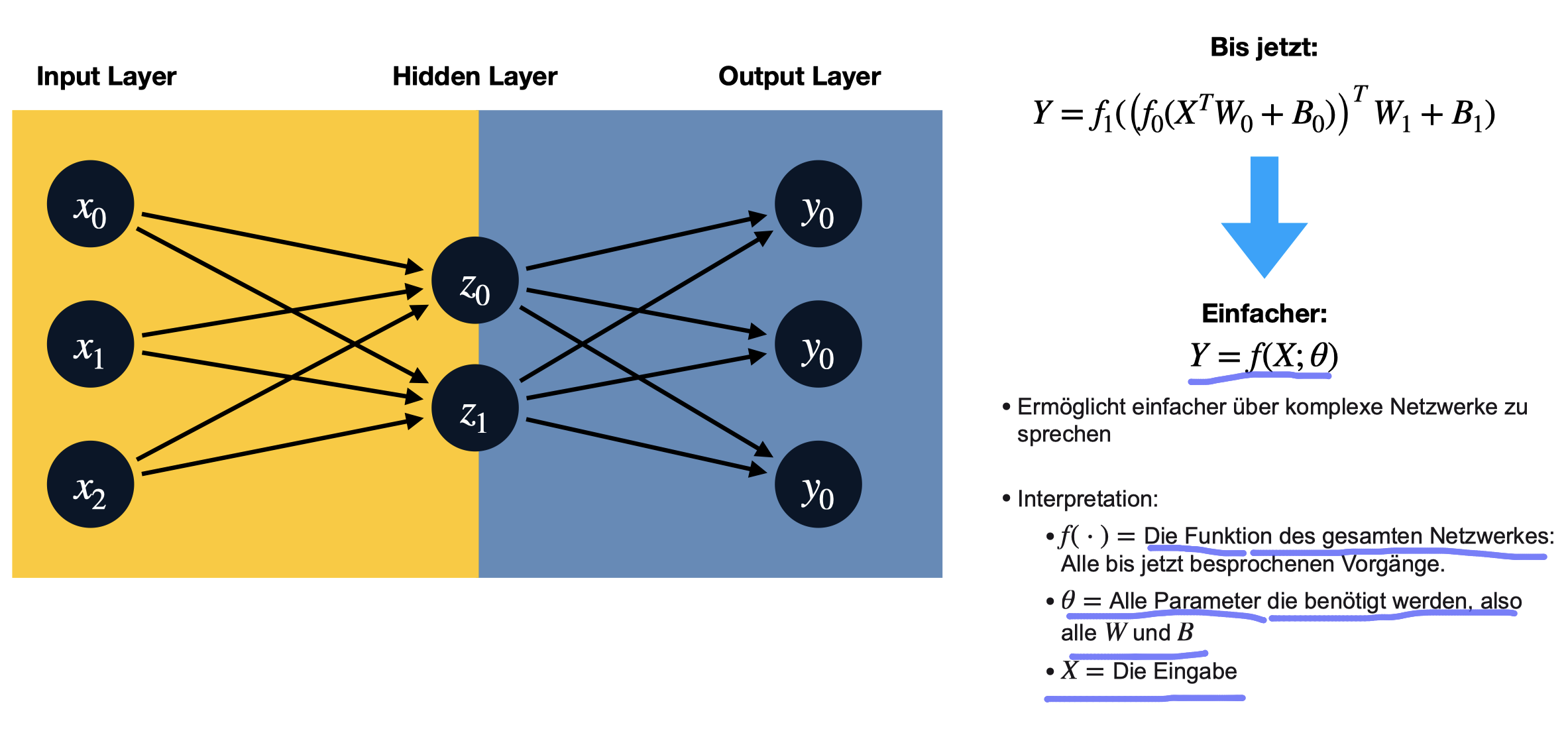
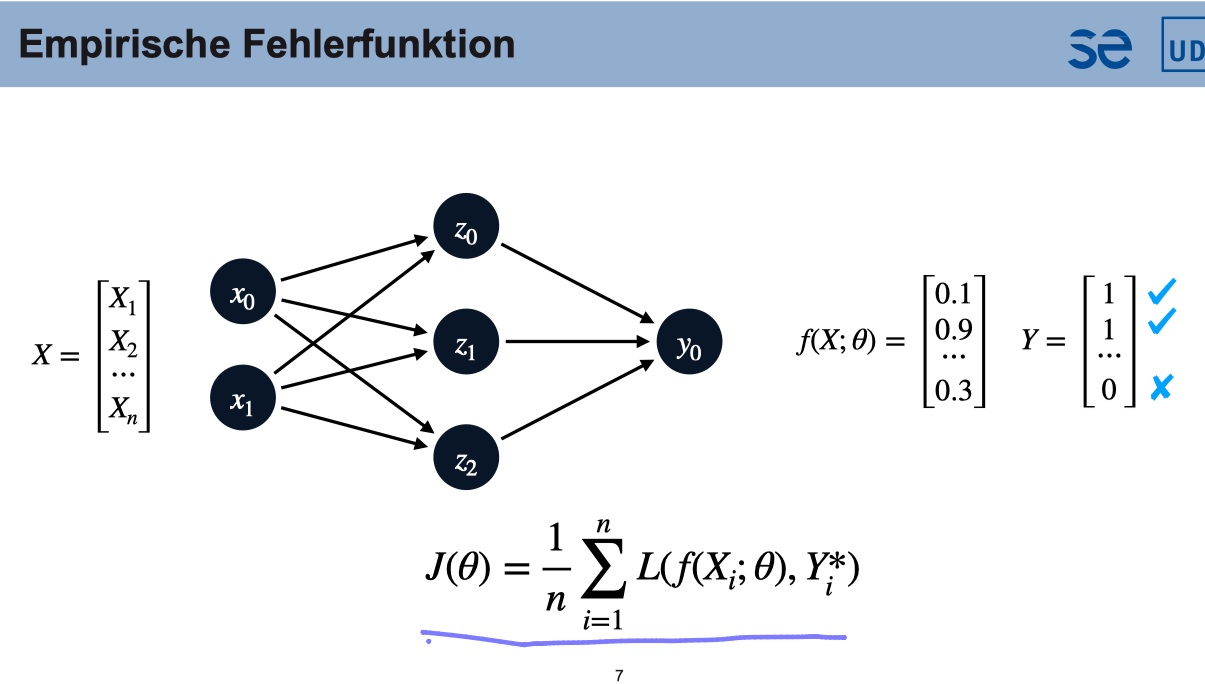
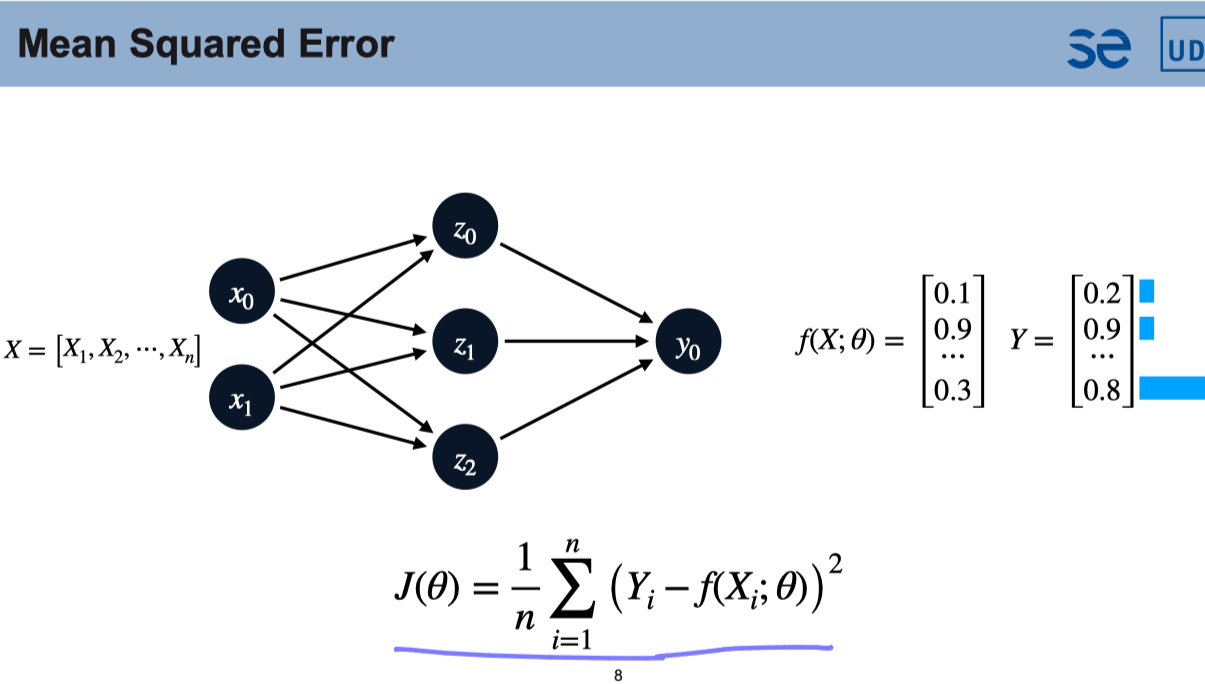
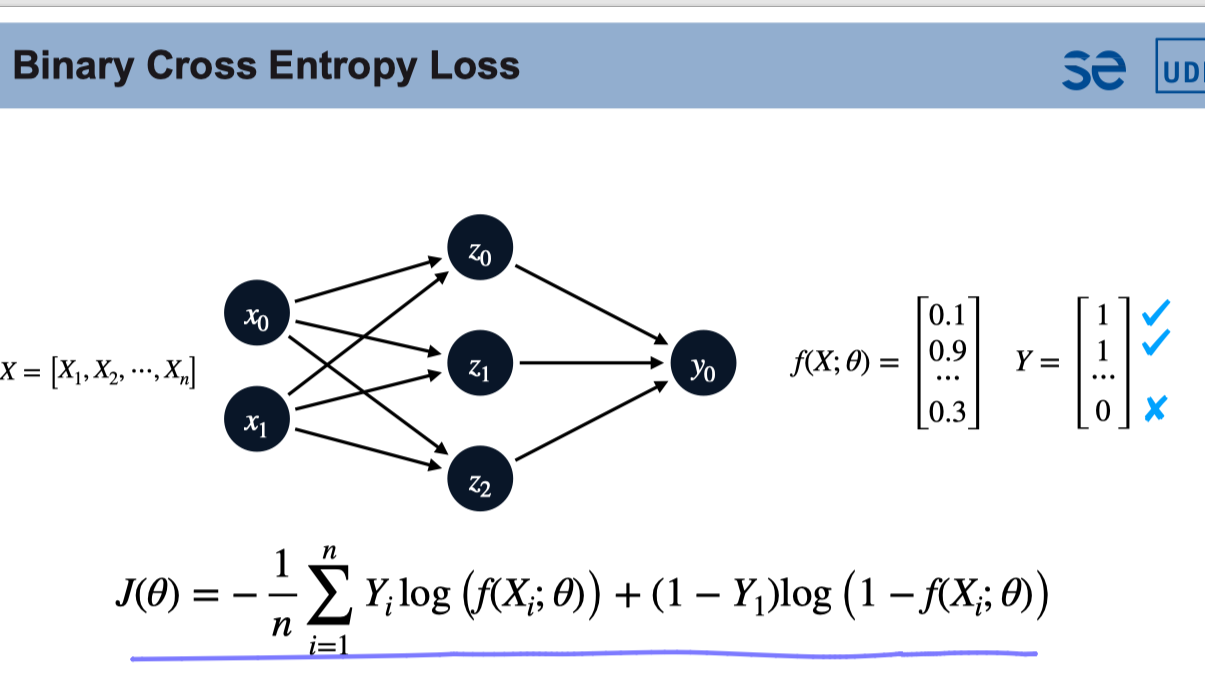
**Neural Network 2 Backpropagation and Optimization**

****

**1. Fehlerfunktion:**

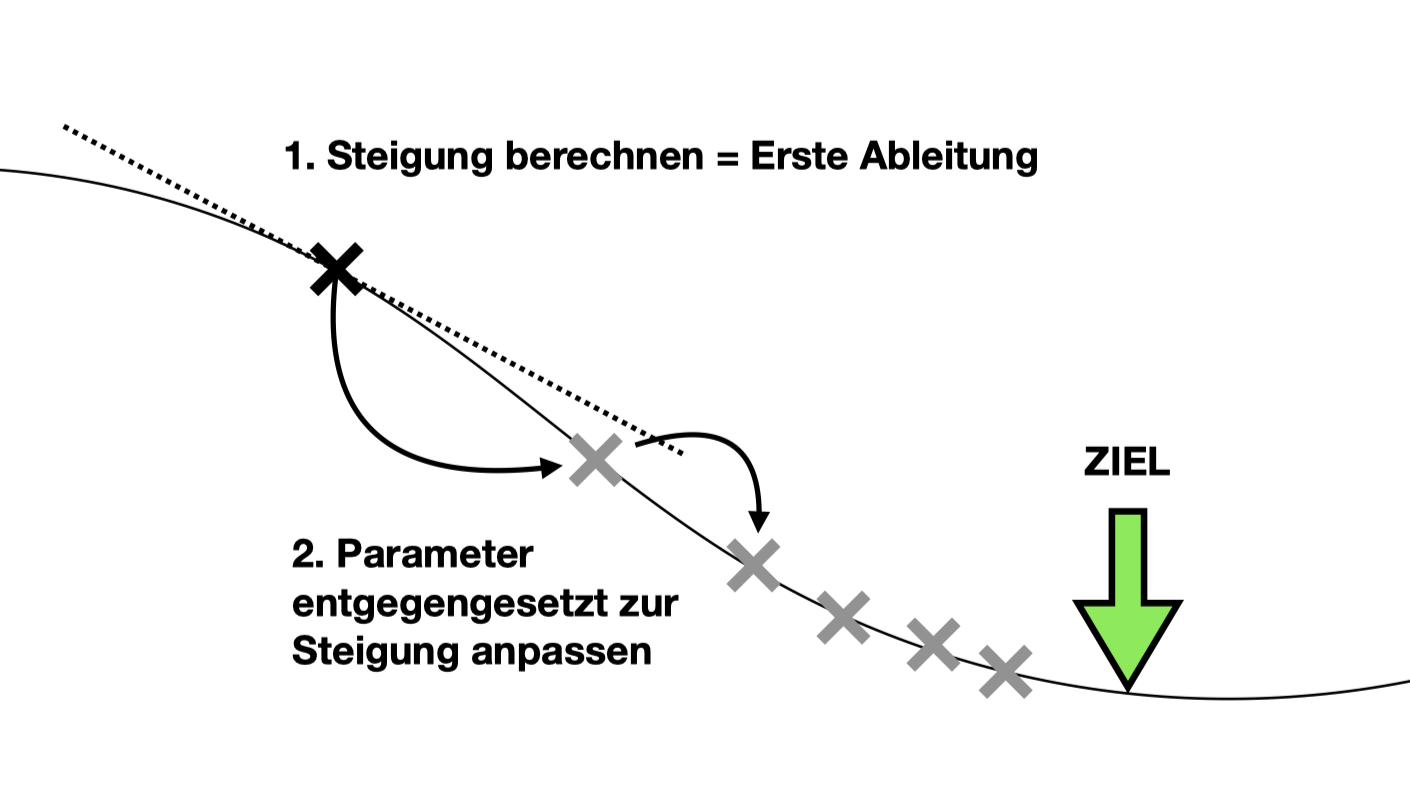
****

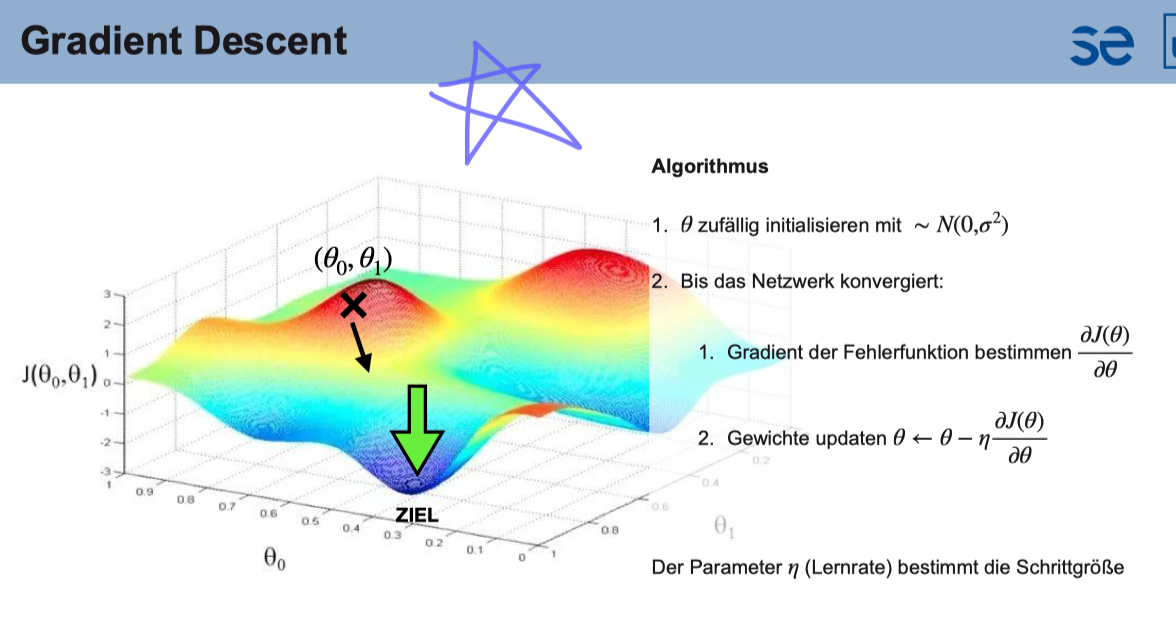
**2. Optimierung der Fehlerfunktion:**

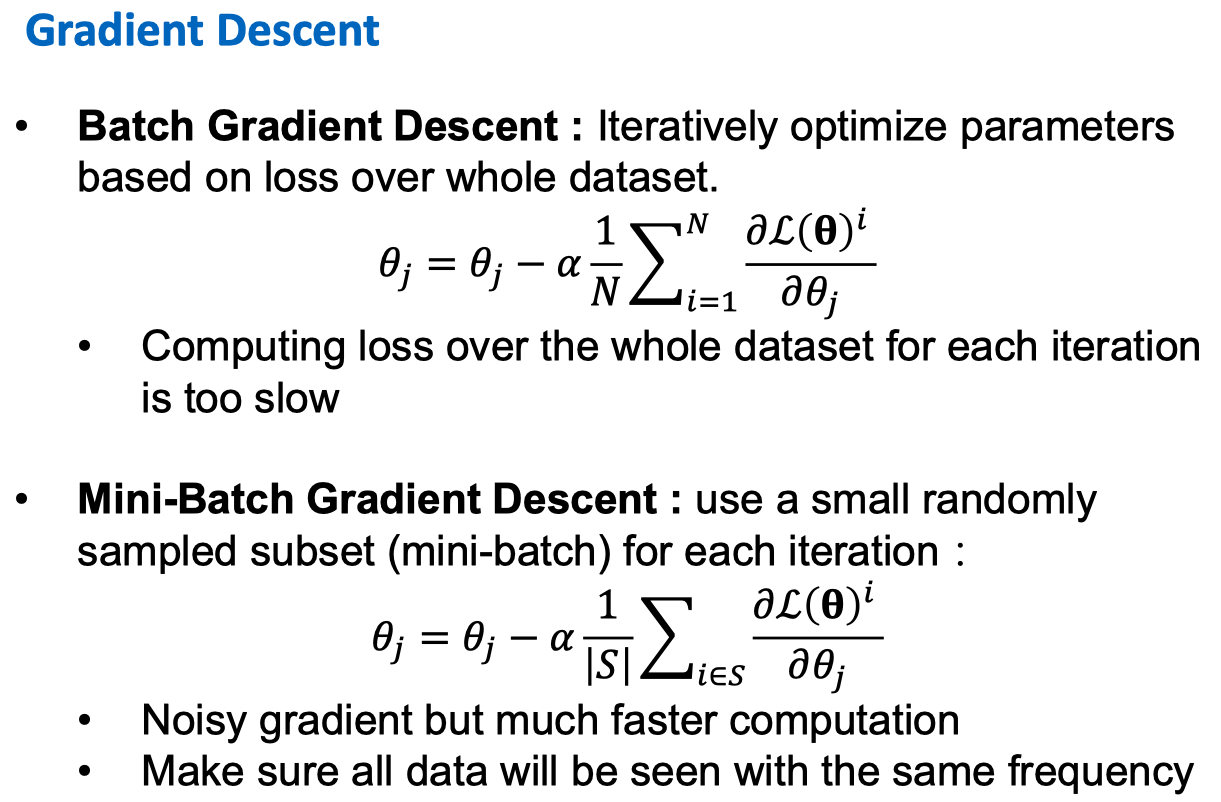
Ziel: den geringsten Fehler erzeugen.

**Gradient Descent:**

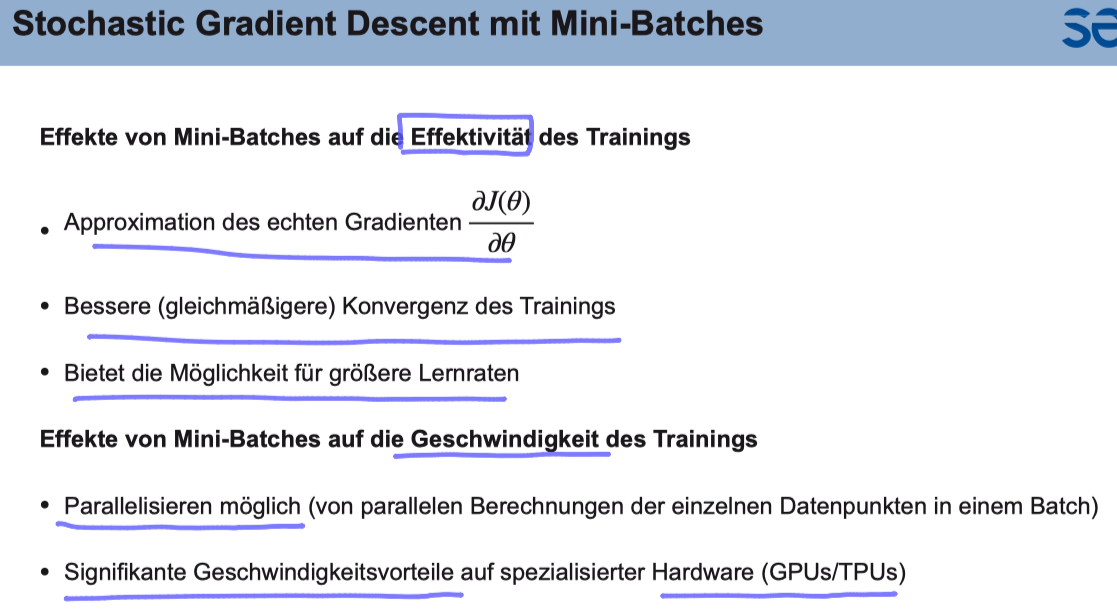
(1). Grundidee: Steigung berechnen, Parameter entgegengesetzt zur Steigung anpassen.







(2). Mini-Batch Gradient Descent:



(3). Lernrate

a. Kleine Lernraten führen zur langsamen Konvergenz und bleiben möglicherweise in einem lokalen Minimum

b. Große Lernraten führen können über das Ziel „hinausschießen“ und verhindern so möglicherweise eine ausreichende Konvergenz.

c. Viele verschiedene Werte ausprobieren.

d. Adaptive Lernraten, abhängig von Fehlerfunktion (Gradient, Verbesserung der Fehlerfunktion, Gewicht oder Verteilung).

**Backpropagation:**

