

**Kompensationsarbeit im Modul Software Entwicklung**

Master:

“Data Science & Intelligent Analytics”

Betreuerin :

Huber Stephan, MA

Autor:

Jochen Hollich | 1810837475

Datum

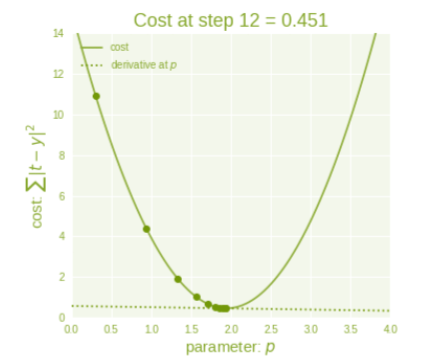
10.12.2019

Inhalt

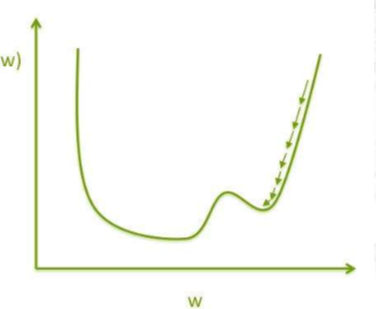
# Gradient Descent

Verwendung für die Optimierung der Cost-Functino in NN, mit anderen Worten auswähle für Hyperparameter um die Cost-Function runterzuschrauben. Dazu geht man schrittweise vor.

Optimales BSP Fehlerfunktion geplotet ist Konvex (nach oben geöffnet):



Leider häufig nicht der fall sondern



# Varianten von Gradient Descent

Batch GD

Bleibt häufig in Suboptimalen Minima hängen

Hier wird pro einmal durchlaufen des gesamten DS der Gradient aktualisiert

|  |  |
| --- | --- |
| + | - |
| Nicht Rechenintensiv | Gefahr lokales Minimum |

Stochastic GD || SGD

Nach jeder Instanz im Datensatz wird der SGD aktualisiert

Im Vergleichzu BGD sehr rechenintensiv

|  |  |
| --- | --- |
| + | - |
| Gefahr lokales Minimum klein | Sehr rechenintesiv |

Mini Batch SGD

Mehre Instanzen zu einer Gradientenaktualisierung zusammenfassen

Kompromiss zwischen Batch GD &

Unterschiede von Gradient Descent

1. Rechenintensiviät