Overfitting – Teil 3

Weight Decay

INNOVET
KI B³
Künstliche Intelligenz in die berufliche Bildung bringen

INNOVET
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung
bibb Bundesinstitut für Berufsbildung

Gefördert als InnoVET-Projekt aus Mitteln des Bundesministeriums für Bildung und Forschung.

L1- und L2-Regularisierung

Idee

- Große Gewichte können Anzeichen von Overfitting sein:
- Netz reagiert stark auf ein bestimmtes Input
- Idee: große Gewichte nur erlauben, wenn sie sich positiv auf den Fehler auswirken

Implementierung

- Minimiere nicht nur den Fehler, sondern auch die Beträge der Gewichte
- Addiere zum Verlust noch
  - die Beträge der Gewichte (L1-Regularisierung) oder
  - die Quadrate der Gewichte (L2-Regularisierung)
- Minimiere diese Summe

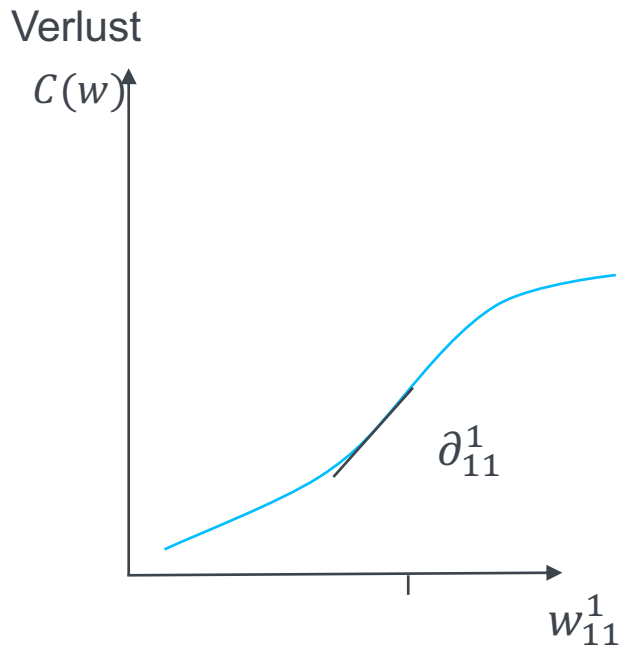
multipliziert mit der Regularisierungs-rate (sehr klein!)

Bei der L2-Regularisierung werden zum Verlust noch die quadrierten Gewichte addiert, multipliziert mit einer Regularisierungsrate. Die Regularisierungsrate ist normalerweise nahe Null.

Dadurch werden große Gewichte „bestraft“. Die kleine Regularisierungsrate bewirkt, dass die Minimierung des Fehlers dabei aber wichtiger bleibt.

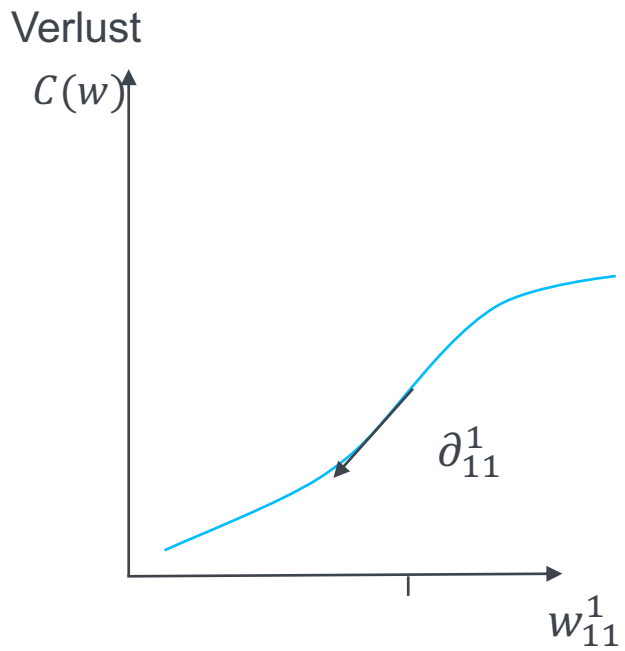
Wiederholung: Gradientenabstiegsverfahren

- Beim Training interessant:
  - Wie verändert sich der Verlust, wenn die Gewichte verändert werden?
  - Verlust als Funktion über alle Parameter des Netzes w
    $C(w) = C(w_{11}^1, w_{12}^1, \dots, b_1^1, b_2^1, \dots)$
- Für jeden Parameter
  - Berechne den Gradienten
  - z.B. ∂_{ij}^k für w_{ij}^k



Anpassung der Parameter beim Gradientenabstiegsverfahren

- Anpassung der Gewichte für Schritt t+1:
  - $w_{ij}^{k(t+1)} = w_{ij}^{k} - \eta \partial_{ij}^{k}$
  - η Lernrate, z.B. 0.001
- Beispiel rechts:
  - $\partial_{ij}^{k} = 1$, also
  - $w_{ij}^{k(t+1)} = w_{ij}^{k} - \eta$



Verlustfunktion bei L2-Regularisierung

der kleine Faktor

L2-normierte Gewichts matrix

Ohne Regularisierung

$C(w)$

Gradient für w_{ij}^k

∂_{ij}^k

Anpassung für w_{ij}^k

$$w_{ij}^{k(t+1)} = w_{ij}^k - \eta \partial_{ij}^k$$

L2-Regularisierung

$$C_{L2}(w) = C(w) + \frac{1}{2}\lambda \|w\|^2$$

Alle Gewichte quadriert und aufsummiert

$$\partial_{L2}{}_{ij}^{k} = \partial_{ij}^k + \lambda w_{ij}^k$$

$$\begin{aligned} w_{ij}^{k(t+1)} &= w_{ij}^k - \eta \partial_{L2}{}_{ij}^{k} \\ &= w_{ij}^k - \eta \left(\partial_{ij}^k + \lambda w_{ij}^k\right) \\ &= \underbrace{w_{ij}^k - \eta \partial_{ij}^k}_{\text{wie bisher}} \underbrace{- \eta \lambda w_{ij}^k}_{\text{lässt Gewicht gegen Null gehen}} \end{aligned}$$

L2-Regularisierung wird auch **Weight Decay** genannt

Bei der Anpassung der Gewichte nach dem Gradientenabstiegsverfahren wird durch den Gradienten des L2-Regularisierungs-Terms immer auch das Gewicht selbst abgezogen, gewichtet mit Lernrate η und Regularisierungsrate λ.

Dadurch gehen die Gewichte immer weiter gegen Null. L2-Regularisierung wird daher auch als Weight Decay (zu Deutsch: “Gewichteverfall”) bezeichnet.

INNOVET
KI B³
Künstliche Intelligenz in die berufliche Bildung bringen

Dr. Antje Schweitzer

Universität Stuttgart
Institut für Maschinelle Sprachverarbeitung

Universität Stuttgart
Institut für Maschinelle Sprachverarbeitung
Institut für Software Engineering

IHK Industrie- und Handelskammer Reutlingen
Reutlingen | Tübingen | Zollernalb

IHK Region Stuttgart

IHK Industrie- und Handelskammer Karlsruhe

LMU Ludwig-Maximilians-Universität München

GEFÖRDERT VOM
INNOVET
Bundesministerium für Bildung und Forschung
bibb Bundesinstitut für Berufsbildung

Gefördert als InnoVET-Projekt aus Mitteln des Bundesministeriums für Bildung und Forschung.

Lizenzbestimmungen

“Overfitting – Teil 3: Weight Decay" von Antje Schweitzer, KI B³ / Uni Stuttgart

Das Werk - mit Ausnahme der folgenden Elemente:

- Logos der Verbundpartner und des Förderprogramms
- im Quellenverzeichnis aufgeführte Medien

ist lizenziert unter:

 CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/deed.de)
(Namensnennung 4.0 International)

Quellenverzeichnis

Titelfoto: Evelyn Bertrand(https://unsplash.com/de/@evelyn_bertrand) auf Unsplash (https://unsplash.com/de/fotos/zwei-verwelkte-rosa-rosen-mit-schwarzem-hintergrund-pwfY1mIm2EM), Zwei verwelkte rosa Rosen mit schwarzem Hintergrund, lizenziert unter Unsplash-Lizenz (https://unsplash.com/license). Bildausschnitt verändert.