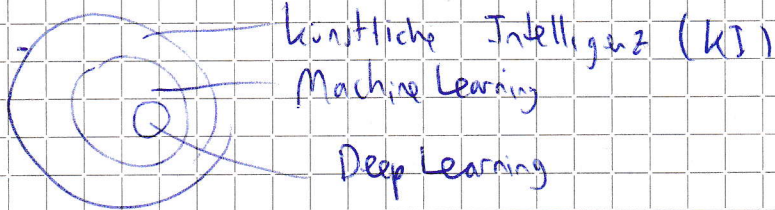


Deep Learning (KLR-341)

Monday, 04.12.2023

Bernd Eberbach

TensorFlow



supervised vs unsupervised Learning
überwachtes vs unüberwachtes Lernen

{
Regression,
Klassifikation

{
Clustering
Ausreißererkennung
(outlier)

One Hot Encoding

Ordinal-Encoding

Skalierung

Anaconda

Dienstag, 05.12.2023

Tensor Flow & PyTorch

MATLAB & Octave

MATLAB vs TensorFlow in Google Trend \rightarrow MATLAB more

künstliche neuronale Netze / Artificial Neural Networks

Aktivierungsfunktion: - Logistische Funktion (Sigmoid)

- tanh

- ReLU = rectified linear unit

MLP = Multi Layer Perceptron / Mehrschichtiges Perzeptron

MLP Classifier in sklearn

~~Loss~~ Loss function \rightarrow minimization

\rightarrow Least Squares / Methode der kleinsten Quadrate

Gradientenabstieg / Gradient Descent.

$$\text{loss} = \frac{1}{n} \sum_{i=1}^N [x_i - (w x_i + b)]^2 \rightarrow f(g(w))$$

$$\left. \begin{array}{l} g(w) = y - w x - b \\ f(g) = g^2 \end{array} \right\} \left\{ \begin{array}{l} \frac{\partial f}{\partial w} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial w} \end{array} \right\} \text{RWS}$$

$$\frac{\partial f}{\partial g} = 2 \cdot g$$

$$\frac{\partial g}{\partial w} = -x$$

$$\frac{\partial g}{\partial b} = -1$$

Aufgabe: - numpy, ~~manuell~~ Netzwerk mit Bias

\rightarrow then with TensorFlow

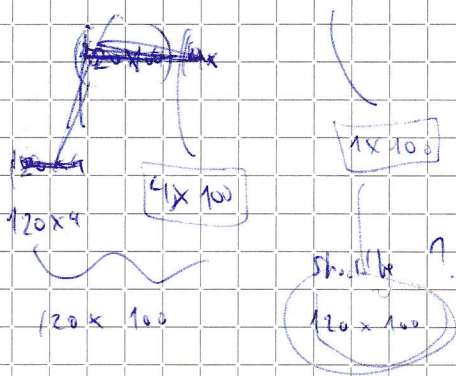
(and Gradient Type)

Mittwoch, 06.12.2023

$$\frac{\partial \text{loss}}{\partial w_1} = \frac{\partial \text{loss}}{\partial y_2} \cdot \frac{\partial y_2}{\partial f} \cdot \frac{\partial f}{\partial w_1} \cdot \frac{\partial y_1}{\partial f} = w_1 \times b$$

\downarrow \downarrow \downarrow
 $-2(y_{\text{pred}} - y_2)$ $\begin{cases} 1 \\ 0 \end{cases}$ $\frac{\partial f}{\partial w_1}$

$$y_1 = X \cdot w_1 + b_1$$



Pytorch

Theano

Keras

Tensorflow

Glorot Uniform \rightarrow book page 393

\rightarrow Xavier Glorot & Yoshua Bengio

} Initialization of parameters

Am Freitag, 08.12.2024, zu Hause...

Verzweigung = Branching

Resume

Hyperparameter tuning

Vermeidung von Overfitting.

Underfitting & Overfitting

Kalibrierung der Wahrscheinlichkeit \sim Calibration curve
(sklearn)

Unbalancierte ~~Labels~~ Labels : Sampling Methoden

- └ Under-sampling
- └ Over-sampling
- └ Gewichtung der Klassen

Gesetz der großen Zahlen

Autoencoder \rightarrow book page 695 \leftarrow Regularisierung

Overfitting vermeiden:

- └ L1 & L2 Regularisierung, Book page 428-429, 187-193
- └ Drop-Out, Book page 429-432
- └ Early-Stopping, Book page 373, 193-194
- └ Optimierung der Netz-Größe, Book page 378-385

Keras-tuner

Wie kann L1/L2 Regularisierung Overfitting vermeiden?

- ↳ Die Beträge der Parameter im Modell werden reduziert
- ↳ Das Modell kann sich verstärkt auf wichtige Merkmale konzentrieren

Warum neigt ein Machine-Learning System zum Overfitting?

- ↳ Die Variabilität der Daten ist sehr groß

Was ist ein Indiz für Underfitting?

- ↳ Die Metriken für die Trainingsdaten liegen unter der Anforderung

Was kann bei Dropout beobachtet werden?

- ↳ Die Metriken der Trainingsdaten sind in den Lernkurven verfügbar

~~Was gibt~~

Was gibt es bei Verwendung von Early-Stopping mit der Option ~~restore_best_weights~~

~~Es~~ restore_best_weights = True zu beachten?

- ↳ Es kann zu einer Überanpassung an die Testdaten kommen

- ↳ Die Patience muss genau an das gegebene System angepasst werden.

Zu kurze Patience können die Testmetriken verschlechtern.