Universität Leipzig
Fakultät für Mathematik und Informatik
Institut für Informatik
Abteilung Datenbanken
Dr. B. Uhrich

Ausgabe: 13.11.2026
Besprechung: 25.11.2026

Physikbasiertes Maschinelles Lernen

Wintersemester 2026/27

—

Übungsblatt 3

Aufgabe 1: Architekturwahl und Modellannahmen

Ordnen Sie jeder Anwendung eine geeignete Architektur zu: vollständig verbundenes Netz (MLP), CNN, RNN/LSTM oder Graph-Netzwerk.

(a) Vorhersage einer Materialeigenschaft aus zehn skalaren Messgrößen.

(b) Erkennung räumlicher Strukturen in einem zweidimensionalen Temperaturfeld auf einem regelmäßigen Gitter.

(c) Vorhersage des nächsten Zeitschritts einer Messreihe mit zeitlichen Abhängigkeiten.

(d) Vorhersage von Knotengrößen auf einem unstrukturierten Finite-Elemente-Netz mit wechselnder Knotenanzahl.

Aufgabe 2: CNN: Dimensionen und Parameter

Ein Temperaturfeld $X \in^{32\times32\times1}$ wird durch folgende Schichten verarbeitet:

$$\text{Conv}_{3\times3,\,8}\ (p=1, s=1)\ \rightarrow\ \text{ReLU}\ \rightarrow\ \text{MaxPool}_{2\times2}\ (p=0, s=2)\ \rightarrow\ \text{Conv}_{3\times3,\,16}\ (p=0, s=1$$

Jede Faltung besitzt einen eigenen Bias je Ausgabekanal.

(a) Bestimmen Sie nach jeder Schicht die vollständige Tensorform. Geben Sie die allgemeine Formel für die räumliche Ausgabegröße einer Faltung an.

(b) Bestimmen Sie die Anzahl trainierbarer Parameter je Schicht und insgesamt. Warum besitzt Max-Pooling keine trainierbaren Parameter?

(c) Ersetzen Sie die erste Faltung durch eine vollständig verbundene Schicht, die aus dem vektorisierten Eingang ebenso viele skalare Ausgaben erzeugt. Wie viele Parameter benötigt diese Schicht einschließlich Bias?

Aufgabe 3: Faltung und diskrete Differentialoperatoren

Gegeben seien das skalare Feld X und der Kernel K:

$$X = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 2 & 3 \end{pmatrix}, \qquad K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(a) Berechnen Sie die vollständige Ausgabe für Padding $p = 0$ und Stride $s = 1$. Es gibt keinen Bias und keine Aktivierungsfunktion.

(b) Berechnen Sie die vollständige Ausgabe für $p = 0$ und $s = 2$. Welche Auswertungspositionen entfallen gegenüber Teil a)?

(c) Auf einem regelmäßigen Gitter mit Abstand Δx wird der Kernel

$$K_\Delta = \frac{1}{(\Delta x)^2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

verwendet. Leiten Sie aus den zentralen Differenzen für u_{xx} und u_{yy} her, welchen Differentialoperator er approximiert.

Aufgabe 4: Graph-Netzwerke und Graph-Laplace Operator

Ein ungerichteter, ungewichteter Graph besitzt die Knoten $V = \{1, 2, 3, 4\}$ und die Kanten

$$E = \{\{1, 2\}, \{2, 3\}, \{2, 4\}\}.$$

Die skalaren Anfangsmerkmale sind $h^{(0)} = (1, 2, 0, 3)^\top$. Selbstkanten werden in der Adjazenzmatrix A nicht berücksichtigt. Es gilt $A_{ij} = 1$, wenn die Knoten i und j verbunden sind, andernfalls $A_{ij} = 0$.

Der Graph-Laplace-Operator ist definiert durch

$$L = D - A, \qquad D = \mathrm{diag}(d_1, d_2, d_3, d_4), \qquad d_i = |\mathcal{N}(i)|.$$

$\mathcal{N}(i)$ bezeichnet die Nachbarschaft von Knoten i. Wir betrachten eine residuale Graph-Schicht mit Laplace-Term:

$$h^{(\ell+1)} = \mathrm{ReLU}\Big(h^{(\ell)} - \tau L h^{(\ell)}\Big), \qquad \tau = \tfrac{1}{4}.$$

ReLU wirkt komponentenweise. Alle Knoten werden gleichzeitig aus der vorherigen Schicht aktualisiert.

(a) Bestimmen Sie die Adjazenzmatrix A, die Gradmatrix D und den Graph-Laplace-Operator L. Prüfen Sie $L\mathbf{1} = 0$ mit $\mathbf{1} = (1, 1, 1, 1)^\top$.

(b) Zeigen Sie, dass die Aktualisierung für Knoten i äquivalent zu

$$h_i^{(\ell+1)} = \mathrm{ReLU}\left((1 - \tau d_i)h_i^{(\ell)} + \tau \sum_{j \in \mathcal{N}(i)} h_j^{(\ell)}\right)$$

(c) Berechnen Sie $Lh^{(0)}$, $h^{(1)}$ und $h^{(2)}$. Welche Knoten können nach zwei Schichten Knoten 1 beeinflussen? Begründen Sie dies anhand der Graphabstände.

(d) Interpretieren Sie den linearen Schritt $h^{(\ell+1)} = h^{(\ell)} - \tau L h^{(\ell)}$ als expliziten Euler-Schritt. Welche Differentialgleichung wird diskretisiert?

Aufgabe 5: RNN, LSTM, Attention Score

(a) Ein skalares RNN sei durch

$$a_t = w_h h_{t-1} + w_x x_t + b, \qquad h_t = \tanh(a_t), \qquad h_0 = 0$$

definiert. Es gelten $w_h = 0{,}5$, $w_x = 1$, $b = 0$ sowie $x_1 = 1$, $x_2 = 0$. Berechnen Sie h_1 und h_2 näherungsweise. Welche Parameter werden beim Entrollen des RNN über die Zeit geteilt?

(b) Ein LSTM verwendet unter anderem den Zellzustand

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t.$$

Erläutern Sie die Rollen von Forget-Gate f_t, Input-Gate i_t und Kandidatenzustand $\widetilde{c}_t$.

(c) Für eine einzelne Query seien folgende Größen gegeben:

$$Q = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad K = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad V = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 2 \end{pmatrix}, \qquad d_k = 2.$$

Die Zeilen von K und V gehören zu drei Positionen einer Sequenz. Berechnen Sie die Scores $s = qK^\top/\sqrt{d_k}$, die Gewichte $\alpha_j = e^{s_j}/\sum_{\ell=1}^{3} e^{s_\ell}$ und die Ausgabe $y = \alpha V$. Runden Sie auf drei Nachkommastellen.

Aufgabe 6: Faltung mit Tensorflow und PyTorch

Gegeben sind die Eingabematrix X und der Faltungskernel K aus dem Vorlesungsbeispiel:

$$X=\begin{pmatrix}0 & 1 & 2\\ 3 & 4 & 5\\ 6 & 7 & 8\end{pmatrix}, \qquad K=\begin{pmatrix}0 & 1\\ 2 & 3\end{pmatrix}.$$

(a) Implementieren Sie die Faltung von X mit K mit Schrittweite (s=1) und ohne Padding. Nutzen Sie hierfür entweder Tensorflow oder PyTorch. Geben Sie die Ausgabematrix aus.

(b) Erweitern Sie X mithilfe von Zero-Padding an allen vier Rändern um jeweils eine Schicht Nullen. Führen Sie anschließend dieselbe Faltung wie in a) mit Schrittweite (s=1) aus und geben die resultierende Ergebnismatrix aus.

(c) Erhöhen Sie die Schrittweite in beiden räumlichen Richtungen auf (s=2). Berechnen Sie die Faltung mit dem Zero-Padding aus b). Geben Sie die Ausgabematrix aus.