LEHRERINFO ZUM VIDEO

KI und maschinelles Lernen

FACH	STUFE	DAUER	VIDEO
Informatik	Gymnasium · Sek II	15 min	youtube.com/@Erklaerbaer

01 Lernziele

Nach dem Video können die Schülerinnen und Schüler …

- überwachtes, unüberwachtes und bestärkendes Lernen an Beispielen unterscheiden,
- Merkmale, k nächste Nachbarn und die Aufteilung in Trainings-, Validierungs- und Testdaten erklären,
- ein Perzeptron mit der Lernregel von Hand trainieren und seine Entscheidung als Gerade deuten,
- begründen, warum XOR eine verborgene Schicht braucht, und XOR aus ODER, NAND und UND bauen,
- Gradientenabstieg an einem Beispiel rechnen und Overfitting an Fehlerkurven erkennen,
- Genauigkeit, Präzision und Recall aus einer Konfusionsmatrix berechnen und je nach Einsatz gewichten.

02 Voraussetzungen

Trainingsdaten, Neuron, Sprachmodell, Bias (Folge 14 · Grundlagen); Logikgatter und XOR (Folge 4 · Advanced); Geradengleichung aus der Mathematik. Für Kapitel 9: SQL mit `JOIN` und `GROUP BY` (Folge 12 · Advanced).

03 Ablauf des Videos

ZEIT	KAPITEL	INHALT UND BEISPIEL
0:00	Einstieg	Minsky & Papert 1969: ein Neuron lernt XOR nicht · Ident · Fahrplan
0:37	Was du brauchst	Rückblick 14G, 04A, Geraden
0:57	Drei Arten zu lernen	Spam-Labels · Gruppen ähnlicher Songs · Belohnungen, Go 2016

1:48	Merkmale und Daten	Äpfel und Birnen nach Gewicht und Farbe · k = 3: 160 g, gelbgrün → Birne · 70/15/15 · Merksatz
2:50	Das Perzeptron	gewichtete Summe + Bias · UND mit w = (1, 1), b = −1,5 · Lernregel von Hand PAUSE-FRAGE b = −0,5 · Merksatz
4:40	XOR und die Schichten	nicht linear trennbar · ODER + NAND → UND = XOR · Sigmoid, ReLU · Merksatz
5:51	Gradientenabstieg und Overfitting	Fehlerlandschaft · Backpropagation · $F(w) = (w - 3)^2$ · Lernrate · Fehlerkurven, früher aufhören
7:05	Ein Modell bewerten	Spamfilter, 1000 Mails · Konfusionsmatrix · 96 %, 75 %, 90 % · Schwelle PAUSE-FRAGE medizinischer Suchtest
8:54	Deep Learning und Sprachmodelle	CNN: Kanten → Formen → Objekte · Transformer, Tokens, Aufmerksamkeit · Parameter · Rechenzentren
9:43	Grenzen und Verantwortung	Bias, Halluzination, Erklärbarkeit, Urheberrecht, Deepfakes, Datenschutz · AI Act · Betrugserkennung
10:43	Projekt-Baustein	Empfehlung in der Mediathek per SQL (Bau in 16A)
11:07	Prüfungsfallen	fünf Fallen (siehe 04)
12:05	Zusammenfassung	Kernaussagen · zurück zum Einstieg
12:51	Quiz	4 Fragen · Endkarte Folge 15

04 Prüfungsfallen

- Hohe Genauigkeit heisst gutes Modell – bei ungleich grossen Klassen täuscht sie; Präzision und Recall ansehen.
- Je kleiner der Trainingsfehler, desto besser – entscheidend ist der Fehler auf neuen Testdaten (Overfitting).
- Ein Perzeptron lernt jede logische Funktion – nur linear trennbare; XOR braucht eine verborgene Schicht.
- Testdaten darf man fürs Training mitbenutzen – dann misst der Test nichts mehr.
- Ein Sprachmodell schlägt Fakten nach – es erzeugt wahrscheinliche Tokens; zuverlässig nur mit angebundenen Quellen.

05 Einsatz im Unterricht

Perzeptron von Hand: Aufgabe 3 in Gruppen an der Tafel; jede Gruppe führt eine Zeile der Trainingstabelle.

Ausprobieren: Im TensorFlow Playground (playground.tensorflow.org) XOR mit null und mit einer verborgenen Schicht trainieren.

Konfusionsmatrix: Klassen-Experiment «Spam oder nicht?» mit 20 Beispielmails,

Diskussion: Wo ist Präzision wichtiger, wo Recall? Spamfilter, Brandmelder, Gesichtserkennung am Flughafen.

dann Präzision und Recall der Klasse ausrechnen.

Hinweis: Die Zahlen sind nachgerechnet. Perzeptron-Training für UND mit Lernrate 1, Start bei 0 und Ausgabe 1 bei Summe > 0: erster Fehler bei (1, 1) → w = (1, 1), b = 1; zweiter bei (0, 0) → b = 0; ab Durchgang 6 fehlerfrei mit w = (2, 1), b = −2. Gradientenabstieg: w ← w − 0,1 · 2(w − 3) ergibt 0 → 0,6 → 1,08 → 1,464. Spamfilter: (90 + 870) / 1000 = 96 %, 90 / 120 = 75 %, 90 / 100 = 90 %. Pause-Frage: Recall 45 / 50 = 90 %, Präzision 45 / 145 ≈ 31 %. Die Wahrscheinlichkeiten fürs nächste Token und die Ausleihzahlen der Mediathek sind erfunden. Der EU AI Act ist seit August 2024 in Kraft; seine Pflichten gelten gestaffelt (Stand der Termine bei der EU nachsehen).

06 Quellen

- F. Rosenblatt (1958): The Perceptron, Psychological Review 65(6) · M. Minsky, S. Papert (1969): Perceptrons, MIT Press
- D. Rumelhart, G. Hinton, R. Williams (1986): Learning representations by back-propagating errors, Nature 323
- A. Vaswani u. a. (2017): Attention Is All You Need, NeurIPS · D. Silver u. a. (2016): Mastering the game of Go, Nature 529
- Verordnung (EU) 2024/1689 (KI-Verordnung, «AI Act»), Amtsblatt der EU, 12.7.2024
- TensorFlow Playground – playground.tensorflow.org · Lehrplan 21 / Rahmenlehrplan Informatik Gymnasium

NAME _____ KLASSE _____ DATUM _____

KI und maschinelles Lernen

1 Lernarten ●○○

Ordne zu (überwacht, unüberwacht, bestärkend): a) Ein Shop sortiert seine Kundschaft in Gruppen mit ähnlichem Kaufverhalten. b) Ein Roboter lernt Laufen und bekommt Punkte für jeden Schritt ohne Sturz. c) Ein Modell lernt aus 10 000 Röntgenbildern mit der Diagnose dazu.

2 k nächste Nachbarn ●○○

Eine neue Frucht hat diese fünf nächsten Nachbarn, nach Abstand sortiert: Apfel, Birne, Birne, Apfel, Apfel. **Wie** wird sie bei k = 1, k = 3 und k = 5 eingeordnet?

3 Perzeptron ●●○

Ein Perzeptron hat $w_1 = 1$, $w_2 = 1$ und $b = 0$ (Ausgabe 1 bei Summe > 0). Es soll UND lernen. a) **Welche** Eingaben gibt es falsch aus? b) **Wende** die Lernregel (Lernrate 1) auf den ersten Fehler in der Reihenfolge (0,0), (0,1), (1,0), (1,1) an.

4 XOR ●●○

Zeige mit einer Tabelle, dass UND(ODER(x_1, x_2), NAND(x_1, x_2)) für alle vier Eingaben XOR ergibt, und **erkläre**, warum ein einzelnes Perzeptron das nicht kann.

5 Gradientenabstieg

●●●

$F(w) = (w - 3)^2$, Steigung $2 \cdot (w - 3)$. **Rechne** zwei Schritte ab $w = 5$ mit der Lernrate 0,25 und dann zwei Schritte mit der Lernrate 1,1. **Was** fällt auf?

6 Modell bewerten

●●○

Ein Bilderkenner soll Katzen finden. Von 200 Bildern zeigen 40 eine Katze. Er markiert 50 Bilder als Katze, davon sind 36 richtig. **Berechne** Genauigkeit, Präzision und Recall.

Lösungen

1 Lernarten

a) unüberwacht (Gruppen ohne Labels) · b) bestärkend (Belohnung) · c) überwacht (Bilder mit Label).

2 k nächste Nachbarn

k = 1: Apfel · k = 3: Birne (2 Birnen gegen 1 Apfel) · k = 5: Apfel (3 Äpfel gegen 2 Birnen). Darum stellt man k mit den Validierungsdaten ein.

3 Perzeptron

a) (0,1) und (1,0): Summe 1 > 0, also Ausgabe 1 statt 0; (0,0) gibt 0 und (1,1) gibt 1, beides richtig. b) Erster Fehler bei (0,1): Soll − Ist = 0 − 1 = −1 → $w_1 = 1 + (-1) \cdot 0 = 1$, $w_2 = 1 + (-1) \cdot 1 = 0$, $b = 0 + (-1) = -1$.

4 XOR

(0,0): ODER 0, NAND 1 → UND 0 · (0,1): 1, 1 → 1 · (1,0): 1, 1 → 1 · (1,1): 1, 0 → 0, also 0, 1, 1, 0 = XOR. Ein Perzeptron zieht nur eine Gerade; die beiden Einsen liegen diagonal gegenüber, keine Gerade trennt sie von den Nullen.

5 Gradientenabstieg

Lernrate 0,25: $w = 5 - 0{,}25 \cdot 4 = 4 \rightarrow 4 - 0{,}25 \cdot 2 = 3{,}5$ (nähert sich 3). Lernrate 1,1: $w = 5 - 1{,}1 \cdot 4 = 0{,}6 \rightarrow 0{,}6 - 1{,}1 \cdot (-4{,}8) = 5{,}88$ (springt über das Tal und entfernt sich immer weiter: zu grosse Lernrate).

6 Modell bewerten

richtig positiv 36, falsch positiv 14, falsch negativ 4, richtig negativ 146 · Genauigkeit (36 + 146) / 200 = 91 % · Präzision 36 / 50 = 72 % · Recall 36 / 40 = 90 %.

Q Quiz im Video

1 B (bestärkendes Lernen) · 2 A (nicht linear trennbar) · 3 C (auswendig gelernt) · 4 A (Anteil der gefundenen echten Positiven)