

■ KIKON – Lernübersicht Kurzarbeit (10.11.2025)

1. Grundbegriffe

Begriff	Bedeutung	Beispiel
Feature (Merkmal)	Messbare Eigenschaft eines Objekts	Farbe, Größe, Pixelwert
Datenpunkt	Vollständige Sammlung aller Merkmale	Ein Bild, eine Tabelle
Label	Zielwert beim Training	Katze, Hund, defekt
Klasse	Gruppe aller möglichen Labels	Tiere: Katze, Hund, Pferd
Klassifikation	Zuordnung eines Datenpunkts zu einer Klasse	Dieses Bild ist eine Katze
Regression	Vorhersage eines Zahlenwerts	Preis, Temperatur
Batch	Anzahl Daten pro Schritt	Batchgröße = 32
Epoche	1 Durchlauf aller Daten	10.000 Bilder × 5 Epochen

2. Der KI-Zyklus (6 Stufen)

Stufe	Beschreibung
1. Problemstellung	Ziel definieren (z. B. Bilder richtig erkennen)
2. Datensammlung	Daten von Quellen wie Kaggle, Roboflow, Sensoren
3. Daten prüfen	Qualität und Ausgewogenheit der Daten bewerten
4. Modellauswahl + Training	Algorithmus auswählen und trainieren
5. Test / Evaluation	Modellleistung prüfen (z. B. Konfusionsmatrix)
6. Ausrollen	Modell produktiv einsetzen und weiter trainieren

3. Prüfkriterien für gute Datensätze

- Haben wir ausreichend Daten von jeder Klasse?
- Sind die Daten innerhalb einer Klasse verschieden genug?
- Stimmen Helligkeit, Farben, Hintergrund?
- Fehler und Ausreißer bereinigen → nur saubere Daten führen zu gutem Modell.

4. KI-Familien (Lernarten)

Lernart	Beschreibung	Typische Verfahren	Beispiele
Überwachtes Lernen	Eingabedaten + Labels, Feedback	Klassifikation, Regression	Spamfilter, Preisvorhersage
Unüberwachtes Lernen	Keine Labels, Mustererkennung	Clustering, Dimensionsreduktion	Kundengruppen, Anomalieerkennung
Verstärkendes Lernen	Belohnung und Bestrafung	Q-Learning, Deep RL	Robotik, autonomes Fahren

5. KI-Domänen

Domäne	Typische Daten	Kernaufgaben	Beispiele
Computer Vision	Bilder, Videos	Objekterkennung, Segmentierung	Gesichtserkennung, Verkehrszeichen
NLP	Texte, Sprache	Sprachverständnis, Übersetzung	Chatbots, Übersetzer
Data / Analytics	Tabellen, Sensorwerte	Vorhersagen, Mustererkennung	Preis- und Nachfrageprognosen

6. Evaluationsmethoden

Realität / Test	Positiv vorhergesagt	Negativ vorhergesagt
Positiv (tatsächlich)	TP (True Positive)	FN (False Negative)
Negativ (tatsächlich)	FP (False Positive)	TN (True Negative)

Formeln (auswendig können):

- $\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$
- $\text{Precision} = TP / (TP + FP)$
- $\text{Recall} = TP / (TP + FN)$
- Recall wichtig bei Krankheitserkennung (keine Positiven übersehen)
- Precision wichtig bei Fehlalarmvermeidung (z. B. Spamfilter)

7. Overfitting vs. Underfitting

Typ	Beschreibung	Merkmale	Gegenmaßnahmen
Underfitting	Lernt zu wenig	Hoher Train- & Testfehler	Komplexeres Modell, längeres Training, mehr Features
Overfitting	Lernt auswendig	Niedriger Trainingsfehler, hoher Testfehler	Mehr Daten, Dropout, Regularisierung, Early Stopping
Gutes Modell	Gleichgewicht	Train & Test niedrig	–

8. Epochen & Batchgröße

Begriff	Bedeutung	Wirkung
Epoche	1 vollständiger Durchlauf durch alle Trainingsdaten	Zu wenige → Underfitting, zu viele → Overfitting
Batchgröße	Anzahl Daten pro Schritt	Kleine Batchgröße = genauer, große Batchgröße = schneller

9. Merksätze

- Feature = Input, Label = gewünschte Ausgabe, Datenpunkt = Beispiel
- KI-Zyklus: Problem → Daten → Prüfen → Trainieren → Testen → Ausrollen
- Überwacht / Unüberwacht / Verstärkend → unterscheiden & erklären
- Precision / Recall / Accuracy → Formeln + Bedeutung
- Underfitting = zu wenig gelernt, Overfitting = zu viel gelernt
- Epochen = wie oft der Datensatz trainiert wird
- Batchgröße = wie viele Daten gleichzeitig verarbeitet werden
- Recall wichtig, wenn nichts übersehen werden darf
- Precision wichtig, wenn Fehlalarme teuer sind