

Die hier aufgelisteten Inhalte der Data-Science-II-Veranstaltungen sollen die Prüfungsvorbereitung erleichtern. Diese Zusammenstellung dient lediglich der Orientierung über die Schwerpunktsetzung im zweiten, theoretischen Teil der Prüfung. Nicht enthaltene Themen können trotzdem in der Prüfung eine Rolle spielen.

1. Überblick über Visualisierungsbibliotheken (Matplotlib, Seaborn, Cartopy, Folium, Plotly). Welche Bibliothek für welche Aufgaben?
2. Problemstellung und allgemeiner Lösungsansatz beim überwachten Lernen (Parameter vs. Hyperparameter, Train-Test-Split, usw.).
3. Funktionsweise von k-NN.
4. Metriken zur Bewertung der Vorhersagequalität von Modellen (nur Bedeutung/Interpretation, keine Formeln).
5. PCA (Grundidee, Wozu?, keine Formeln).
6. Kreuzvalidierung (Cross-Validation) zur Hyperparameteroptimierung.
7. Lineare Regression mit nichtlinearen Funktionen (Definition, Vorteile/Nachteile, Regularisierung).
8. RANSAC (Grundidee und Schritte zur Umsetzung, keine Formeln oder Details).
9. Grundlagen der künstlichen neuronalen Netze (Aufbau von Neuronen, geschichtete Netzstruktur, Training mit Gradientenverfahren).
10. Faltende neuronale Netze (Grundidee, Struktur, Vorteile)

Projekte für Prüfungsvorleistung

Zur Erfüllung der Prüfungsvorleistung ist der funktionsfähige Quellcode zu einem der folgenden Projekte aus dem Praktikum abzugeben. Der Quellcode wird nicht bewertet, sondern dient ausschließlich als Grundlage für die Diskussion des Projekts im ersten Teil der mündlichen Prüfung.

- QMNIST Feature Reduction (week 7)
- PCA and ANN for QMNIST (week 11)
- CNN for QMNIST (week 12)