



Digital Applications & Data Management

WS25/26

Dr. Jens Kohl



Roadmap Vorlesung

1. Grundlagen Künstlicher Intelligenz, Machine Learning und Daten
2. Grundlagen Data Science
3. Angewandte Data Science (Teil 1)
4. Angewandte Data Science (Teil 2)
5. Data Science Use Case
6. Grundlagen unüberwachtes Lernen
7. Grundlagen überwachtes Lernen (tabellarische Daten)
8. Machine Learning Use Case
9. Grundlagen überwachtes Lernen (Bilddaten)
10. Transfer Learning Bilddaten und Fallbeispiel
11. Grundlagen Generative AI
12. Prompt Engineering, Agenten
13. Ausblick, Wiederholung, Fragestunde
14. Fragestunde



Vorlesung 5:

Data Science Use Case



Was machen wir heute?

Motivation

Umsetzen eines Data Science Use Cases (in Gruppenarbeit/ selbstständig)



Data Science Use Case

Aufgabenstellung „Kreditvergabe“

1. **Data Engineering und cleaning:** Laden Sie den Datensatz “Data Science Use Case Loan.csv” in einen Dataframe und verwenden Sie dabei folgende Spalten

Loan_ID	Geschlecht	Verheiratet	Nutzungsintensität TikTok	Bildung	Selbstständig	Einkommen	Darlehenshöhe	Darlehensdauer (Monate)	Darlehenshistorie (ja/ nein)	Entscheidung Kreditvergabe	Lage des Objekts
---------	------------	-------------	------------------------------	---------	---------------	-----------	---------------	----------------------------	---------------------------------	-------------------------------	------------------

2. **Feature engineering:** erstellen Sie eine neue Spalte “Darlehenshöhe in Klassen”.
Ergeben sich damit weitere/ neue Erkenntnisse?
3. **Deskriptive Analyse:** was fällt Ihnen auf? Welche Spalten sind relevant?
4. **Explorative Analyse:** welche Erkenntnisse können Sie gewinnen? Welche Eigenschaften sind für eine Kreditvergabe eher oder weniger relevant? Nutzen Sie matplotlib/ pygwalker
5. **Outlier detection und deletion:** in welcher Spalte sind Ausreißer vorhanden? Was passiert wenn Sie diese Ausreißer entfernen?
6. **Unsupervised Learning:** Clustern Sie Datensatz? Was ist die ideale Anzahl Cluster?
Was für Erkenntnisse können Sie daraus gewinnen?
7. **Supervised Learning:** erstellen Sie ein Modell für die Vorhersage einer Kreditvergabe
Wie hoch ist die Genauigkeit? Was sind die wichtigsten Eigenschaften für Vorhersage?

Nächste
Stunde

Übernächste
Stunde