



Digital Applications & Data Management

WS25/26

Dr. Jens Kohl



Roadmap Vorlesung

1. Einführung und Übersicht
2. Grundlagen Data Science
3. Vorgehen Data Science Use Case
4. Case Study Data Science
5. Grundlagen unüberwachtes Lernen
6. Grundlagen überwachtes Lernen
(tabellarische Daten)
7. Case Study überwachtes Lernen
(tabellarische Daten)
8. Grundlagen überwachtes Lernen (Bilddaten)
9. Case Study überwachtes Lernen und Transfer Learning (Bilddaten)
10. Grundlagen Generative AI
11. Generative AI mit Texten und Prompt Engineering
12. Agentic AI
13. Ausblick: Machine Learning in der Cloud und Reinforcement Learning



Vorlesung 4: Case Study Data Science



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	Darlehendsdauer (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?

Nächste Stunde
7. **Supervised Learning:** erstellen Sie ein Modell für die Vorhersage einer Kreditvergabe Wie hoch ist die Genauigkeit? Was sind die wichtigsten Eigenschaften für Vorhersage?

Übernächste Stunde