



# Digital Applications & Data Management

WS25/26

Modul 11

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



# **8. KÜNSTLICHE INTELLIGENZ: ÜBERWACHTES (SUPERVISED) LERNEN**

## **Case Study**



# Was machen wir heute?

Motivation

Umsetzen eines Data Supervised Learning Use Cases (in Gruppenarbeit/ selbstständig)

# Supervised Learning – Case study Direktmarketing



## Beschreibung

- Die Daten beziehen sich auf Direktmarketing-Kampagnen (Telefonanrufe) eines portugiesischen Bankinstituts. Das Klassifizierungsziel ist die Vorhersage, ob der Kunde eine FESTGELD-ANLAGE TÄTIGEN wird (Variable  $y$ ).
- Dabei war oft mehr als ein Kontakt mit demselben Kunden erforderlich, um zu erfahren, ob das Produkt (Festgeld-ANLAGE) ANGENOMMEN wird ('ja') oder nicht ('nein').
- Datensatz Direktmarketing.csv mit 41188 ZEILEN und 21 SPALTEN, geordnet nach Datum (von Mai 2008 bis November 2010).
- Ziel ist, Algorithmen für maschinelles Lernen zu testen um Vorhersagen zu treffen, ob ein neuer Kunde einen Vertrag unterzeichnet oder nicht.



# Supervised Learning – Case study Direktmarketing

## Übersicht Aufgabenstellung

1. Data Engineering und Cleaning (als Vorbereitung ML)
  - Umwandeln der Features in Zahlenbasierte Datentypen
  - Beseitigen leerer Werte und Nicht-Zahlenwerte
  - Ggf. löschen nicht relevanter Spalten
2. Vorbereitung maschinelles Lernen
  - Aufteilen in X-Matrix und y-Vektor
  - Aufteilen in Trainings- und Testmenge
3. Anwenden von 2 ML-Verfahren inkl. Hyperparametertuning
  - Decision Tree/ Logistische Regression
  - Random Forest
4. Visuelle Darstellung der wichtigsten Features