

Stundenplan CAS Data Science FS 2022 - Klasse 2

Tag	Datum	Raum	Zeit	Unterrichtsblock	Themen
1	Freitag, 25.2.2022		08:45 - 16:30	Einleitung zum CAS DS Kurs Einführung in die Thematik von Data Science	<ul style="list-style-type: none"> Begrüssung und Vorstellung der Teilnehmenden Diskussion der Lernziele und der Gesamtplanung Wieso ist Data Science wichtig geworden? Einführung der Begriffe Data Science und Big Data. Wie sind sie gebunden? Herausforderungen, Limitierungen, Chancen und Risiken Grundaufbau einer Data Science Lösung: Ansätze und Vorgehen Beispiele aus Forschung und Wirtschaft
2	Samstag, 26.2.2022	online	08:45 - 12:15 13:00 - 16:30	Statistik	<ul style="list-style-type: none"> Deskriptive Statistik Wahrscheinlichkeitsverteilungen Hypothesentests
3	Freitag, 4.3.2022		08:45 - 12:15	SQL	<ul style="list-style-type: none"> Relationales Datenbankmodell Abfrage und Gruppierung von Daten aus einer oder mehreren Tabellen Abfrageoptimierungen
			13:00 - 16:30	Lineare Algebra	<ul style="list-style-type: none"> Vektoren und ihre geometrische Bedeutung Matrizen als mathematische Operatoren Lineare Abbildungen und ihre geometrische Bedeutung Eigenwerte, Eigenvektoren und Diagonalisierung symmetrischer Matrizen <i>Auseinandersetzung mit dem Numpy-Paket für Python</i>
4	Samstag, 5.3.2022	online	08:45 - 12:15 13:00 - 16:30	Data Wrangling	<ul style="list-style-type: none"> Datentypen und Datenformate (strukturierte, unstrukturierte & geographische Daten) Ausreisser-Analyse Herausforderungen bei der Anonymisierung von Daten Umgang mit unvollständigen / unbalancierten Daten Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn
5	Freitag, 18.3.2022		08:45 - 12:15 13:00 - 16:30	Information Visualization	<ul style="list-style-type: none"> Einleitung: Wofür visualisieren? Daten und grundlegende Grafiken Grammar of Graphics Visuelle Wahrnehmung und Informationsdesign Interaktive Visualisierungssysteme Übung mit Jupyter notebooks, vega-lite
6	Samstag, 19.3.2022		08:45 - 12:15 13:00 - 16:30	Data Wrangling	<ul style="list-style-type: none"> Anreicherung von Daten Erstellen von Daten Pipelines Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn
7	Freitag, 25.3.2022		08:45 - 12:15 13:00 - 16:30	Data Mining	<ul style="list-style-type: none"> Einführung Data Mining <ul style="list-style-type: none"> Was bringt Data Mining? Die Aufwertungskette des Data Minings. Data Mining an Objektmengen – Market Basket Analysis Regeln suchen Data Mining Modelle <i>Übungen zu den vorgestellten Verfahren</i>
8	Samstag, 26.3.2022		08:45 - 12:15 13:00 - 16:30	Regression	<ul style="list-style-type: none"> Einfache lineare Regression Residuenanalyse Multiple lineare Regression Sensitivität und Robustheit Variablenselektion und Modellierung
9	Samstag, 2.4.2022		08:45 - 12:15 13:00 - 16:30	Grundlagen und Terminologie Machine Learning	<ul style="list-style-type: none"> Grundbegriffe Eigenschaften von Daten : Daten laden und visualisieren Metriken zur Evaluation der Güte von ML Modellen Kostenfunktion und Optimierung Modell-Selektion Lineare Regression, Kernel Regression Übungen mit Jupyter notebooks, pandas, matplotlib, seaborn, scikit-learn
10	Freitag, 8.4.2022		08:45 - 12:15 13:00 - 16:30	Data Product Definition & Architektur	<ul style="list-style-type: none"> Vom Business-Case zum Data Science Problem
11	Samstag, 9.4.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning Lab	Anwenden von Machine Learning an bereitgestellten Datensätze im Übungsbetrieb. (kein Frontalunterricht, jedoch fachliche Unterstützung) <ul style="list-style-type: none"> Machine Learning Aufgaben zu Regression, Klassifikation, Clustering Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn
12	Freitag, 29.4.2022	online	08:45 - 09:30	Prüfung 1	Prüfung über die Themen: <ul style="list-style-type: none"> Statistik & Regression SQL Lineare Algebra Data Wrangling Data Mining Information Visualization

			10:00 - 12:15 13:00 - 16:30	Graphen Analyse	<ul style="list-style-type: none"> • Modellierung, Influencer & Communities entdecken, Link-Prediction, Statistik in Netzwerken • Übungen mit Gephi, python (pandas, networkX)
13	Samstag, 30.4.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning	<ul style="list-style-type: none"> • Methoden zur Klassifikation • Logistische Regression, Naïve Bayes • Decision Trees, Random Forests • Support Vector Machines • Methoden zur Dimensionsreduktion: PCA, NMF • Clustering: KMeans • Übungen mit Jupyter notebooks, pandas, matplotlib, seaborn, scikit-learn
14	Freitag, 6.5.2022		08:45 - 12:15 13:00 - 16:30	Recommender Systems	<ul style="list-style-type: none"> • Klassische Ansätze und Umsetzung für Recommender-Systems (u.a. Collaborative Filtering, Content-based Filtering, Hybrid-Ansätze) • <i>Übungen mit Jupyter notebook, Google collab</i>
14	Freitag, 13.5.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning	<ul style="list-style-type: none"> • Einführung in Neuronale Netzwerke • Multi-Layer Perceptron • Deep Learning Konzepte und Architekturen • Übersicht ML Projekte in der Praxis • Übungen mit Jupyter notebooks, pandas, matplotlib, seaborn, scikit-learn, keras
15	Samstag, 14.5.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning Lab	<p>Anwenden von Machine Learning an bereitgestellten Datensätze im Übungsbetrieb sowie das begleitete Arbeiten an der Projektarbeit (kein Frontalunterricht, jedoch fachliche Unterstützung)</p> <ul style="list-style-type: none"> • <i>Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn</i>
16	Freitag, 20.5.2022	online	08:45 - 12:15 13:00 - 16:30	Information Retrieval & Natural Language Processing	<ul style="list-style-type: none"> • Schwierigkeiten und Grenzen, Arten der Repräsentation • Textvorverarbeitung (Tokenization, Stemming, Normalization, Eigennamenerkennung) • Reguläre Ausdrücke • Übungen mit Jupyter notebook
17	Samstag, 21.5.2022	online	08:45 - 12:15 13:00 - 16:30	Information Retrieval & Natural Language Processing	<ul style="list-style-type: none"> • Suchmaschinen (BOW Modell, VectorSpace Model, Elasticsearch) • Textklassifizierung • Language Models (word2vec, ...) • <i>Übungen mit Jupyter notebook, Elasticsearch server, Google collab</i>
19	Samstag, 11.6.2022		08:45 - 12:15 13:00 - 16:30	NLP Lab	<p>Anwenden von NLP an bereitgestellten Datensätze im Übungsbetrieb sowie das begleitete Arbeiten an der Projektarbeit (kein Frontalunterricht, jedoch fachliche Unterstützung)</p>
20	Freitag, 17.6.2022		09:00 - 11:00	Prüfung 2	<p>Prüfung über die Themen:</p> <ul style="list-style-type: none"> • Data Product Definition • Machine Learning • IR & NLP • Graphen Analyse • Recommender Systems
21	Samstag, 2.7.2022	online	08:45 - 12:15 13:00 - 16:30	Projektabschluss & Ausblick	<ul style="list-style-type: none"> • <i>Abschluss & Präsentation der praktischen Projektarbeit</i>