

### Stundenplan CAS Data Science FS 2022 - Klasse 1

Tag	Datum	Raum	Zeit	Unterrichtsblock	Themen
1	Freitag, 18.2.2022		08:45 - 16:30	Einleitung zum CAS DS Kurs Einführung in die Thematik von Data Science	<ul style="list-style-type: none"> <li>Begrüssung und Vorstellung der Teilnehmenden</li> <li>Diskussion der Lernziele und der Gesamtplanung</li> <li>Wieso ist Data Science wichtig geworden?</li> <li>Einführung der Begriffe Data Science und Big Data. Wie sind sie gebunden?</li> <li>Herausforderungen, Limitierungen, Chancen und Risiken</li> <li>Grundaufbau einer Data Science Lösung: Ansätze und Vorgehen</li> <li>Beispiele aus Forschung und Wirtschaft</li> </ul>
2	Samstag, 19.2.2022		08:45 - 12:15 13:00 - 16:30	Statistik	<ul style="list-style-type: none"> <li>Deskriptive Statistik</li> <li>Wahrscheinlichkeitsverteilungen</li> <li>Hypothesentests</li> </ul>
3	Freitag, 25.2.2022		08:45 - 12:15	SQL	<ul style="list-style-type: none"> <li>Relationales Datenbankmodell</li> <li>Abfrage und Gruppierung von Daten aus einer oder mehreren Tabellen</li> <li>Abfrageoptimierungen</li> </ul>
			13:00 - 16:30	Lineare Algebra	<ul style="list-style-type: none"> <li>Vektoren und ihre geometrische Bedeutung</li> <li>Matrizen als mathematische Operatoren</li> <li>Lineare Abbildungen und ihre geometrische Bedeutung</li> <li>Eigenwerte, Eigenvektoren und Diagonalisierung symmetrischer Matrizen</li> <li><i>Auseinandersetzung mit dem Numpy-Paket für Python</i></li> </ul>
4	Samstag, 26.2.2022		08:45 - 12:15 13:00 - 16:30	Data Wrangling	<ul style="list-style-type: none"> <li>Datentypen und Datenformate (strukturierte, unstrukturierte &amp; geographische Daten)</li> <li>Ausreisser-Analyse</li> <li>Herausforderungen bei der Anonymisierung von Daten</li> <li>Umgang mit unvollständigen / unbalancierten Daten</li> <li>Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn</li> </ul>
5	Samstag, 5.3.2022		08:45 - 12:15 13:00 - 16:30	Regression	<ul style="list-style-type: none"> <li>Einfache lineare Regression</li> <li>Residuenanalyse</li> <li>Multiple lineare Regression</li> <li>Sensitivität und Robustheit</li> <li>Variablenselektion und Modellierung</li> </ul>
6	Freitag, 11.3.2022		08:45 - 12:15 13:00 - 16:30	Information Visualization	<ul style="list-style-type: none"> <li>Einleitung: Wofür visualisieren?</li> <li>Daten und grundlegende Grafiken</li> <li>Grammar of Graphics</li> <li>Visuelle Wahrnehmung und Informationsdesign</li> <li>Interaktive Visualisierungssysteme</li> <li>Übung mit Jupyter notebooks, vega-lite</li> </ul>
7	Samstag, 12.3.2022		08:45 - 12:15 13:00 - 16:30	Data Wrangling	<ul style="list-style-type: none"> <li>Anreicherung von Daten</li> <li>Erstellen von Daten Pipelines</li> <li>Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn</li> </ul>
8	Freitag, 18.3.2022		08:45 - 12:15 13:00 - 16:30	Data Mining	<ul style="list-style-type: none"> <li>Einführung Data Mining <ul style="list-style-type: none"> <li>Was bringt Data Mining?</li> <li>Die Aufwertungskette des Data Minings.</li> </ul> </li> <li>Data Mining an Objektmengen – Market Basket Analysis</li> <li>Regeln suchen</li> <li>Data Mining Modelle</li> <li><i>Übungen zu den vorgestellten Verfahren</i></li> </ul>
9	Freitag, 25.3.2022		08:45 - 12:15 13:00 - 16:30	Grundlagen und Terminologie Machine Learning	<ul style="list-style-type: none"> <li>Grundbegriffe</li> <li>Eigenschaften von Daten : Daten laden und visualisieren</li> <li>Metriken zur Evaluation der Güte von ML Modellen</li> <li>Kostenfunktion und Optimierung</li> <li>Modell-Selektion</li> <li>Lineare Regression, Kernel Regression</li> <li>Übungen mit Jupyter notebooks, pandas, matplotlib, seaborn, scikit-learn</li> </ul>
10	Freitag, 1.4.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning Lab	<ul style="list-style-type: none"> <li>Anwenden von Machine Learning an bereitgestellten Datensätze im Übungsbetrieb. (kein Frontalunterricht, jedoch fachliche Unterstützung)</li> <li>Machine Learning Aufgaben zu Regression, Klassifikation, Clustering</li> <li>Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn</li> </ul>
11	Samstag, 2.4.2022		08:45 - 12:15 13:00 - 16:30	Data Product Definition & Architektur	<ul style="list-style-type: none"> <li>Vom Business-Case zum Data Science Problem</li> </ul>
12	Freitag, 29.4.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning	<ul style="list-style-type: none"> <li>Methoden zur Klassifikation</li> <li>Logistische Regression, Naive Bayes</li> <li>Decision Trees, Random Forests</li> <li>Support Vector Machines</li> <li>Methoden zur Dimensionsreduktion: PCA, NMF</li> <li>Clustering: KMeans</li> <li>Übungen mit Jupyter notebooks, pandas, matplotlib, seaborn, scikit-learn</li> </ul>

13	Samstag, 30.4.2022		08:45 - 09:30	Prüfung 1	Prüfung über die Themen: <ul style="list-style-type: none"> <li>• Statistik &amp; Regression</li> <li>• SQL</li> <li>• Lineare Algebra</li> <li>• Data Wrangling</li> <li>• Data Mining</li> <li>• Information Visualization</li> </ul>
			10:00 - 12:15 13:00 - 16:30	Graphen Analyse	<ul style="list-style-type: none"> <li>• Modellierung, Influencer &amp; Communities entdecken, Link-Prediction, Statistik in Netzwerken</li> <li>• Übungen mit Gephi, python (pandas, networkX)</li> </ul>
14	Freitag, 6.5.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning	<ul style="list-style-type: none"> <li>• Einführung in Neuronale Netzwerke</li> <li>• Multi-Layer Perceptron</li> <li>• Deep Learning Konzepte und Architekturen</li> <li>• Übersicht ML Projekte in der Praxis</li> <li>• Übungen mit Jupyter notebooks, pandas, matplotlib, seaborn, scikit-learn, keras</li> </ul>
15	Samstag, 7.5.2022		08:45 - 12:15 13:00 - 16:30	Machine Learning Lab	Anwenden von Machine Learning an bereitgestellten Datensätze im Übungsbetrieb sowie das begleitete Arbeiten an der Projektarbeit (kein Frontalunterricht, jedoch fachliche Unterstützung) <ul style="list-style-type: none"> <li>• <i>Übungen mit Jupyter notebooks, pandas, matplotlib, scikit-learn</i></li> </ul>
16	Freitag, 13.5.2022		08:45 - 12:15 13:00 - 16:30	Information Retrieval & Natural Language Processing	<ul style="list-style-type: none"> <li>• Schwierigkeiten und Grenzen, Arten der Repräsentation</li> <li>• Textvorverarbeitung (Tokenization, Stemming, Normalization, Eigennamenerkennung)</li> <li>• Reguläre Ausdrücke</li> <li>• Übungen mit Jupyter notebook</li> </ul>
17	Samstag, 14.5.2022		08:45 - 12:15 13:00 - 16:30	Information Retrieval & Natural Language Processing	<ul style="list-style-type: none"> <li>• Suchmaschinen (BOW Modell, VectorSpace Model, Elasticsearch)</li> <li>• Textklassifizierung</li> <li>• Language Models (word2vec, ...)</li> <li>• <i>Übungen mit Jupyter notebook, Elasticsearch server, Google collab</i></li> </ul>
18	Samstag, 21.5.2022		08:45 - 12:15 13:00 - 16:30	Recommender Systems	<ul style="list-style-type: none"> <li>• Klassische Ansätze und Umsetzung für Recommender-Systems (u.a. Collaborative Filtering, Content-based Filtering, Hybrid-Ansätze)</li> <li>• <i>Übungen mit Jupyter notebook, Google collab</i></li> </ul>
19	Freitag, 3.6.2022		08:45 - 12:15 13:00 - 16:30	NLP Lab	Anwenden von NLP an bereitgestellten Datensätze im Übungsbetrieb sowie das begleitete Arbeiten an der Projektarbeit (kein Frontalunterricht, jedoch fachliche Unterstützung)
20	Samstag, 18.6.2022		09:00 - 11:00	Prüfung 2	Prüfung über die Themen: <ul style="list-style-type: none"> <li>• Data Product Definition</li> <li>• Machine Learning</li> <li>• IR &amp; NLP</li> <li>• Graphen Analyse</li> <li>• Recommender Systems</li> </ul>
21	Freitag, 1.7.2022		08:45 - 12:15 13:00 - 16:30	Projektabschluss & Ausblick	<ul style="list-style-type: none"> <li>• <i>Abschluss &amp; Präsentation der praktischen Projektarbeit</i></li> </ul>