

Modulbeschreibung Text Analytics

Autor/in:	Albert Weichselbraun
Ausgabestelle:	Schweizerisches Institut für Informationswissenschaft (SII)
Geltungsbereich:	Studiengang MSc User Experience Design & Data Visualization Studienrichtung Data Visualization
Klassifizierung:	Intern
Version:	V02.00
Ausgabedatum:	01.09.2025

Modul

Name	Text Analytics		
Kürzel	TA	ECTS-Punkte	4
Typ	Pflichtmodul		
Verantwortliche/r	Prof. Dr. Albert Weichselbraun		
Leitidee	Dieses Modul bietet eine Einführung in das Natural Language Processing und stellt relevante Prinzipien, Methoden und Werkzeuge vor. Die Studierenden werden mit grundlegenden Konzepten zur Beschreibung der Struktur von Sprache und Schritten zu deren maschinellen Bearbeitung vertraut gemacht. Anschliessend werden ausgewählte Methoden des Natural Language Processings besprochen und auf aktuelle Problemstellungen angewandt. Es werden fortschrittliche Methoden diskutiert, welche für die Wissensextraktion aus unstrukturierten Texten zur Anwendung kommen können. Ziel ist es, dass die Studierenden einen Überblick über aktuelle Entwicklungen erhalten und das vermittelte Wissen auf praktische Aufgabenstellungen anwenden können.		
Voraussetzungen	Programming Python		
Lernergebnisse	<p>Nach Abschluss des Moduls sind die Studierenden in der Lage:</p> <ul style="list-style-type: none">• die Grundkonzepte von Natural Language Processing wiederzugeben, kennen verschiedene Methoden für das Text Pre-Processing und Feature Engineering sowie deren Vor- und Nachteile,• Techniken für das Pre-Processing von Text und das anschliessende Feature Engineering auf einfache Aufgabenstellungen anzuwenden,• unterschiedliche Algorithmen für die Textklassifikation und deren Funktionsweise zu nennen und einfach Programme zu erstellen, welche von diesen Gebrauch machen, <p>kennen verschiedene Methoden zur Berechnung der Text Ähnlichkeit sowie Möglichkeiten zum Clustern von Text und sind in der Lage diese miteinander</p>		

	<p>zu kombinieren und anzuwenden.</p> <ul style="list-style-type: none"> • die Funktionsweise von fortschrittlichen Feature Engineering Techniken sowie deren Vor- und Nachteile zu erläutern, • Methoden für die automatische Zusammenfassung von Texten zu nennen und diese prototypisch anzuwenden, • verschiedene semantische Analysetechniken, deren Anwendungen und Funktionsweise zu nennen, <p>Techniken für die Sentiment Analyse zu beschreiben und diese auch praktisch anzuwenden.</p>
Leistungsnachweis	<ul style="list-style-type: none"> • Mündliche Stundenwiederholungen • Mitarbeit und Hausübungen • Schriftliche Prüfungen
Nachprüfung	Nein

Unterrichtssprache	Deutsch
Eingangskompetenzen	Allgemeine Studierfähigkeit
Inhalte	<p>1. Grundlagen Text Analytics (Auswahl von):</p> <ul style="list-style-type: none"> • Einführung: Sprache (Syntax, Struktur und Semantik), Text-Corpora und Pre-Processing • Embeddings <p>2. Zentrale Methoden, z.B.:</p> <ul style="list-style-type: none"> • Part-of-Speech Tagging und Dependency/Constituent Parsing • Feature Engineering (z. B. TF-IDF, Word Embeddings) • Text-Klassifikation (z. B. Naive Bayes, Logistic Regression) • Text-Ähnlichkeit und Clustering <p>3. Erweiterte Ansätze</p> <ul style="list-style-type: none"> • Large Language Models
Lehr- und Lernmethoden	<ul style="list-style-type: none"> • Offener, fragend-entwickelnder Unterricht • Gemeinsames Erarbeiten von Beispielaufgaben und deren Lösung
Struktur	<p>56h Präsenzunterricht</p> <p>64h Selbststudium</p>
Literatur	Dipanjan Sarkar (2019): Text Analytics with Python: A Practitioner's Guide to Natural Language Processing, Apress, 2nd ed. Edition