
Spring School Digital Humanities 2022

in Regensburg

KURSBESCHREIBUNG

TITEL	To watch or not to watch? (Transformerbasierte) Sprachklassifikation für die Filmwahl
BESCHREIBUNG	<p>Textdaten aus Filmdatenbanken wie IMDb können uns viele nützliche Informationen über Filme und deren Qualität liefern, die sich mithilfe von Machine Learning-Methoden systematisch untersuchen lassen. Dieser Kurs bietet eine Einführung in die Methoden der Sprachverarbeitung und Künstlichen Intelligenz. Zudem werden State-of-the-Art End-to-End Deep Learning-Ansätze aus dem Bereich NLP besprochen, die wir nutzen, um zu untersuchen, ob Reviewer Filme sehenswert finden oder nicht.</p> <p>Der Kurs beinhaltet Praxisteile, in denen gelernte Inhalte geübt werden. Die erworbenen Kompetenzen erlauben es den Teilnehmer*innen am Ende des Kurses einen eigenen Klassifikator für die eingeführten Probleme zu erstellen.</p>
UMFANG	Der Workshop umfasst ca. 12 Stunden, verteilt über 2 Tage.
EINHEITEN	<p>Tag 1:</p> <ul style="list-style-type: none">• 120 Min.: Grundlagen in Natural Language Processing (Feature Engineering, Text Pre-processing)• 120 Min.: Einführung Klassifikation (SVM, Random Forest) und Word Vectors• 120 Min.: Praktischer Teil 1: Gruppenarbeit <p>Tag 2:</p> <ul style="list-style-type: none">• 120 Min.: Grundlagen Deep Learning• 120 Min.: Architekturen im NLP-Bereich (CNN + RNN) und insbesondere Transformers• 120 Min.: Praktischer Teil 2: Gruppenarbeit
VORAUSSETZUNGEN	<ul style="list-style-type: none">• Internetfähiges Notebook (Gruppenbildung möglich)• Kenntnisse in Python• Python- und PyCharm-Installationen• Erfahrung mit Maschinellem Lernen hilfreich

TEILNEHMER-ZAHL	Maximal 15 Teilnehmer
DOZENTEN	Prof. Dr. Bernd Ludwig, Selina Meyer, Gregor Donabauer
KOMMENTAR	Für Studierende der Informationswissenschaft an der Universität Regensburg besteht die Möglichkeit den Spring School-Kurs im Rahmen eines fortführenden Projekts in das Studium einzubringen. Bei Interesse und für weitere Informationen nehmen Sie bitte Kontakt mit den leitenden Dozenten auf.