Sprache: European Credit Points (ECTS):

Deutsch + Englisch 5

Kursziele:

Fach- und Methodenkompetenz:

* Kenntnisse und Fähigkeiten grundlegender Vorgehensweise bei der Entwicklung von Anwendungen für Machine Learning/KI und Natural Language Processing (NLP)
* Programmiersprache Python

Handlungskompetenz:

* Die Studierenden erlernen die Entwicklung und Nutzung von Machine Learning/Künstlicher Intelligenz für Textbasierte Anwendungen, u.a.
  + Wortvorkommen + Entity-Erkennung: Analyse der häufigsten Worte + Begriffe/Namen
  + Text-Klassifikation: Zuordnung zu Themenfeldern
  + Text-Zusammenfassung (automatisiert)
  + Text-Übersetzung (automatisiert)
  + Sentimentanalyse: Gefühle im Text identifizieren
  + ChatBot erstellen

Sozialkompetenz:

* Informations- und Kommunikationsverhalten quantitativ und qualitativ optimieren (z.B. Nutzung von ChatBots wie ChatGPT <https://de.wikipedia.org/wiki/ChatGPT>)

Kursinhalte:

* Grundlagen der strukturierten Programmierung mit Python.
* Algorithmen für Machine Learning/KI und Natural Language Processing (NLP) u.a. für Text Cleaning, Tokenization, Stemming, Named Entity Recognition, Text Classification, Text Summarization und Text Translation
* Anwendung von Machine Learning Algorithmen mit Hilfe der Pakete Tensorflow, NLTK und spaCy

Literatur:

Pflichtliteratur:

* Lee, R. S. T. (2024). Natural language processing: A textbook with Python implementation. Singapore: Springer.

Hinweis: Das Buch von Lee ist über die Bibliothek ausleihbar, aber auch eine gute Anschaffung

Optionale Literatur:

* Caelen. O. & Blete, M-A. (2024). Anwendungen mit GPT-4 und ChatGPT entwickeln: Intelligente Chatbots, Content-Generatoren und mehr erstellen. O’Reilly.
* Geron, A. (2022). Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow. O'Reilly.
* Goswami & Sinha (2023). Statistical Modeling in Machine Learning: Concepts and Applications. Academic Press.
* IRMA (2020). Natural language processing : concepts, methodologies, tools, and applications. IGI. Volltext: <https://www.igi-global.com/gateway/book/230042>
* Kulkarni, A. & Shivananda, A. (2021). NLP recipes: Unlocking text data with machine learning and deep learning using Python. APress.
* Kwartler, T. (2017). Text Mining in Practice with R. Hoboken, New Jersey: Wiley.
* Malik, U. M. (2020). Natural Language Processing Crash Course for Beginners: Theory and Application of NLP using TensorFlow and Keras. AI Publishing.
* Ozdemir, S. (2024). Large Language Models: Strategien und Best Practices für den Einsatz von ChatGPT und anderen LLMs. O’Reilly.
* Patel, A. A. & Arasanipalai, A. U. (2021). Applied natural language processing in the enterprise teaching machines to read, write and understand. O’Reilly.
* Tunstall, L., von Werra, L. & Wolf, T. (2023). NLP mit Transformern: Sprachanwendungen mit Hugging Face erstellen. O’Reilly.
* Vajjala, S., Majumder, B., Gupta, A. & Surana, H. (2020). Practical NLP: A comprehensive guide to building real-world NLP systems. O’Reilly.

Leistungsnachweise:

|  |  |  |  |
| --- | --- | --- | --- |
| Nr. | Aufgabe | Fälligkeit | Anteil an Note |
| 1 | Microsoft Powerpoint Präsentation und Live-Demonstration von Anwendung mit Python für NLP (60 Minuten je Thema) | 9. Nov. | 100% |
|  | **Total** |  | **100%** |

Hinweis: Jeder benotete Leistungsnachweis muss mit der Note 4 oder besser abgeschlossen werden, um den Kurs zu bestehen.

Details zum Leistungsnachweis 1 (Präsentation + Demo):

Erstellen sie eine Powerpoint Präsentation, in der sie einen Ausschnitt der Literatur des Kurses und dessen Anwendung (Demo) anhand der Beispieldaten und des Codings vom Buch vorstellen (Das Buch ist bereits in Workshops mit Codebeispielen strukturiert und der Code kann von github heruntergeladen werden). Laden sie die Präsentation (.PPTX), die Datasets und das Coding (.PY) bis zur Deadline in moodle hoch.

Da die Kapitel aufeinander aufbauen, können sich die Studierenden eines der folgenden Themen (ein Thema kann mehrere Kapitel umfassen) für ihre Präsentation in moodle auswählen. Erst wenn alle Themen vergeben sind, können sich Studierende einem bereits vergebenen Thema als zusätzliches Gruppenmitglied anschließen. Die Sprache für Präsentation und Demo ist Englisch (Dies erleichtert die Präsentation, da auch die Literatur Englisch ist).

Folgende Themen stehen zur Auswahl:

1. Workshop #1: Basics of NLP-Toolkit (Lee, Kapitel 10)
2. Workshop #2: N-grams in NLTK and tokenization in SpaCy (Lee, Kapitel 11)
3. Workshop #3: POS Tagging using NLTK (Lee, Kapitel 12)
4. Workshop #4: Semantic analysis and word vectors (Lee, Kapitel 13)
5. Workshop #5: Sentiment analysis and text classification (Lee, Kapitel 14)
6. Workshop #6: Transformers with spaCy and TensorFlow (Lee, Kapitel 15)
7. Workshop #7: Building Chatbot with TensorFlow and Transformer Technology (Lee, Kapitel 16)