

Künstliche Intelligenz in der (Bio-)Chemie (KIChem)					Stand: 01.07.2024	
Studiengang: B. Sc. Chemie					Modus: Qualifizierung	
ECTS-Punkte	Arbeitsaufwand [h]	Dauer	Turnus	Studiensemester		
8	240	1 Semester	SoSe	6.		
Lehrveranstaltungen		Typ	Umfang [SWS]	Kontaktzeit [h]	Selbststudium[h]	Gruppengröße
KIChem-Vorlesung		V	2	30	60	100
KIChem-Übungen		Üb	2	30	75	20
KIChem-Seminar		Sem	1	15	30	20
Modulverantwortliche:r		Prof. Dr. Birgit Strodel				
Beteiligte Dozierende		Dozenten des Arbeitskreises "Computergestützte Biochemie"				
Sprache		deutsch, englisch auf Wunsch				
Weitere Verwendbarkeit des Moduls		Studiengang			Modus	
		B. Sc. Biochemie			Qualifizierungsmodul	
		B. Sc. Interdisziplinäre Naturwissenschaften			Qualifizierungsmodul	
		M. Sc. Chemie			Wahlpflichtmodul	
		M. Sc. Wirtschaftschemie			Wahlpflichtmodul	
		M. Sc. Biochemie			Wahlpflichtmodul	
Lernergebnisse und Kompetenzen						
Studierende können nach erfolgreichem Abschluss des Moduls						
<ul style="list-style-type: none"> • Algorithmen und Techniken des maschinellen Lernens (ML) verstehen, • Python-Module für ML benutzen und grundlegende Dinge selbst programmieren, • chemische Daten in einem maschinenlesbaren Format erstellen, • ML-Modelle für (bio-)chemische Daten generieren, • beurteilen / bewerten, ob maschinelle Lernmodelle für eine bestimmte (bio-)chemische Fragestellung geeignet sind, • Herausforderungen und Grenzen des ML im Bereich der (Bio-)Chemie erkennen, • gezielt nach existierender ML-/KI-Software im Internet für ein (bio-)chemisches Problem suchen und diese anwenden, • wissenschaftliche Publikationen zur Anwendung von ML in der (Bio-)Chemie analysieren und in einem Vortrag vorstellen. 						
Inhalte						
Vorlesung:						
Die Vorlesung besteht aus Theorie und Beispielen (in Python) zu folgenden Themen:						
<u>1. Python in der Chemie:</u>						
<ul style="list-style-type: none"> • Einführung in die häufig verwendeten Python-Module zur Modellentwicklung für (bio-)chemischen Daten. 						
<u>2. Einführung in das maschinelle Lernen:</u>						
<ul style="list-style-type: none"> • Überblick über die Grundlagen und Konzepte des maschinellen Lernens (ML) und wie die verschiedenen Arten von Problemen mit maschinellem Lernen angegangen. • Einführung der am häufigsten verwendeten ML-Methoden (überwachtes Lernen, unüberwachtes Lernen, teilüberwachtes Lernen und bestärkendes Lernen) und ML-Algorithmen (lineare und logistische Regression, Entscheidungsbäume, Naive Bayes, <i>Random Forest</i>, k-nächste-Nachbarn, Support-Vektor-Maschine, Gradientenverstärkung) und neuronale Netze (<i>deep learning</i>). • Implementierung dieser Algorithmen in Python zur Analyse und Vorhersagen für chemische Daten. 						

3. Evaluierung und Interpretation von ML-Modellen:

- Schlüsselkonzepte und -techniken für die Auswertung und Interpretation der Ergebnisse von ML-Modellen, wie z. B. Genauigkeit, Präzision, Wiedererkennung und Konfusionsmatrizen.
- Evaluation der Zuverlässigkeit und Verallgemeinerbarkeit der ML-Modelle.
- Visualisierung der ML-Modell-Ergebnisse .

4. Maschinelles Lernen in der (Bio-)Chemie:

- Typische Anwendung von ML im Bereich der (Bio-)Chemie zur Analyse großer Datensätze sowie Identifizierung von Mustern und Trends in chemischen Daten.
- Vorstellung von typischen ML-/KI-Anwendungen in der Chemie und Biochemie: *und Biochemie*, z.B. *Chemprop* zur Vorhersage chemischer Eigenschaften, *Reaxys Predictive Retrosynthesis* für Retrosynthese, *IR Spectra Interpreter* zur Analyse von IR-Spektren, *AlphaFold* für Proteinstrukturvorhersage, *TransMEP* für Proteinmutationsvorhersage.

Übungen:

Die Übungen werden wöchentlich ausgegeben. Die Lösungen (Python-Programme oder Protokolle) werden kontrolliert und anschließend gemeinsam in der Übungsstunde besprochen.

- Begleitende Programmierübungen (Python) zu den Themen 1. bis 3. der Vorlesung-
- Übungen zum Erlernen des Umgangs der unter Vorlesungsthema 4. genannten Programme; die Studierenden können die Programme nach ihren Interessen auswählen; die Bearbeitung dieser Übungen wird durch Protokolle belegt.

Seminar:

Wöchentliches Seminar, in dem die Studierenden wissenschaftliche Publikationen, die zum jeweiligen Stoff der Vorlesung ausgegeben werden, vorstellen.

Teilnahmevoraussetzungen	Nachweis von Mathematikkenntnissen: MMC1 + MMC2 (Chemie & Biochemie) bzw. MMWC (Wirtschaftschemie) bzw. Lineare Algebra 1 + Analysis 1 (Naturwissenschaften) bzw. vergleichbare Leistungen		
Studienleistungen	Bearbeitung von Übungen inklusive Protokolle, Seminarvortrag.		
Zulassungsvoraussetzung zur Modulprüfung	Erfolgreicher Abschluss von min. 75% der Übungen.		
Prüfungen	Prüfungsform	Dauer [min]	benotet/unbenotet
	Klausur	60	benotet
Stellenwert der Note für die Gesamtnote			8/180
Sonstige Informationen			
Aktuelle Informationen finden Sie auf ILIAS und im HIS-LSF und auf der Webseite des Instituts. Wenn das Modul bereits als Qualifikationsmodul im Bachelorstudiengang gewählt wurde, ist eine Belegung als Wahlpflichtmodul im Masterstudiengang ausgeschlossen.			
Literatur			
<ul style="list-style-type: none">• Skript zur Vorlesung• Fachbücher:<ol style="list-style-type: none">1. F. A. Dalwigk, <i>Python für Einsteiger: Programmieren lernen mit dem großen Python Buch - Schritt für Schritt zum Python Profi – auch ohne Vorkenntnisse!</i>, Eulogia Verlag, 2022.2. J. Frochte, <i>Maschinelles Lernen: Grundlagen und Algorithmen in Python</i>, Carl Hanser Verlag GmbH & Co. KG, 2020.3. Ch. Hill, <i>Python for Chemists</i>, Cambridge University Press, 2023.4. S. Raschka, Y. Liu, V. Mirjalili, <i>Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python</i>, Packt Publishing, 2022.• Spezialliteratur zu Thema 4 der Vorlesung und Seminarthemen wird ausgegeben.			