

# Antrag zur Einführung/Änderung eines Moduls am Institut für Informatik

Akademischer Grad	Modulnummer	Modulform
(z.B. Bachelor of Science, Master of Science...)	10-...	(z.B. Pflicht, Wahlpflicht, Wahl)
<b>Modultitel</b>	<b>Einführung in die Informatik für Data Science</b>	
<b>Modultitel (englisch)</b>	<b>Introductory Computing for Data Science</b>	
<b>Empfohlen für:</b>	1.-3. Semester	
<b>Verantwortlich</b>	Humboldt Professur für Künstliche Intelligenz	
<b>Dauer</b>	1 Semester	
<b>Modulturnus</b>	Jedes Sommersemester	
<b>Lehrformen</b> <i>Bitte Angabe zum ZKK Kriterium 52 beachten*1</i>	<ul style="list-style-type: none"> <li>• Vorlesung „Introductory computing for Data Science“ (3 SWS) = 45 h Präsenzzeit und 100 h Selbststudium = 145 h</li> <li>• Praktikum „Introductory computing for Data Science“ (2 SWS) = 30 h Präsenzzeit und 125 h Selbststudium = 155 h</li> </ul>	
<b>Arbeitsaufwand</b>	10 LP = 300 Arbeitsstunden (Workload)	
<b>Verwendbarkeit</b>	Module im M.Sc. Data Science	
<b>Qualifikationsziele</b>	<p>At the end of this course, students should be able to use Git and write efficient Python code to analyze and visualize data using recent ML techniques such as Deep Neural Networks, PCA, k-means, Reinforcement learning, Manifold learning, etc... This course aims to provide the students with the algorithmic concepts, the approaches to translating their algorithms into Python code and finally, how to optimize their code or find the appropriate Python library or module adapted to their needs. Several Python modules will be introduced in this course, such as NumPy, Scipy, Pandas, Matplotlib, Scikitlearn, Tensorflow and more. And to allow students to practice during the course constantly, individual homework will be given, and group projects along this course. The sum of students' effort in contributing to the lecture, group projects and homework will yield their overall marks.</p>	
<b>Inhalt</b>	<ul style="list-style-type: none"> <li>● Introduction to Git</li> <li>● Introduction to Python <ul style="list-style-type: none"> <li>○ Control flow, list comprehension, and functions</li> <li>○ Data structures: list, array, set and dict</li> <li>○ Numpy basics</li> <li>○ Advanced indexing and broadcasting</li> <li>○ Scipy</li> <li>○ Pandas: basics and advanced usages</li> </ul> </li> <li>● Structuring an ML project <ul style="list-style-type: none"> <li>○ Introduction to ML Strategy</li> <li>○ Setting up the project goal</li> <li>○ Comparing the model's performance to human-level</li> <li>○ Carrying out the error analysis</li> <li>○ Mismatched training and dev/test sets</li> </ul> </li> <li>● Introduction to visualization tools <ul style="list-style-type: none"> <li>○ Pyplot</li> <li>○ Seaborn</li> </ul> </li> </ul>	

# Antrag zur Einführung/Änderung eines Moduls am Institut für Informatik

	<ul style="list-style-type: none"> <li>● Numerical optimization methods             <ul style="list-style-type: none"> <li>○ Line search algorithm</li> <li>○ Naive Gradient Descent</li> <li>○ Improving the Gradient Descent algorithm</li> <li>○ Taylor expansion and Newton's Method from 1d to nd</li> <li>○ Scipy modules for numerical optimization</li> </ul> </li> <li>● OOP (Object Oriented programming)             <ul style="list-style-type: none"> <li>○ Notions of classes and objects in Python</li> <li>○ Polymorphism and Inheritance</li> </ul> </li> <li>● Machine learning tools             <ul style="list-style-type: none"> <li>○ Scikit learn</li> <li>○ PyTorch and TensorFlow</li> </ul> </li> <li>● Extra resources:             <ul style="list-style-type: none"> <li>○ Statmodels and patsy</li> <li>○ PyMc3 and Arviz</li> <li>○ Apach arrow</li> <li>○ SQL and NoSQL databases</li> <li>○ Linux Admin System</li> </ul> </li> </ul> <p>Das Modul wird in englischer Sprache gehalten. Lehr- und Prüfungssprache ist englisch.</p>	
<p><b>Teilnahmevoraussetzungen</b></p>	<p>Kenntnisse über Grundrechenarten, Statistik und Algorithmus</p> <p>Basic knowledge of arithmetic, statistics, calculus, linear algebra and algorithms is required to attend this course. In case the student does not have any of the above knowledge, I will recommend the following resources:</p> <p>1. <i>Stroud and Booth, (2020) "Advanced Engineering Mathematics", Bloomsbury</i></p> <p>2. <i>Silberschatz, A. et al. (2012), Operating System Concepts, Wiley.</i></p>	
<p><b>Literaturangabe</b></p>	<p>Homepage des MSc Data Science, Vorlesungsskript</p>	
<p><b>Vergabe von Leistungspunkten</b></p>	<p>Die Vergabe von Leistungspunkten richtet sich nach den in der Prüfungsordnung zum Studiengang getroffenen Regelungen.</p>	
<p><b>Prüfungsleistungen und Prüfungsvorleistungen*2</b></p>	<p>Lösen von Aufgaben mit Wichtung 3 (6 Programmieraufgaben mit je 1 Woche Bearbeitungszeit)</p>	<p>Vorlesung (3 SWS)</p>
	<p>Gruppen-Projektarbeit mit Wichtung 2 (Projektplanungspräsentation (20 Min.), Projektarbeit (Programmieraufgabe), Abschlusspräsentation mit Diskussion (20 Min.) und schriftlicher Ausarbeitung (Foliensatz), Bearbeitungszeit 6 Wochen)</p>	<p>Praktikum (2 SWS)</p>