

Course Manual DDML

Data Mining

Version: 1 | Last Change: 12.02.2021 13:46 | Draft: 0 | Status: vom verantwortlichen Dozent freigegeben

– General information

Long name	Data Mining
Approving CModule	<u>DML BaET</u>
Responsible	Prof. Dr. Beate Rhein Professor Fakultät IME
Valid from	summer semester 2023
Level	Bachelor
Semester in the year	winter semester
Duration	Semester
Hours in self-study	78
ECTS	5
Professors	Prof. Dr. Beate Rhein Professor Fakultät IME
Requirements	From Mathematics 1 and 2 the ability to construct mathematical models as well as knowledge of differential calculus and linear algebra is required.
Language	German
Separate final exam	Yes

Literature

Final exam

Details

Depending on the number of participants:
For a small number of participants:
combination of exam or oral examination and evaluation of the mini-project.
For many participants, examination by written examination; mini-project as prerequisite for participation in the examination.

In the written or oral examination, the methods, procedures, pitfalls and legal foundations of data mining are examined.

In the mini-project the ability to act independently and on one's own responsibility and the use of suitable software will be tested.

Minimum standard

Basic knowledge of the general approach to data mining, the procedures covered and their limitations.

Exam Type

EN andere summarische Prüfungsform

– Lecture / Exercises

Learning goals

Goal type	Description
Knowledge	<p>Introduction to a suitable software, e.g. Python</p> <p>Introduction to descriptive statistics and possibly also probability calculation</p> <p>Supervised learning:</p> <ul style="list-style-type: none"> - Classification procedure: Procedure, performance measures, application of a method of instance-based learning, e.g. k-nearest-neighbor and a method of model-based learning, e.g. decision trees - Possibly regression analysis: about machine learning and classical <p>Unsupervised learning:</p> <ul style="list-style-type: none"> - Cluster analysis: k-means, possibly also DBSCAN <p>Preprocessing of the data:</p> <ul style="list-style-type: none"> - Handling Damaged / Missing Data - Runaway or noise - problems - Scaling - Visualization of data <p>- Possible dimension reduction</p> <ul style="list-style-type: none"> - Assessment of data quality - possibly look at different types of data records, make reference to NoSql databases, Outlook on current research, e.g. image recognition, Natural Language Processing, Reinforcement Learning
Skills	<p>Be able to name and apply a suitable method and overall approach to tasks</p> <p>Select and evaluate a suitable performance measure</p> <p>Apply Privacy Policy</p>

Expenditure classroom teaching

Type	Attendance (h/Wk.)
Lecture	2
Exercises (whole course)	0

Special literature

A. Geron: Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme, Heidelberg, o'Reilly Verlag 2017, 978-3960090618, S. Raschka, V. Mirjalili: Machine Learning mit Python und Scikit-Learn und TensorFlow: Das umfassende Praxis-Handbuch für Data Science, Predictive Analytics und Deep Learning, mitp Verlag, 2018, 978-3958457331, J. Frochte, Jörg: Maschinelles Lernen, München, Carl Hanser Verlag GmbH & Co. KG, 2018, eBook ISBN: 978-3-446-45705-8, Print ISBN: 978-3-446-45291-6, A. Müller: Einführung in Machine Learning mit Python: Praxiswissen Data Science, Heidelberg, o'Reilly Verlag 2017, eBook: 978-3-96010-111-6

Special requirements

none

Accompanying material	<p>Script or set of slides</p> <p>Tasks (expected to be integrated into the script)</p> <p>Mini project task with data set</p>
------------------------------	--

Separate exam	No
----------------------	----

Exercises (shared course)	2
---------------------------	---

Tutorial (voluntary)	0
----------------------	---