

# Module Description

## 39-M-Inf-ABDA Advanced Big Data Analytics / Big Data Machine Learning

Faculty of Technology

*Version dated Feb 9, 2026*

This module guide reflects the current state and is subject to change. Up-to-date information and the latest version of this document can be found online via the page

<https://ekvv.uni-bielefeld.de/sinfo/publ/modul/308598306>

The current and valid provisions in the module guide are binding and further specify the subject-related regulations (German "FsB") published in the Official Announcements of Bielefeld University.

## 39-M-Inf-ABDA Advanced Big Data Analytics / Big Data Machine Learning

---

### Faculty

---

Faculty of Technology

### Person responsible for module

---

Prof. Dr. Alexander Schönhuth

### Regular cycle (beginning)

---

To be discontinued

### Credit points

---

5 Credit points

### Competencies

---

Das Modul legt den Schwerpunkt auf den Einsatz von Techniken des Maschinellen Lernens und der Künstlichen Intelligenz zur Analyse grosser Datenmengen. Beispiele sind oft mit populären Anwendungen im Internet verbunden, wie zum Beispiel Recommender Systems oder Online-Werbung, die unser tägliches Leben in vielerlei Bereichen mitbestimmen.

Die Idee ist, ein tieferes algorithmischen und statistisch-mathematisches Verständnis für die zugrundeliegenden Vorgänge zu entwickeln. Das Ziel ist, dass sich Teilnehmer nach der Teilnahme an der Veranstaltung in der Lage fühlen, selber an den bestehenden Prozessen mitwirken zu können, sowohl in Theorie als auch Praxis.

Je nach konkreter Themenwahl kann das Seminar/Projekt mit vom Lehrenden gehaltenen Tutorien starten, die an die entsprechenden Themen hinführen. Im Rahmen der Vorträge/Projekte und Ausarbeitungen soll dann auch generell, je nach Aufgabe, das Verstehen, Interpretieren und Bewerten von Originalliteratur bzw. die konkrete Umsetzung von Theorie in Praxis geübt werden.

Abhängig von der jeweiligen Variante sollen entweder Präsentationstechniken oder Programmierstechniken erarbeitet werden. Schliesslich soll das Anfertigen einer schriftlichen Ausarbeitung, die wissenschaftlichen und entsprechenden ethischen Standards (zum Beispiel die Reproduzierbarkeit der Implementierungen gewährleisten) genügen, erlernt werden.

The emphasis of this module is on understanding and application of machine learning and artificial intelligence techniques for the analysis of large datasets. Examples of such techniques often deal with popular internet based applications, such as recommender systems or web advertisements, in any case techniques that are influential for modern daily life.

The basic motivation is to develop a deeper understanding for the algorithms, the statistics that underlie such techniques and the processes they drive. The goal is that seminar participants feel well equipped, and poised to actively engage in related processes, both in theory and in professional practice.

Depending on concrete choice of topics, the seminar/project can start with tutorials run by the teacher, to introduce seminar participants to the topics of choice. Presentations or projects then are supposed to practice how to understand, how to interpret and evaluate original literature, or to implement relevant theory in practical applications. While in case of a presentation, the focus is on presentation techniques, programming techniques are stressed in case of programming projects. Drafting a scientific report is supposed to generate text that complies with actual scientific and ethical standards; an important aspect is the exact reproducibility of the facts and processes discussed.

## Content of teaching

Inhaltliche Elemente der Tutorien und der im Seminar behandelten Literatur beziehen sich auf Themen, bei denen der Einsatz künstlicher Intelligenz und des maschinellen Lernens zur Analyse großer Datenmengen eine prominente Rolle spielt. Populäre und allgegenwärtige Beispiele sind zum Beispiel die Bestimmung von 'Communities', oder das Vorhersagen von (noch nicht erkenntlichen) Links in sozialen Netzwerken, der Einsatz von deep neural networks in der Vorhersage einer Passung von Werbung zu Nutzer, oder in der Empfehlung von Produkten in Online-Stores, die einen maximalen Umsatz versprechen. Klassifizierungssysteme, die (sehr) große Datenströme (data streams) auswerten, und aufgrund der fortlaufend eintreffenden Daten regelmäßig aktualisiert werden müssen, sind ebenso von Interesse.

Contents of tutorials and literature relevant for the seminar cover topics where the employment of artificial intelligence and machine learning for the analysis of big datasets plays a prevalent/pivotal role. Popular, ubiquitous examples are the identification of communities, or the prediction of (not yet recognizable, because hidden) links in social networks, the employment of deep neural networks in predicting the fit of advertisements with (e.g. search engine) users, or the recommendation of products in online stores that promise maximal profits. Classification systems that evaluate big data streams, which need to be regularly updated, are equally relevant.

## Recommended previous knowledge

Eine Teilnahme an der Vorlesung Big Data Analytics ist hilfreich. Desweiteren sind grundlegende Kenntnisse in Algorithmen und Datenstrukturen und der Künstlichen Intelligenz für das Verständnis der besprochenen Techniken nötig.

Participation in the lecture 'Big Data Analytics' is helpful. Basic knowledge of algorithms, data structures and artificial intelligence is required.

## Necessary requirements

—

## Explanation regarding the elements of the module

Module structure: 1 uPr <sup>1</sup>

## Courses

Title	Type	Regular cycle	Workload <sup>5</sup>	LP <sup>2</sup>
Machine Learning and AI in Advanced Big Data Analytics	project o. seminar	Wird unregelmäßig angeboten, vorzugsweise im Wintersemester	60 h (30 + 30)	2 [Pr]

## Examinations

Allocated examiner	Type	Weighting	Workload	LP <sup>2</sup>
--------------------	------	-----------	----------	-----------------

<p>Teaching staff of the course <b>Machine Learning and AI in Advanced Big Data Analytics (project o. seminar)</b></p> <p><i>Referat im Umfang von 20 bis 30 Minuten mit schriftlicher Ausarbeitung im Umfang von 8 bis 10 Seiten</i></p>	Referat mit Ausarbeitung	without grades	90h	3
---	--------------------------	----------------	-----	---

## Further notices

---

Das Modul kann in folgenden Wahlpflichtbereichen (WP) anerkannt werden:

- WP im Master Intelligente Systeme
- WP im Master Bioinformatik und Genomforschung
- WP im Master Naturwissenschaftliche Informatik

Bisheriger Angebotsturnus war unregelmäßig.

## Legend

---

- 1 The module structure displays the required number of study requirements and examinations.
  - 2 LP is the short form for credit points.
  - 3 The figures in this column are the specialist semesters in which it is recommended to start the module. Depending on the individual study schedule, entirely different courses of study are possible and advisable.
  - 4 Explanations on mandatory option: "Obligation" means: This module is mandatory for the course of the studies; "Optional obligation" means: This module belongs to a number of modules available for selection under certain circumstances. This is more precisely regulated by the "Subject-related regulations" (see navigation).
  - 5 Workload (contact time + self-study)
- SoSe** Summer semester
- WiSe** Winter semester
- SL** study requirement
- Pr** Examination
- bPr** Number of examinations with grades
- uPr** Number of examinations without grades