

1.40 Machinelles Lernen und Data Mining (INF-MLDM)

Zugeordnete Modulteile:

Lehrform	Veranstaltung	Tu- rus	Präsenzzeit	Selbst- studium	ECTS
Vorlesung	Vorlesung zu Machinellem Lernen und Data Mining	SoSe	45 h (3 SWS)	45 h	3 CP
Übung	Übungen zu Machinellem Lernen und Data Mining	SoSe	30 h (2 SWS)	60 h	3 CP

Im Modul müssen insgesamt 3 ECTS-Punkte erworben werden. Die Präsenzzeit beträgt 2 Wochenstunden. Inklusive Selbststudium sind etwa 90 Stunden aufzuwenden.

**Art des Mo-
duls**

Dauer Das Modul erstreckt sich über 1 Semester.

**Art der Be-
wertung** benotet ()

**Form der
Modulprüfung** Klausur (90-180 Minute) oder mündlich (15-30 Minute)
Wiederholbarkeit: beliebig, Zulassungsvoraussetzung: keine

**Modul-
verantwort-
licher** Prof. Dr. Volker Tresp

Anbieter Ludwig-Maximilians-Universität München
Fakultät für Mathematik, Informatik und Statistik
Institut für Informatik
Kerninformatik
LFE Mobile und Verteilte Systeme

**Unterrichts-
sprache(n)** Englisch

Inhalte

Ziel des Maschinellen Lernens ist die Entwicklung von Methoden zur Realisierung lernfähiger technischer Systeme. In der Vorlesung werden die wichtigsten Zweige des Maschinellen Lernens vorgestellt:

- Beim überwachten Lernen werden basierend auf vorhandenen Beispieldaten Systeme zur Klassifikation und zur Modellierung funktioneller Abhängigkeiten trainiert

(typische Vertreter: Neuronales Netz, Support Vector Machine).

- Das Ziel des unüberwachten Lernens ist es, autonom relevante Strukturen in Daten zu finden (typische Vertreter: Cluster Analyse, Independent Component Analyse).
- Kausal Probabilistische Netze (Bayesian Networks) stellen einen Rahmen zur Beschreibung und zum Erlernen von komplexen probabilistischen Abhängigkeiten dar, bis hin zur Analyse von Kausalität.
- Reinforcement Learning ist das wichtigste Lernverfahren zur Optimierung des Verhaltens von Agenten.

Datamining bezeichnet die Analyse von Datenbanken mit Methoden des Maschinellen Lernens und motiviert Forschungsaktivitäten mit neuen Anforderungen an Skalierbarkeit und Bedienbarkeit der Lösungen. Ein typisches Ziel ist die Nutzbarmachung von Kundendaten (Customer Relationship Management). In der Vorlesung werden eine Reihe von industriellen Anwendungen des Maschinellen Lernens, mit Schwerpunkt auf Anwendungen im Data Mining, vorgestellt.

Literaturhinweise

Die Studierenden arbeiten weitgehend selbständig und haben die Möglichkeit sich zu kleinen Teams zusammenschließen. Nach einer Einführung in das jeweilige Themengebiet werden praktische Aufgaben eigenständig gelöst.

Vorkenntnisse

Die Beherrschung mindestens einer Programmiersprache.

Qualifikationsziele

1.40 Machine Learning and Data Mining (INF-MLDM)

Associated Module Components:

Teaching Component	Rota	Attendance	Selfstudy	ECTS	
lecture	Lecture: Machine Learning and Data Mining	SoSe	45 h (3 SWS)	45 h	3 CP
exercise	Exercises: Machine Learning and Data Mining	SoSe	30 h (2 SWS)	60 h	3 CP

3 credit points are awarded for this module. The attendance time is 2 hours a week. Including self-study, there are about 90 hours to be spent.

Type

Duration The module comprises 1 semester.

Grading marked ()

Type of Examination Klausur (90-180 Minute) oder mündlich (15-30 Minute)
Repeatability: arbitrary, Admission Requirements: none

Responsible for Module Prof. Dr. Volker Tresp

Provider Ludwig-Maximilians-University Munich
Faculty for Mathematics, Computer Science and Statistics
Institute for Computer Science
Core Computer Science
Mobile and Distributed Systems Group

Teaching Lang. English

Contents

Machine Learning is a data-driven approach for the development of technical solutions. Initially motivated by the adaptive capabilities of biological systems, machine learning has increasing impact in many fields, such as vision, speech recognition, machine translation, and bioinformatics, and is a technological basis for the emerging field of Big Data.

Data mining is closely related to machine learning but has developed as a separate field, more directly motivated by the challenges in working with large databases and Web data.

The lecture will cover:

- Supervised learning: the goal here is to learn functional dependencies for classification and regression. We cover linear systems, basis function approaches, kernel approaches and neural networks
- Unsupervised Learning: the goal here is to compactly describe important structures in the data. Typical representatives are clustering and principal component analysis
- Graphical models (Bayesian networks, Markov networks), which permit a unified description of high-dimensional probabilistic dependencies
- Reinforcement Learning as the basis for the learning-based optimization of autonomous agents
- Some theoretical aspects: frequentist statistics, Bayesian statistics, statistical learning theory

The technical topics will be illustrated with a number of real-world applications.

Recommended Literature

Students work independently and have the opportunity to form small teams. After an introduction to the specific topic area practical problems are solved independently.

Prior Knowledge

At least one programming language should be mastered.

Qualifikation Aims