Tárgy neve: Applied Data Science

Tárgyfelelős neve: Angyalné Dr. Alexy Márta

Tárgyfelelős tudományos fokozata: PhD

Tárgyfelelős MAB szerinti akkreditációs státusza: AT

Az oktatás célja angolul / Aim of the subject:
The goal is to show the students how to approach a data science project, introduce the most common issues data scientists have to deal with during their work, discuss the issues of data quality and privacy, and provide an overview of open-source as well as commercial data science technologies.

Knowledge
- Know the general and specific mathematical and computational principles, facts, rules, contexts and procedures that form the basis of computer science. The relevant fields are: analysis (calculus), numerical analysis, discrete mathematics, linear algebra, operations research, probability and statistics, foundations of logic, computation theory, the design and analysis of algorithms, formal languages and automata theory, and the foundations of artificial intelligence
- Know and understand the most important general theories, contexts and facts, as well as the related concepts in the field of computer science, in particular: the methodological foundations of programming, programming languages, compilers, application development, and programming environment; computer architectures, operating systems, computer networks, distributed systems, and the theoretical foundations of databases
- Know the basic principles, methods and procedures of IT design, development, operation and management, in particular – according to their chosen specialisation – in the following areas: programming technology, database design and management, design and management of enterprise information systems, development of Internet tools and services, development of geographic information systems, design and management of distributed systems, information security, and applications of logic in IT

Abilities:
- Apply general and specific mathematical and computational principles, facts, rules and contexts in the field of information technology
- Apply formal models of computer science
- Apply computer science to design, analyse and implement algorithms, while taking into account the main programming paradigms
- Utilise computer science to interpret existing systems designs and to apply software development methodologies and technologies, with particular attention to software development, documentation, testing, code quality control and validation in project work

Attitude:

- Be committed to and authentically represent the professional principles of the IT discipline
- Be open to learning about and embracing professional and technological developments and innovations in their field of competence and specialisation
- Attach importance to the dissemination of IT achievements to both professional and non-professional audiences
- Attach importance to the promotion and implementation of environmental awareness and social responsibility
- Adopt the ethical principles and the rules of work and organisational culture of the IT profession
- Reflect on their own professional competencies and activities
- Strive for continuous professional development and general self-education
- Seek cooperation with professionals of other disciplines
- Respect the legal regulations of the IT field

Autonomy, responsibility:
- Assume responsibility for their professional activities
- Approach professional conflicts with a constructive attitude
- Strive for efficiency and high quality at work
- Take responsibility for sub-tasks in complex software development projects
- Make responsible decisions on developing their own knowledge and building their career
- Perform their work with respect for information security

Az oktatás tartalma angolul / Major topics:

The course is concerned with introducing the basic concepts in data science through real-world examples from various application areas with special focus on telecommunication, agriculture and enterprise management.

Data science techniques will be dealt with in a black-box manner, i.e. the focus is to show what is on their input and output as well as how to interpret the resulting models. During the course, students work on a semester project which quality will determine the final grade.

Modules: The main modules/topics of the course include the following:
1. Data Science methodologies (KDD, CRISP-DM, etc.)
2. Unsupervised techniques (clustering, frequent pattern mining)
3. Supervised techniques (classification, regression, recommendation)
4. Data types, data quality and data privacy
5. Data streams and sensor fusion techniques
6. Open-source technologies for data science (Python libraries)
7. Commercial technologies for data science (Azure, SPSS, etc.)
8. Introduction to various domains (telecommunication, agriculture, enterprise management, etc.) and their characteristics from the point of view of data scientists
9. Project work

A számonkérés és értékelés rendszere angolul / Requirements and evaluation:

Contineous evaluation

Irodalom / Literature:

1. Sandeep Koranne: Handbook of OpenSource Tools. ISBN-13 978-1441977182. Springer, 2011
2. John D. Kelleher, Brendan Tierney: Data Science. ISBN-13 978-0262535434, MIT Press, 2018
3. Joe Reis, Matt Housley: Fundamentals of Data Engineering: Plan and Build Robust Data Systems. ISBN-13 978-1098108304, O'Reilly, 2022
4. Albert-László Barabási and Márton Pósfai: Network Science. ISBN-13 978-1107076266, Cambridge University Press, 2015
5. Information Technology Essentials Volume 1: Introduction to Information Systems
6. Brian Williams, Stacey Sawyer: Using Information Technology. ISBN-13 978-0073516882, McGraw-Hill, 2015

Dátum: 2023. 04. 20.

tárgyfelelős oktató aláírása   tanszékvezető aláírása