COURSE DESCRIPTION CARD - SYLLABUS

Course name

Data Analysis and Semantic Networks fo the Internet of Things

Course

Field of study	Year/Semester
Computing	1/1
Area of study (specialization)	Profile of study
Internet of Things	general academic
Level of study	Course offered in
Second-cycle studies	Polish
Form of study	Requirements
full-time	compulsory

Number of hours

Lecture	Laboratory classes	Other (e.g. online)
30	30	
Tutorials	Projects/seminars	

Number of credit points

5

Lecturers

Responsible for the course/lecturer:

dr hab. inż. Jerzy Stefanowski, prof. PP

dr hab. inż. Agnieszka Ławrynowicz

dr inż. Tomasz Łukaszewski

Responsible for the course/lecturer:

Prerequisites

A student starting this course should have basic knowledge of Internet technologies (including XML), the basics of logic and databases, and Python programming. He should also have the ability to obtain information from the indicated sources and be ready to handle cooperation within the team.

Course objective

Introduction to data analysis in the field of machine learning, including classification, feature selection, cluster analysis, data streams and Big Data. Provide students with basic knowledge in the field of semantic technologies, including the concept of Web 3.0 (semantic web). Developing students' problem-solving skills in the field of using and designing systems using machine learning and semantic technologies.

Course-related learning outcomes

Knowledge

1. Has advanced detailed knowledge of classification, feature selection, cluster analysis, data streams, large data volumes, semantic technologies, Web 3.0.

2. Has knowledge of development trends and new achievements in machine learning and semantic technologies.

3. Knows advanced methods, techniques and tools used in solving complex engineering tasks in the field of computing related to machine learning and semantic technologies.

Skills

1. Can plan and carry out experiments in the area of machine learning, interpret the obtained results and draw conclusions.

2. Can evaluate the usefulness of methods and tools of semantic technologies in the Internet of Things.

Social competences

Understands that knowledge and skills become obsolete very quickly in computing.

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

The knowledge acquired during the lecture is verified by a written exam. Passing threshold: 50% of points. Final issues, on the basis of which the questions are developed, will be sent to students by e-mail using the university's e-mail system. The skills acquired during the laboratory classes are verified on the basis of the presentation resulting from the analysis of the indicated problem related to machine learning and semantic technologies.

Programme content

The data analysis lecture program includes: nearest neighbor classifier, decision trees, Bayesian classification, classifier assessment, feature selection, hierarchical and k-means grouping, classifier enssemble, data streams and large data volumes. In the field of semantic networks, the program covers the following issues: the concept of the semantic web and the World Web of Things, layered architecture of the semantic web languages, ternary data model, RDF resource description language, knowledge representation using ontology, SPARQL query language, access to data via ontology, metadata modeling and knowledge engineering, examples of current initiatives related to semantic data and knowledge integration (e.g. http://schema.org), ontologies and metadata schemas for sensor networks (W3C SSN) and the use of semantic technologies in the Internet of Things.

The laboratory program covers the issues discussed during the lectures. Data analysis using modules for Python: data visualization, nearest neighbor classifier, decision trees, naive Bayesian classifier, feature selection, hierarchical and k-means grouping. Data representation in the RDF model. Modeling ontology using the ontology editor (Protégé). Metadata modeling (schema.org, JSON-LD). Processing of semantic data (triple repositories, MongoDB). Transformation of data into knowledge graph format. Querying heterogeneous sources of knowledge with SPARQL. Access to data via ontology (R2RML). Summary of acquired knowledge in semantic technologies within the mini-project.

Teaching methods

Lecture: multimedia presentation

Laboratory exercises: practical exercises, discussion, team work

Bibliography

Basic

1. Python: uczenie maszynowe, Sebastian Raschka, Helion 2018

2. Ontologie w systemach informatycznych, Krzysztof Goczyła, EXIT 2011

3. Linked Data: Evolving the Web into a Global Data Space (1st edition). Tom Heath and Christian Bizer, Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, 2011, http://linkeddatabook.com/book

Additional

1. Naczelny Algorytm. Jak jego odkrycie zmieni nasz świat, Pedro Domingos, Helion 2016

2. Semantic Web for the Working Ontologist, Dean Allemang and Jim Hendler, Morgan Kaufmann 2008

3. Demystifying OWL for the Enterprise, Michael Uschold, Morgan & Claypool Publishers, 2018

4. An Introduction to Ontology Engineering. Keet, C.M. College Publications, volume 20, November 2018

5. Internet rzeczy. Budowa sieci z wykorzystaniem technologii webowych i Raspberry Pi, Dominique Guinard, Vlad Trifa, Helion, 2017

6. Semantic data mining. An ontology-based approach. Agnieszka Ławrynowicz. Studies on the Semantic Web, Vol. 29. IOS Pres/AKA Verlag 2017

Breakdown of average student's workload

	Hours	ECTS
Total workload	125	5,0
Classes requiring direct contact with the teacher	65	2,6
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests/exam, project preparation) [1]	60	2,4

[1] delete or add other activities as appropriate