Tárgyleírás

Tárgy neve: Natural Language Processing

Tárgyfelelős neve: Dr. habil Lőrincz András

Tárgyfelelős tudományos fokozata: CSC

Tárgyfelelős MAB szerinti akkreditációs státusza: AT

Az oktatás célja angolul / Aim of the subject:

Knowledge

- They have comprehensive and up-to-date knowledge of general mathematical and computing principles
- Posess the knowledge of specific tools and methods of Natural Language Processing

Abilities:

- They are able to apply their mathematical, computer science and informatics skills in a novel way in order to solve tasks in IT research and development.
- They are able to formalize complex IT tasks, to identify and study their theoretical and practical background and then to solve them.

Attitude:

- They follow professional and technological developments in their IT field.
- They are committed to lifelong learning and are open to acquiring new IT competencies.

Autonomy, responsibility:

- They take responsibility for their professional decisions made in their IT-related activities.
- They undertake to meet deadlines and to have deadlines met.
- They bear responsibility for their own work as well as for the work of their colleagues they work together with in a project.
- Regarding mission critical IT systems, they can be entrusted with developing and operational responsibilities that are in accordance with their professional competencies.

Az oktatás tartalma angolul / Major topics:

During the semester, student learn concepts of natural language processing and applications towards human-machine interaction. Topics in the course are as follows:

- Introduction to NLP
- Computational Linguistics: syntax, parsing, semantics, POS tagging, tokenization
- Databases: WordNet, Wikipedia, ConceptNet
- Basic methods: bag of words, word sense disambiguation, latent semantic analysis
- Overview of applications: document retrieval, knowledge extraction and summarization, sentiment analysis, chatbots
- Introduction and Word Vectors
- Word Vectors and Word Senses
- Linguistic Structure: Dependency Parsing
- N-Grams, Recurrent Neural Networks and Language Models
- Machine Translation, Seq2Seq and Attention
- Contextual Word Representations: BERT
- Grounding Language on Visual Concepts
- Grounding Language on Visual Programs
- Grounding Language through Multi-Agent Collaboration
- Conversational Agents

A számonkérés és értékelés rendszere angolul / Requirements and evaluation:

Mixed assessment, exam

Irodalom / Literature:

Recommended:

1. Jacob Eisenstein: Introduction to Natural Language Processing, MIT. 2019, ISBN 9780262042840
2. Computational linguistics, Stanford Encyclopedia, 2014 https://plato.stanford.edu/entries/computational-linguistics/
3. Chris Manning: Natural Language Processing with Deep Learning, CS224n Stanford Course, 2020. http://web.stanford.edu/class/cs224n/
4. Katerina Fragkiadaki: Language Grounding to Vision and Control, CMU course 10-808, 2017. https://katefvision.github.io/LanguageGrounding/