EE-724 **Human language technology: applications to information access**

Popescu-Belis Andrei

Cursus	Sem.	Type
Electrical Engineering		Opt.

Language of teaching	English
Credits	4
Session	
Exam	Multiple
Workload	120h
Hours	**56**
Courses	28
TP	28
Number of positions	

Frequency

Every 2 years

Remark

Next time: Spring 2022

Summary

The HLT course introduces applications of human language technology focusing on accessing text information across three types of barriers: the quantity barrier (large repositories), the cross-lingual barrier, and the subjective barrier (human interactions).

Content

The following technologies will be studied for each barrier to information access:

- The quantity barrier: information retrieval, web search, document classification, topic models, learning to rank, question answering, recommender systems.
- The crosslingual barrier: machine translation (history of the field, presentation of rule-based and of statistical systems including phrase-based and tree-based ones, domain adaptation, the use of syntax and semantics), methods for text alignment, issues and metrics for MT evaluation, cross-language information retrieval.
- The subjective barrier: sentiment analysis, subjectivity detection, analysis of human exchanges (spoken or written) for information access, search within multimedia archives.
- Conclusion on the bases of HLT research: defining a problem, building reference data, finding features for machine learning algorithms, training the algorithms, evaluating and analyzing the performance.

Keywords

Human language technology, language engineering, information retrieval, machine translation.

Learning Prerequisites

Recommended courses

At least one prior course in statistics, machine learning, computational linguistics, or artificial intelligence. Programming proficiency in a language such as Perl or Java.

Assessment methods

Project report and oral presentation.

Resources

Websites

- http://iict-space.heig-vd.ch/apu/hlt-course/