linfo2263

2022

UCLouvain

Computational Linguistics

5.00 credits	30.0 h + 15.0 h	Q1

Teacher(s)	Dupont Pierre ;Tack Anaïs (compensates Dupont Pierre) ;
Language :	English > French-friendly
Place of the course	Louvain-la-Neuve
Main themes	• Various levels of linguistic analysis • Corpus processing • Part-of-speech tagging • Probabilistic language modeling (N-grams and Hidden Markov Models) • Formal grammars and parsing algorithms • Machine translation, deep learning • Linguistics engineering applications such as automatic completion software, POS tagging, parsing or machine translation
Learning outcomes	**At the end of this learning unit, the student is able to :** 1 Given the learning outcomes of the "Master in Computer Science and Engineering" program, this course contributes to the development, acquisition and evaluation of the following learning outcomes: INFO1.1-3 INFO2.3-4 INFO5.3-5 INFO6.1, INFO6.4 Given the learning outcomes of the "Master [120] in Computer Science" program, this course contributes to the development, acquisition and evaluation of the following learning outcomes: SINF1.M4 SINF2.3-4 SINF5.3-5 SINF6.1, SINF6.4 Students completing successfully this course should be able to • describe the fundamental concepts of natural language modeling • master the methodology of using linguistic resources, in particular large scale corpora, possibly annotated or structured • apply in a relevant way statistical language modeling techniques • implement recent machine learning methods applied to language processing • develop linguistic engineering applications Students will have developed skills and operational methodology. In particular, they have developed their ability to • integrate a multidisciplinary approach between computer science and linguistics, using wisely the terminology, tools and existing methods, • manage the time available to complete projects of medium size, • manipulate and exploit large amounts of data.
Evaluation methods	The projects are worth 30 % of the final grade, 70 % for the final exam (closed-book). The projects **cannot** be implemented again in second session. The project grades are fixed at the end of the semester and included as such in the global score for the second session. The final exam is, by default, a written exam (on paper or, when appropriate, on a computer).
Teaching methods	• Lectures • Practical projects implemented in Python on the *Inginious* platform.
Content	• Various levels of linguistic analysis • (Automated) corpus processing: formating, tokenization, data tagging • Probabilistic language models: N-grams, HMMs • Part-of-Speech Tagging • (Probabilistic) Context-Free Grammars: parameter estimation and parsing algorithms • Introduction to Machine Translation

	• Introduction to Deep Learning • Typical linguistic applications such as automated completion, POS taggers, parsing or machine translation.
Inline resources	moodle.uclouvain.be/course/view.php?id=1182
Bibliography	One recommended textbook - un ouvrage conseillé : • Speech and Language Processing, D. Jurafsky and J.H. Martin, Prentice Hall.
Faculty or entity in charge	INFO

Programmes containing this learning unit (UE)				
Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	5		🔍
Master [120] in Linguistics	LING2M	5		🔍
Master [120] in Computer Science and Engineering	INFO2M	5		🔍
Master [120] in Computer Science	SINF2M	5		🔍
Master [120] in Data Science Engineering	DATE2M	5		🔍
Master [120] in Data Science: Information Technology	DATI2M	5		🔍