

- Faculté des sciences
- www.unine.ch/sciences

Natural language processing (3IN2010)

| Filières concernées | Nombre d'heures | Validation | Crédits ECTS |
|-------------------------------|---------------------------------------|-------------------|--------------|
| Master en informatique | Cours: 2 ph Exercice: 2 ph | écrit: 2 h | 5 |

ph=période hebdomadaire, pg=période globale, j=jour, dj=demi-jour, h=heure, min=minute

Période d'enseignement:

- Semestre Printemps

Equipe enseignante:

prof. Jacques Savoy

Objectifs:

The main objective of this course is to introduce the students to the underlying problems when facing with natural languages data.

- Representation and standards;
- Statistical methods for natural language processing;
- Prolog and parsing;
- Markov Models and Hidden Markov Models

Practical exercises will complete the theoretical presentation.

Contenu:

Introduction to natural language processing and their applications; Perl and XML; Introduction to linguistics (morphology, syntax, semantics); simple statistical approaches (KWIC, concordances); Automata and natural language (FSTN, RTN, ATN); Spelling detection and correction; Statistical models (counting words, bigramns, entropy); Markov chains; Hidden Markov chains; Information retrieval. The final mark is based on both a final written exam and the results of the practical exercises.

References

- Christopher D. Manning, Hinrich Schütze: Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge (MA).
- Pierre Nugues: An Introduction to Language Processing with Perl and Prolog. Springer, Berlin.

Forme de l'évaluation:

Examination two hours

Documentation:

Copies of the slides available.

Pré-requis:

None

Forme de l'enseignement:

2 hours of lectures and 2 hours of exercises