

# Natural Language Processing

## Rationale

Natural Language Processing (NLP) is one of the most important areas within Artificial Intelligence. It is deeply connected with Algorithms, Machine Learning, Programming Languages and Compiler Theory, and Automata and Formal Language Theory.

## Course Description

Computers process massive quantities of information every day in the form of human language, yet machine understanding of human language remains one of the great challenges of computer science. How can advances in computing technology enable more intelligent processing of all this language data? Will computers ever be able to use this data to learn language like humans do? This course provides a systematic introduction to statistical models of human language, with particular attention to the structures of human language that inform them and the structured learning and inference algorithms that drive them. This is a lecture course, not a seminar course, but aims to cover both fundamental and cutting-edge research issues.

## Topic List

Topics can include but are not limited to:

- Natural Language Understanding
- Hidden Markov Models
- Structured learning
- Inference Algorithms
- Coreference Resolution
- Keyphrase Extraction

- Extraction Based Summarization
- Text Mining

## **Learning Goals**

- be able to write simple programs that understand natural language text by implementing classical NLP algorithms such as Viterbi and CKY
- be able to understand the mathematical theory of noisy-channel model
- be able to understand the formal machineries of describing natural language, such as finite automata and context-free grammars
- be able to understand current NLP research

## **Assessment**

- Python Programming Exercises (50%)
- Quizzes (25%)
- Final Project (25%)