

# COMP 7022 NATURAL LANGUAGE UNDERSTANDING

**Credit Points** 10

**Legacy Code** 301313

**Coordinator** Manas Patra ([https://directory.westernsydney.edu.au/search/name/Manas Patra/](https://directory.westernsydney.edu.au/search/name/Manas%20Patra/))

**Description** Natural Language Understanding involves machine reading comprehension and the technologies using it are becoming increasingly widespread. This unit provides a foundation in using the Natural Language Toolkit, which is a leading platform for building Python programs working with 'human language' data, as well as an introduction to Python for Natural Language Processing. Students will use algorithms and explore accessing text corpora and processing raw text; categorising words and classifying text; understanding information from text and analysing sentence structures; and understanding semantic meanings of sentences. Students also gain real-world hands-on experience with Natural Language Understanding through the practical tasks and assignments.

**School** Computer, Data & Math Sciences

**Discipline** Artificial Intelligence

**Student Contribution Band** HECS Band 2 10cp

**Level** Postgraduate Coursework Level 7 subject

## Assumed Knowledge

There are no assumed knowledge for this subject although an undergraduate degree with some probability and statistics is advantageous.

## Learning Outcomes

On successful completion of this subject, students should be able to:

1. Apply the Natural Language Toolkit to real world Natural Language Processing tasks.
2. Use text corpora for the engineering and evaluation of Natural Language Processing systems.
3. Determine the role of classification of words and text in Natural Language Processing algorithms.
4. Clarify the function of Natural Language Processing algorithms in understanding information from text and the semantic meanings of sentences.
5. Analyse artificial intelligence techniques employed in Natural Language Processing algorithms.
6. Manage linguistic data.

## Subject Content

1. Installing and introducing the Natural Language Toolkit (NLKT):
  - Introducing the Python programming language with emphasise for Natural Language Processing (NLP);
  - Using Python string datatypes for the processing of words and texts.
2. Accessing text corpora (i.e., ?text repositories?f) and the processing of raw text:
  - Introduces the Gutenberg, Reuters, Inaugural Address and Annotated text corpus as well as Web and Chat text;
  - Looking at text corpus structure.
3. Categorizing words and classifying text:
  - Tagging corpora;

- Mapping words to properties through Python Dictionaries;
- Transformation based tagging;
- Determining the category of a word;
- Supervised, decision trees, naive bayes and ent

## Assessment

The following table summarises the standard assessment tasks for this subject. Please note this is a guide only. Assessment tasks are regularly updated, where there is a difference your Learning Guide takes precedence.

Item	Length	Percent	Threshold	Individual/Group Task
Quiz	1 hour (per Quiz)	30	Y	Individual
Practical Report	2 hours 1,500 words	30 40	N Y	Individual Individual

### Prescribed Texts

- Jurafsky, D., & Martin, J. H. (2020). Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition (3rd Draft ed.). Stanford. <https://web.stanford.edu/~jurafsky/slp3/>

### Teaching Periods

## Autumn

### Parramatta - Victoria Rd

#### Day

**Subject Contact** Manas Patra ([https://directory.westernsydney.edu.au/search/name/Manas Patra/](https://directory.westernsydney.edu.au/search/name/Manas%20Patra/))

View timetable ([https://classregistration.westernsydney.edu.au/event/timetable/?subject\\_code=COMP7022\\_22-AUT\\_PS\\_D#subjects](https://classregistration.westernsydney.edu.au/event/timetable/?subject_code=COMP7022_22-AUT_PS_D#subjects))