# SYLLABUS – A COURSE DESCRIPTION

## I. General information
1. Course name: **Natural Language Processing**
2. Course code: **NLP**
3. Course type (compulsory or optional): **compulsory**
4. Study programme name: **Language, Mind, Technology**
5. Cycle of studies (1st or 2nd cycle of studies or full master's programme): **2nd cycle**
6. Educational profile (general academic profile or practical profile): **general academic**
7. Year of studies (if relevant): **2nd**
8. Type of classes and number of contact hours: **lectures: 30 hours; practical classes: 30 hours**
9. Number of ECTS credits: **6**
10. Name, surname, academic degree/title of the course lecturer/other teaching staff: **Rafał Jaworski, PhD, rjawor@amu.edu.pl**
11. Language of classes: **English**
12. Online learning – yes (partly – online / fully – online) / no: **no**

## II. Detailed information
1. Course aim (aims):

   The aim of the subject is to introduce the students to the ideas of automatic natural language processing with the use of computer software. The first part of the course will describe the elementary techniques of language and text processing, including sentence splitting, tokenization, text normalization. The second part of the course will focus of techniques of deeper language understanding, such as shallow and deep parsing, classification algorithms based on machine learning and machine translation. After completing the course students will be able to implement basic of natural language processing and take advantage of more advanced tools.

2. Pre-requisites in terms of knowledge, skills and social competences (if relevant):

   The course requires basic knowledge of algorithms and computer programming.

3. Course learning outcomes (EU) in terms of knowledge, skills and social competences and their reference to study programme learning outcomes (EK):

| Course learning outcome symbol (EU) | On successful completion of this course, a student will be able to: | Reference to study programme learning outcomes (EK) |
|---|---|---|
| NLP_01 | Knows the basic applications of natural language processing, knows the limits of technological possibilities, can list the basic techniques of language processing | K_W07 K_W09 K_W11 K_U02 |
| NLP_02 | Write a regular expression, use regular expressions for practical tasks in the field of text processing, use regular expressions in Python. | K_W11 K_U08 K_U14 |
| NLP_03 | Knows the concept of formal grammar, can design his own grammar and modify an existing one, implements context-free grammar in Python. | K_W11 K_U08 K_U14 |
| NLP_04 | Run a Linux shell, knows the basic commands for text processing under Linux, understands the concept of the command pipeline. | K_W11 K_U08 K_U14 |
| NLP_05 | Knows the techniques of splitting a text into sentences and words, understands the challenges associated with it (especially in languages that do not use spaces), can define the rules of splitting into sentences and words. | K_W11 K_U08 |
| NLP_06 | Knows what the morphological analysis of a text is, knows electronic morphological dictionaries for various languages, knows disambiguation techniques. | K_W11 K_W13 K_U08 |
| NLP_07 | Knows the most popular NLP toolkits, can install and run them. | K_W11 K_U08 |

| NLP_08 | Knows the applications of machine learning techniques in language processing, can implement the Bayesian classifier using a ready implementation of the algorithm. | K_W11<br>K_U08<br>K_U16 |
|---|---|---|
| NLP_09 | Knows statistical regression techniques and their application in natural language processing. | K_W11<br>K_U08<br>K_U16 |
| NLP_10 | Understands the concept and knows the use of neural networks in language modeling and classification problems. | K_W11<br>K_W13<br>K_U08<br>K_U14 |
| NLP_11 | Knows statistical methods of language modeling, can train a model of any language based on text resources. | K_W11<br>K_U08<br>K_U16 |
| NLP_12 | Knows the techniques of automatic spell checking, can run a spell checker. | K_W11<br>K_U07<br>K_U08 |
| NLP_13 | Understands the concept of shallow parsing, can run a shallow parser. | K_W11<br>K_W13<br>K_U08 |
| NLP_14 | Knows the concept of deep parsing and syntax analysis. Can run a deep parser and analyse its results. | K_W11<br>K_W13<br>K_U08 |
| NLP_15 | Knows the techniques of automatic web scraping of language resources, can implement a simple web crawler. | K_W11<br>K_U08 |
| NLP_16 | Is able to cooperate with prospective employers. | K_W08<br>K_K05<br>K_K11 |

4. Learning content with reference to course learning outcomes (EU)

| Course learning content: | Course learning outcome symbol (EU) |
|---|---|
| Applications of natural language processing, basic terminology related to this field, the limits of NLP, running NLP programs. | NLP_01 |
| Regular expressions - defining the language of regular expressions, using it to search and modify text. | NLP_02 |
| Grammars and formal languages, building your own contextual grammar, using formal grammars for text analysis. | NLP_03 |
| Linux commands, the concept of a pipe, application in word processing. | NLP_04 |
| Text segmentation - splitting into sentences (segments) and words, SRX formalism, tokenization algorithms. | NLP_05 |
| Morphological analysis - morphological dictionaries, disambiguation problem, running tools for morphological analysis. | NLP_06 |
| NLP Toolkits - overview and application of the most popular NLP toolkits available for Python: NLTK and spaCy. | NLP_07 |
| Introduction to machine learning, the use of machine learning in language processing, Bayes classifier. | NLP_08 |
| Linear and logistic regression methods in natural language processing with examples of their practical application. | NLP_09 |
| The concept of neural networks, their role and application in natural language processing. | NLP_10 |
| Statistical methods of language modelling and their practical application. | NLP_11 |
| The problem of automatic spelling correction, the use of spelling correction programs in practice, the efficiency of algorithms | NLP_12 |
| Shallow parsing as an efficient method of superficial text analysis with applications. | NLP_13 |
| Deep parsing, differences from shallow parsing, application and examples. | NLP_14 |
| Automated download of linguistic resources from the Internet, practical exercises. | NLP_15 |
| Participation in a study visit. | NLP_16 |

5. Reading list:

- Edward Loper, Steven Bird, Ewan Klein: "Natural Language Processing with Python…

## III. Additional information

1. Teaching and learning methods and activities to enable students to achieve the intended course learning outcomes (please indicate the appropriate methods and activities with a tick and/or suggest different methods)

| Teaching and learning methods and activities | X |
|---|---|
| Lecture with a multimedia presentation | X |
| Interactive lecture | |
| Problem – based lecture | |
| Discussions | X |
| Text-based work | X |
| Case study work | |
| Problem-based learning | X |
| Educational simulation/game | |
| Task – solving learning (eg. calculation, artistic, practical tasks) | X |
| Experiential work | X |
| Laboratory work | X |
| Scientific inquiry method | |
| Workshop method | |
| Project work | X |
| Demonstration and observation | |
| Sound and/or video demonstration | |
| Creative methods (eg. brainstorming, SWOT analysis, decision tree method, snowball technique, concept maps) | |
| Group work | |
| Other (please specify) - | |
| … | |

2. Assessment methods to test if learning outcomes have been achieved (please indicate with a tick the appropriate methods for each LO and/or suggest different methods)

| Assessment methods | Course learning outcome symbol | | | | |
|---|---|---|---|---|---|
| | NLP_01-15 | NLP_01-06 | NLP_07-15 | | |
| Written exam | | | | | |
| Oral exam | | | | | |
| Open book exam | | | | | |
| Written test | x | | | | |
| Oral test | | | | | |
| Multiple choice test | | | | | |
| Project | | | x | | |
| Essay | | | | | |
| Report | | x | | | |
| Individual presentation | | | | | |
| Practical exam (performance observation) | | | | | |
| Portfolio | | | | | |
| Other (please specify) - | | | | | |
| … | | | | | |

3. Student workload and ECTS credits

| Activity types | | Mean number of hours spent on each activity type |
|---|---|---|
| Contact hours with the teacher as specified in the study programme | | 60 |
| Independent study* | Preparation for classes | 20 |
| | Reading for classes | 17 |
| | Essay / report / presentation / demonstration preparation, etc. | 25 |
| | Project preparation | |
| | Term paper preparation | |
| | Exam preparation | 20 |
| | Other (please specify) – study visit | 8 |
| | … | |
| Total hours | | 150 |
| Total ECTS credits for the course | | 6 |

* please indicate the appropriate activity types and/or suggest different activities

4. Assessment criteria in accordance with AMU in Poznan's grading system:
Very good (bdb; 5,0):
Good plus (+db; 4,5):
Good (db; 4,0):
Satisfactory plus (+dst; 3,5):
Satisfactory (dst; 3,0):
Unsatisfactory (ndst; 2,0):