

Course Specification Document

Title	Natural Language Processing
--------------	-----------------------------

Credits	5 ECTS
----------------	--------

Aims	This course aims to provide students with the necessary knowledge to create interactive applications with natural language processing and to help him understand the stages of morphological, syntactic, and semantic text processing, and identify the key issues in each stage. It also enables the student to understand various semantic representation methods, both traditional and those based on deep learning concepts, language models and various forms of semantic embedding.
-------------	---

Intended learning outcomes

On successful completion of this course, the student will be able to:

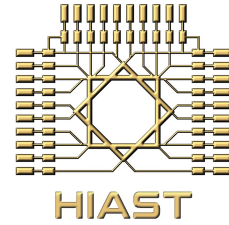
- Identify various applications of natural language processing systems.
- Grasp the fundamental concepts related to linguistic analysis at different levels.
- Familiarize himself with models used in morphological, syntactic, and semantic representation.
- Understand methods and techniques for representing the meaning of words and sentences.
- Acquire knowledge of language models and semantic embedding concepts.
- Utilize popular frameworks for natural language processing in both Arabic and English to build computer applications.
- Employ various semantic embedding models and comparing them.

Syllabus

- **Introduction to Natural Language Processing (NLP):** Definition of the NLP field, introduction to NLP applications and their main challenges, overview of the general structure of NLP applications, their fundamental issues, and the key tools aiding in their implementation.
- **Simple Text Processing:** Sentence tokenization, word tokenization, n-gram concepts, basic linguistic processing using regular expressions (RE).
- **Morphological Analysis:** Definition of morphological analysis, distinction between stemming and lemmatization, popular morphological analysis methods in both Arabic and English.
- **POS-Tagging:** Definition of POS Tagging, rule-based approaches, statistical methods, and machine learning methods.
- **Traditional Syntactic Analysis:** The difference between shallow and deep syntactic analysis, rule-based models, syntactic analysis methods.

Syrian Arab Republic

Higher Institute for Applied Sciences and Technology



- **Advanced Syntactic Analysis:** Shallow parsing, Named Entity Recognition (NER), Dependency Tree structures.
- **Semantic Processing:** Introduction to semantics and its representation, partial issues of semantic processing.
- **Ontology-based Semantic Representation:** Concept of ontology and its structures, popular available resources, semantic similarity measures based on ontology.
- **Vector Representation of Meaning:** Bag of Words (BOW) concept and representation using Vector Space Model (VSM).
- **Topic Modeling:** Use of Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA) in topic modeling.
- **Word-level Semantic Embedding:** The concept of embedding, famous linguistic models (contextual and non-contextual), principles of linguistic models and their importance.
- **Sentence-level Semantic Embedding:** Language models, use of Recurrent Neural Networks (RNN).
- **Advanced Concepts:** Fine-tuning, transformers, seq2seq models.