uc3m | Universidad Carlos III de Madrid

Natural language processing with deep learning

Academic Year: (2022 / 2023)   **Review date:** 28-04-2022

Department assigned to the subject: Computer Science and Engineering Department

Coordinating teacher: SEGURA BEDMAR, ISABEL

Type: Compulsory ECTS Credits : 3.0

Year : 1 Semester : 2

OBJECTIVES

- Create and train deep neural network architectures (such as Convolutional Neural Networks, Recurrent Neural Networks, LSTMs, Transformers) and apply them to tackle NLP applications such as text classification, machine translations, information extraction, text simplification and text summarization.
- Study, implement and use word embeddings to represents texts.

DESCRIPTION OF CONTENTS: PROGRAMME

1. Introduction
2. Basic NLP Tasks básicas de PLN.
   - Word embeddings. Text similarity.
3. Text Classification (CNN; RNN, Transformers)
4. Named Entity Recognition.
5. Relation Extraction.
6 Text simplification.
7 Text summarization.

LEARNING ACTIVITIES AND METHODOLOGY

The flippled classroom model will be applied:

Each week, the teacher will publish in aulaglobal a Jupyter notebook that describes and contains the implementation of a deep architecture for the resolution of a certain NLP application (for example, the detection of fake news or the anonymization of clinical notes). Students should study and run these notebooks before class.

During class, the teacher will explain the notebook and resolve any possible doubts. Once resolved, the teacher will propose possible improvements and raises new challenges or problems to solve. Students should try to extend or adapt the implementations of notebooks to solve the new challenges posed by the teacher.

The methodology is practical. The classes will have a configuration of work in pairs in a computer room.

Tutoring by appointment is limited to 2 hours per week. These can be online or face to face.

ASSESSMENT SYSTEM

CONTINUOUS ASSESSMENT. The activities planned during the classes will be valued.

In particular, it will be necessary to implement at least two of the architectures studied during the course and apply these architectures to solve some NLP task (for example, text classification).
Specifically, it will be necessary to participate in one of the competitions proposed at IberLef 2023 (https://sites.google.com/view/iberlef2022).

In addition, to participate in said competition, it will be necessary to write a scientific article (extension 4-8 pages), following the author guidelines proposed in the IberLef competition in which the students participate. The scientific article must describe the architectures, present and discuss the results obtained.

% end-of-term-examination: 0

% of continuous assessment (assigments, laboratory, practicals…):   100

BASIC BIBLIOGRAPHY
 - Delip Rao, Brian McMahan Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning, o'really, 2019
 - Delip Rao, Brian McMahan Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning, o'really, 2019
 - Li Deng (Redactor), Yang Liu (Redactor) Deep Learning in Natural Language Processing, SPRINGER, 2018

 - Palash Goyal, Sumit Pandey, Karan Jain Deep Learning for Natural Language Processing: Creating Neural Networks with Python, APRESS, 2018
 - Edward Loper, Steven Bird, Ewan Klein Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, 0'REALLY, 2009
ADDITIONAL BIBLIOGRAPHY
 - Amandeep Implement NLP use-cases using BERT: Explore the Implementation of NLP Tasks Using the Deep Learning Framework and Python, bpb, 2021
BASIC ELECTRONIC RESOURCES
 - Isabel Segura Bedmar . Repositorio de implementaciones para aplicaciones de PLN: http://github.com/isegura/