Universidad de Navarra

Procesamiento del Lenguaje Natural (MADI)

Guía docente 2026-27

PRESENTACIÓN

Breve descripción:

- **Titulación**: Máster Universitario en Análisis de Datos en Ingeniería
- **Módulo/Materia**: Análisis de Datos / Aprendizaje Automático
- **ECTS**: 4 ECTS
- **Curso, semestre**: 1º, Primero
- **Carácter**: Obligatorio
- **Profesorado**:
  - Hernantes Apezetxea, Josune - Email: jhernantes@tecnun.es / Catedrático
  - Pey Pérez, Yon - Email: ypey@external.unav.es / Invitado
- **Idioma**: Inglés

RESULTADOS DE APRENDIZAJE (Competencias)

- CG1 - Comprender y aplicar los algoritmos y herramientas que se utilizan en las aplicaciones de Inteligencia Artificial.
- CE05 - Comprender y aplicar los algoritmos de aprendizaje automático en la resolución de problemas, evaluando su rendimiento en base a las técnicas empleadas, los datos disponibles y el contexto en el que se aplica
- CE06 - Aplicar técnicas para extraer información de texto e imágenes

PROGRAMA

- NLP Introduction (2 sessions)
  - General Introduction
  - Application: Machine Learning and Sentiment Analysis
  - Preprocessing: Tokenizer, stop words, Stemming
- NLP and probabilistic models (2 sessions)
  - Edit distance: implementing and autocorrect
  - Hidden markov models: Part of Speech Tagging
  - Dynamic programming: n-grams and language models
- NLP and deep learning (3)
  - Deep learning intro (example neural network // backpropagation)
  - Word Embedding: training and evaluation
  - RNN
  - Transformers
  - Fine tuning, benchmark, reasoning
  - Infrastructure session
- Final course assignment (1 session)
  - RAG (Retrieval-Augmented Generation)
  - Exam & presentation (exam calendar)

ACTIVIDADES FORMATIVAS

La dedicación de 100-150h (5 ECTS) a la asignatura se divide en las siguientes actividades formativas:

Clases presenciales teóricas: 15 horas

Clases presenciales prácticas, laboratorios o talleres: 25 horas

Resolución de casos prácticos: 30 horas

Trabajos individuales y/o equipo: 25 horas

Tutorías: 2 horas

Estudio personal: 10 horas

Evaluación: 4 horas

EVALUACIÓN

CONVOCATORIA ORDINARIA

Continuous Evaluation (2 Points)

Each class starts with a brief quiz on the previous session's material. The lowest score will be dropped. The final score will be the average of the remaining quizzes.

Python Implementations (2 Points)

Students will complete challenges during the applied sections, submitting Python functions. T

Final Project (2 Points)

The final project involves implementing a Retrieval-Augmented Generation (RAG) system. Project details will be provided in the last class.

Exam (5 Points)

The exam, held on the date specified in the course syllabus, will consist of a theoretical section with questions similar to those addressed during the continuous evaluation, as well as an applied section in which students will complete a notebook similar to the one used in class.Students may bring all materials produced during the course, but they may not use the internet.

CONVOCATORIA EXTRAORDINARIA

Se realizarán dos proyectos diferentes a los de la convocatoria ordinaria.

HORARIOS DE ATENCIÓN

Dra. Josune Hernantes Apezetxea (jhernantes@unav.es)

- Despacho O101 Edificio Urdaneta Planta 1ª
- Horario de tutoria: The professors will be available to answer any doubts concerning the class topics with prior appointment made by e-mail.

Universidad de Navarra

BIBLIOGRAFÍA

Deep Learning with Python - Francois Chollet

http://bioserver.cpgei.ct.utfpr.edu.br/disciplinas/eeica/papers/Livros/%5BChollet%5D-Deep_Learning_with_Python.pdf

GitHub - fchollet/deep-learning-with-python-notebooks: Jupyter notebooks for the code samples of the book "Deep Learning with Python" Jupyter notebooks for the code samples of the book "Deep Learning with Python" - GitHub - fchollet/deep-learning-with-python-notebooks: Jupyter notebooks for the code samples of the book ... GitHub

Hands-On Machine Learning with Scikit-Learn & Tensorflow - Aurélien Géron

http://index-of.es/Varios-2/Hands%20on%20Machine%20Learning%20with%20Scikit%20Learn%20and%20Tensorflow.pdf

GitHub - ageron/handson-ml2: A series of Jupyter notebooks that walk you through the fundamentals of Machine Learning and Deep Learning in Python using Scikit-Learn, Keras and TensorFlow 2. A series of Jupyter notebooks that walk you through the fundamentals of Machine Learning and Deep Learning in Python using Scikit-Learn, Keras and TensorFlow 2. - GitHub - ageron/handson-ml2: A ser... GitHub

Natural Language Processing with Python

GitHub - Sturzgefahr/Natural-Language-Processing-with-Python-Analyzing-Text-with-the-Natural-Language-Toolkit: My solutions to selected exercises to "Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit" by Steven Bird, Ewan Klein, and Edward Loper. My solutions to selected exercises to "Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit" by Steven Bird, Ewan Klein, and Edward Loper. - GitHub -... GitHub

Deep Learning - Ian Goodfellow, Yoshoua Bengio, Aaron Courville

GitHub - PacktPublishing/Natural-Language-Processing-with-TensorFlow: Natural Language Processing with TensorFlow, published by Packt Natural Language Processing with TensorFlow, published by Packt - GitHub - PacktPublishing/Natural-Language-Processing-with-TensorFlow: Natural Language Processing with TensorFlow, published by Packt GitHub

Hands-On Natural Language Processing with Python: A practical guide to applying deep learning architectures to your NLP applications

GitHub - PacktPublishing/Hands-On-Natural-Language-Processing-with-Python: Hands On Natural Language Processing with Python, published by Packt Hands On Natural Language Processing with Python, published by Packt - GitHub - PacktPublishing/Hands-On-Natural-Language-Processing-with-Python: Hands On Natural Language Processing with Python, p... GitHub