NATIONAL UNIVERSITY OF SAN MARCOS

FACULTY OF MATHEMATICAL SCIENCES

SYLLABUS

1. GENERAL INFORMATION

1.1 Course name	:	Natural Language Processing
1.2 Code	:	AI262
1.3 Academic Program	:	Scientific Computing
1.4 Credits	:	4
1.5 Prerrequisites	:	AI263. Introduction to Machine Learning. (8^{th} Sem)
1.6 Type of Course	:	Elective
1.7 Area of Studies	:	Estudios Generales
1.8 Number of Weeks	:	16
1.9 Hours per week	:	2 HT; 4 HP;
1.10 Cycle	:	I
1.11 Learning modality	:	Face to face
1.12 Academic Period	:	2026-I
1.13 Semester Start Date	:	
1.14 Semester End Date	:	

2. PROFESSORS

Meetings after coordination with the professor

3. INTRODUCTION TO THE COURSE

This course introduces the theoretical and practical fundamentals of Natural Language Processing (NLP), covering from classical techniques to modern deep learning models. Students will learn to implement complete NLP pipelines for applications like sentiment analysis, machine translation, and chatbots.

4. GOALS

- Understand computational representation of human language.
- Implement text processing algorithms in Python.
- Apply NLP models to real-world problems using modern frameworks.

5. COMPETENCES

ABET-2) An ability to formulate or design a system, process, procedure, or program for the intended purpose. (Usage)

ABET-1) An ability to identify, formulate, and solve broadly defined technical or scientific problems by applying knowledge of mathematics and science and/or technical topics to areas relevant to the discipline. (Usage)

6. TOPICS

Unit 1: NLP Fundamentals (15 hours)	
Competences Expected: ABET-1,ABET-2	
Topics	**Learning Outcomes**
• Tokenization and text representation • N-gram models and smoothing • Feature extraction (TF-IDF, word embeddings)	• Preprocess text for NLP models [Usar] • Explain limitations of classical statistical models [Familiarizarse]
Readings : [**Jurafsky2021**], [**Manning1999**]	

Unit 2: Language Modeling (15 hours)	
Competences Expected: ABET-1,ABET-2	
Topics	**Learning Outcomes**
• Word2Vec, GloVe and FastText • Recurrent Neural Networks (RNNs) for NLP • Sequence-to-sequence models (seq2seq)	• Train word embeddings [Usar] • Implement basic RNNs for text generation [Evaluar]
Readings : [**Goldberg2017**], [**Mikolov2013**]	

Unit 3: Transformers and Modern NLP (15 hours)	
Competences Expected: ABET-1,ABET-2	
Topics	**Learning Outcomes**
• Transformer architecture (self-attention) • BERT and pre-trained models • Fine-tuning for specific tasks	• Apply pre-trained models with HuggingFace [Usar] • Evaluate NLP performance metrics [Evaluar]
Readings : [**Vaswani2017**], [**Devlin2019**]	

Unit 4: Practical Applications (15 hours)	
Competences Expected: ABET-1,ABET-2	
Topics	**Learning Outcomes**
• Sentiment analysis with transformers • Question-answering systems • Model deployment with Flask/FastAPI	• Build a complete NLP pipeline [Usar] • Optimize models for production [Evaluar]
Readings : [**Jurafsky2021**], [**Chollet2021**]	

7. WORKPLAN

7.1 Methodology

Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

7.2 Theory Sessions

The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

7.3 Practical Sessions

The practical sessions are held in class where a series of exercises and/or practical concepts are developed through problem solving, problem solving, specific exercises and/or in application contexts.

8. EVALUATION SYSTEM

********* EVALUATION MISSING ********

9. BASIC BIBLIOGRAPHY