



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
INSTITUTO DE INVESTIGACIÓN



TALLER PROCESAMIENTO DE LENGUAJE NATURAL: DESDE TF-IDF HASTA CHATGPT

FECHAS: MARTES Y VIERNES DEL 03 AL 28 DE MAYO del 2024

HORA: 03:00 a 6:00 PM

EXPOSITOR: Dr. Marco Sobrevilla Cabezudo

Doctor en Ciencias de la Computación por la Universidad de São Paulo, Brasil. Ingeniero de Sistemas FISI UNMSM. Amplia experiencia en investigación en Inteligencia Artificial, con énfasis en Procesamiento de Lenguaje Natural en el ámbito académico y la industria.

Meet:

PROGRAMACIÓN	
Viernes 03 de mayo. Presencial y virtual. Auditorio FISI, Alfredo Celso Bravo	Martes 07 de mayo
TEMA I. PROCESAMIENTO DE LENGUAJE NATURAL (PLN) Y REPRESENTACIONES VECTORIALES (1) <ul style="list-style-type: none">- Introducción a PLN- Representaciones dispersas (TF-IDF)- Aplicaciones en problemas de clasificación de textos	TEMA I. PLN Y REPRESENTACIONES VECTORIALES (2) <ul style="list-style-type: none">- Representaciones densas (word2vec, glove, fasttext)- Aplicaciones en problemas de clasificación de textos
Viernes 10 de mayo	Martes 14 de mayo
TEMA II. REDES NEURONALES (RN) RECURRENTES (1) <ul style="list-style-type: none">- Redes Neuronales Recurrentes (LSTM y GRU)- Aplicaciones de RNRs en tareas de clasificación	TEMA II. RNs RECURRENTES (2) <ul style="list-style-type: none">- Aplicaciones de RNRs en tareas de clasificación- Arquitectura Sequence-to-Sequence y aplicaciones
Viernes 17 de mayo	Martes 21 de mayo
TEMA III. TRANSFORMERS Y APLICACIONES (1) <ul style="list-style-type: none">- Qué es un Transformer?- Conociendo a la familia de transformers (BERT, T5, BART)	TEMA III. TRANSFORMERS Y APLICACIONES (2) <ul style="list-style-type: none">- Transformers (BERT, T5, BART)- Aplicaciones en tareas de clasificación y generación
Viernes 24 de mayo	Martes 28 de mayo
TEMA IV. GRANDES MODELOS DE LENGUAJE (1) <ul style="list-style-type: none">- Introducción a los grandes modelos del lenguaje: chatgpt, llama, flan-t5.	TEMA IV. GRANDES MODELOS DE LENGUAJE (2) <ul style="list-style-type: none">- Aplicaciones en tareas de Clasificación, Generación y Búsqueda- Problemas existentes

Clases: Teoría y Laboratorio. **Herramientas:** Python, Google Colaboratory, transformers, pytorch.

Organizan: Decanato, Instituto de Investigación, Departamentos Ciencias de la Computación e ISW

Dirigido especialmente a: Docentes FISI y UNMSM. Estudiantes de Posgrado FISI y UNMSM.