

**TEMAS AVANZADOS EN
CIENCIAS DE LA
COMPUTACIÓN
2024-1**

I. INFORMACIÓN GENERAL

CURSO	TEMAS AVANZADOS EN CIENCIAS DE LA COMPUTACIÓN
CLAVE	1INF12
CRÉDITOS	3
HORAS DE DICTADO	CLASE: 3 Semanal EXAMEN:
HORARIO	TODOS
PROFESORES	LUIS FERNANDO MUROYA TOKUSHIMA HECTOR ERASMO GOMEZ MONTOYA

II. PLANES CURRICULARES DONDE SE DICTA EL CURSO

ESPECIALIDAD	ETAPA	NIVEL	CARÁCTER	REQUISITOS
INGENIERÍA INFORMÁTICA	PREGRADO EN FACULTAD	0	ELECTIVO	1INF24 INTELIGENCIA ARTIFICIAL [04]

Tipos de requisito

- 04 = Haber cursado o cursar simultáneamente
- 05 = Haber aprobado o cursar simultáneamente
- 06 = Promedio de notas no menor de 08
- 07 = Haber aprobado el curso

III. DESCRIPCIÓN DEL CURSO

El procesamiento de lenguaje natural (en adelante, PLN) es una rama de la inteligencia artificial cuyo objetivo es dotar al computador de la capacidad de entender y generar lenguaje humano. Se trata de una rama que ha tenido un desarrollo vertiginoso en la última década, principalmente debido a la aparición de arquitecturas de aprendizaje profundo con desempeños muy superiores a los obtenidos con métodos tradicionales. En este curso se presentarán conceptos y técnicas del PLN con sus aplicaciones, poniendo énfasis en las técnicas que emplean aprendizaje profundo.

IV. SUMILLA

Este curso abordará las nuevas tendencias en el campo de Ciencias de la Computación y sus contenidos se adaptan a la evolución de los avances tecnológicos

V. OBJETIVOS

El curso contribuye al logro de los siguientes Resultados de Aprendizaje:

- RA1. Identifica y resuelve problemas clásicos de PLN empleando técnicas estadísticas y de aprendizaje profundo, evaluando fuentes bibliográficas y las implicancias éticas de dichas soluciones. (C1, C4, C6)
- RA2. Implementa y optimiza modelos de aprendizaje profundo que permitan resolver problemas complejos de predicción, traducción y generación de textos, analizando la implicancia ética de dichos desarrollos. C1, C2, C4, C6)
- RA3. Diseña e implementa sistemas de software que incorporen técnicas de PLN para dar solución de problemas a un usuario no especializado. (C1, C2, C4)

VI. PROGRAMA ANALÍTICO

CAPÍTULO 1 FUNDAMENTOS DEL PROCESAMIENTO DE LENGUAJE NATURAL Y LINGÜÍSTICA COMPUTACIONAL (3 horas)

Objetivo: Presentar conceptos básicos empleados en la lingüística computacional y en el procesamiento de lenguaje natural.

Contenido: Introducción. Lingüística computacional. Definición de conceptos básicos. Colocación. Concordancias. Procesamiento de Lenguaje Natural. Limpieza de textos. Tokenization, stemming y lemmatization. Stop-word removal. Problemas.

CAPÍTULO 2 REPRESENTACIÓN DE TEXTOS (3 horas)

Objetivo: Estudiar los diferentes métodos para representar textos.

Contenido: Introducción. Bag of Words Representation. TF-IDF. Word Embeddings. Word2Vect. GloVe. FastText.

CAPÍTULO 3 MODELOS DE LENGUAJE (3 horas)

Objetivo: Estudiar el concepto de modelo de lenguaje y los tipos que existen.

Contenido: Introducción. Modelos de lenguaje. Modelos estadísticos. Modelos neuronales simples. Aplicaciones.

CAPÍTULO 4 TAREAS BÁSICAS DE PROCESAMIENTO DE LENGUAJE NATURAL (6 horas)

Objetivo: Presentar las tareas elementales y recurrentes que pueden ser abordadas por el PLN.

Contenido: Introducción. Clasificación de texto. Identificación de lenguaje. Análisis de sentimientos. Part-of-speech tagging (POS-tagging). Name-Entity Recognition (NER). Identificación de tópicos. Resolución de preguntas (Question answering).

CAPÍTULO 5 DEEP LEARNING EN PROCESAMIENTO DE LENGUAJE NATURAL (9 horas)

Objetivo: Presentar las arquitecturas de aprendizaje profundo que pueden ser usadas para el PLN.

Contenido: Introducción. Problemas secuencia a secuencia (Seq2Seq). Redes neuronales recurrentes (RNN). Redes LSTM. Tuning. Atención. Transformers. Modelos generativos.

CAPÍTULO 6 TAREAS AVANZADAS DE PROCESAMIENTO DE LENGUAJE NATURAL (9 horas)

Objetivo: Abordar las tareas complejas que pueden ser abordadas por el PLN.

Contenido: Introducción. Traducción automática. Modelos de lenguaje basados en Transformers. BERT. Variantes de BERT. Generación de lenguaje natural.

CAPÍTULO 7 IMPLEMENTACIÓN DE CHATBOTS (6 horas)

Objetivo: Mostrar los conceptos y técnicas necesarias para construir un chatbot.

Contenido: Introducción. Chatbots básicos. Manejo de intenciones. Manejo de historias. Chatbox conversacionales con IA generativa.

VII. METODOLOGÍA

El curso es de carácter expositivo teórico práctico en donde el docente presenta la teoría de clase, acompañada con la resolución ejercicios prácticos dirigidos. Estos ejercicios serán desarrollados en laboratorio empleando el lenguaje de programación Python. En 5 sesiones se programarán ejercicios prácticos calificados que deberán ser resueltos de forma individual y que ayudarán a fortalecer los conceptos teóricos adquiridos y a aplicarlos en situaciones concretas. Durante el semestre, los alumnos desarrollarán de forma grupal un proyecto de curso donde deberán aplicar las diferentes técnicas y conocimientos adquiridos para resolver un problema específico. Ello permitirá que el estudiante valide su aprendizaje y desarrolle habilidades interpersonales.

VIII. EVALUACIÓN

FACULTAD DE CIENCIAS E INGENIERÍA
1INF12 - TEMAS AVANZADOS EN CIENCIAS DE LA COMPUTACIÓN

Sistema de evaluación

Nº	Codigo	Tipo de Evaluación	Cant. Eval.	Forma de aplicar los pesos	Pesos	Cant. Eval. Eliminables	Consideraciones adicionales	Observaciones
1	Nf	Nota Unica	1	Por Evaluación	Nf1=1			

Modalidad de evaluación: 4

Fórmula para el cálculo de la nota final

$$(1Nf1) / 1$$

Aproximación de la nota final No definido

Consideraciones adicionales

Fórmula para obtener la nota del curso será: $(2EC + 2Ex1 + 3Ex2 + 3Ta)/10$

EC: Evaluación continua en clase

Ex1: Examen Parcial.

Ex2: Examen Final.

Ta: Proyecto de curso.

IX. BIBLIOGRAFÍA

Referencia obligatoria

- Libro
Bengfort, Benjamin, 1984-
2018
Applied text analysis with Python [recurso electrónico] : enabling language-aware data products with machine learning
Sebastopol, CA : O'Reilly Media, 2018.ñ
[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:671345/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:671345/one)
- Libro
Cambria, Erik, ed.
2017
A Practical Guide to Sentiment Analysis [recurso electrónico]
Cham : Springer, 2017.
[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:666066/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:666066/one)
- Libro
Jurafsky, Daniel.
2000
Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition
Upper Saddle River, NJ : Prentice Hall, 2000
[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:365198/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:365198/one)
- Libro
Lee, G.G, ed.
2015
Natural Language Dialog Systems and Intelligent Assistants [recurso electrónico]
Cham : Springer, 2015.
[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:583427/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:583427/one)
- Libro
Liu, Bing, 1963-.

2016

Sentiment analysis : mining opinions, sentiments, and emotions

New York : Cambridge University Press, 2016.

[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:591582/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:591582/one)

- Libro

Manning, Christopher D.

1999

Foundations of statistical natural language processing

Cambridge, Mass. : MIT Press, 1999

[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:364829/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:364829/one)

Referencia complementaria

- Libro

Bishop, Christopher M.

2007

Pattern recognition and machine learning

New York : Springer, 2006.

[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:473859/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:473859/one)

- Libro

Goodfellow, Ian.

2016

Deep learning [recurso electrónico]

Cambridge, Massachusetts : The MIT Press, 2016.

[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:655391/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:655391/one)

- Libro

Sarkar, Dipanjan.

2016

Text Analytics with Python [recurso electrónico] : A Practical Real-World Approach to Gaining Actionable Insights from your Data

Berkeley, CA : Apress, 2016.

[https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:\\$002f\\$002fSD_ILS\\$002f0\\$002fSD_ILS:645995/one](https://pucp.ent.sirsi.net/client/es_ES/campus/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:645995/one)

X. POLÍTICA CONTRA EL PLAGIO

Para la corrección y evaluación de todos los trabajos del curso se va a tomar en cuenta el debido respeto a los derechos de autor, castigando severamente cualquier indicio de plagio con la nota CERO (00). Estas medidas serán independientes del proceso administrativo de sanción que la facultad estime conveniente de acuerdo a cada caso en particular. Para obtener más información, referirse a los siguientes sitios en internet

www.pucp.edu.pe/documento/pucp/plagio.pdf