

Programa de curso

Unidad Académica	:Centro de Informática Médica Y Telemedicina Programa de Anatomía y Biología del Desarrollo Centro de Informática Médica Y Telemedicina Programa de Anatomía y Biología del Desarrollo
Nombre del curso	:Métodos de Análisis Estadístico II
Nombre en inglés del curso	:Methods for Statistical Analysis II
Idioma en que se dicta	:Español
Código ucampus	:CAMAEL
Versión	:v. 2
Modalidad	:A distancia
Semestre	:1
Año	:2025
Días/Horario	:Lun 18:00-21:20, Vier 18:00-21:20, Mie 18:00-21:20,
Fecha inicio	:02/06/2025
Fecha de término	:09/07/2025
Lugar	:
Cupos mínimos	:3
Cupos máximo	:15
Créditos	:3

Tipo de curso

AVANZADO

Datos de contacto

Nombre	: Fabián Villena
Teléfono	: +56963442579
Email	: fabian@villena.cl
Anexo	:

Horas cronológicas

Presenciales:	: 0
A distancia:	: 92

Tipos de actividades(Horas directas estudiante)

Clases(horas)	: 30
Seminarios (horas):	: 0
Evaluaciones (horas)	: 1
taller/trabajo práctico	: 0
Trabajo/proyecto	: 1
investigación:	: 1
Créditos	: 3

PROFESOR ENCARGADO/A DEL CURSO (PEC)

Fabián Bernardo Villena Rodríguez

Docente Participantes	Unidad Academica	Función	Horas directas.	Horas indirectas.	Horas totales
-----------------------	------------------	---------	-----------------	-------------------	---------------

Fundamentos, Antecedentes que justifican la necesidad de dictar el curso

Una importante proporción de la ficha clínica se encuentra en formato de texto libre, lo cual dificulta la extracción de información clave y el uso secundario de datos. El texto clínico tiene desafíos particulares debido al uso extensivo de abreviaciones, la presencia de negación o la variabilidad del lenguaje clínico entre especialidades médicas. Por otro lado, el procesamiento del lenguaje natural (PLN) es una rama de la inteligencia artificial que permite la comunicación entre humanos y máquinas a través del lenguaje. Este curso es una introducción al PLN clínico.

Destinatarios

Estudiantes del programa de Magíster en Informática Médica. También está abierto a recibir estudiantes de otros programas de Magíster o Doctorado que tengan una formación inicial en programación.

Requisitos

Cursos del primer y segundo semestre del Magister en Informática Medica o contar con conocimientos de programación.

Resultado de aprendizaje

El estudiante comprenderá la teoría y aplicación en salud de los principales métodos de procesamiento de lenguaje natural utilizando el lenguaje de programación Python.

- Aplicar métodos de preprocesamiento y exploración de conjuntos de datos de texto.
- Aplicar métodos de representación de documentos.
- Aplicar métodos de modelamiento predictivo con datos de texto.

Metodologías de enseñanza y aprendizaje

Cantidad

Clase teórica

30

Metodologías de evaluación

Cantidad

Duración horas

Ponderación

Prueba teórica

1

1

30.0 %

Informe, trabajo o proyecto de investigación

1

1

70.0 %

Suma (Para nota presentación examen)

100.0 %

Total %

%

Requisitos de aprobación y asistencia.Suma ponderada de la prueba y el trabajo ≥ 4.0

Unidades

Unidad: Métodos de preprocesamiento y exploración de conjuntos de datos de texto

Encargado: Fabián Bernardo Villena Rodríguez

Logros parciales de aprendizajes:

Aplicar métodos de preprocesamiento y exploración de conjuntos de datos de texto.

Acciones Asociadas:

Clases teóricas expositivas participativas, demostraciones de código programado en Python y talleres de resolución de problemas de procesamiento de lenguaje natural en salud utilizando programas en Python.

Contenidos:

- Preprocesamiento de datos de texto - Exploración y descripción de datos de texto

Unidad: Métodos de representación de documentos

Encargado: Fabián Bernardo Villena Rodríguez

Logros parciales de aprendizajes:

Aplicar métodos de representación de documentos

Acciones Asociadas:

Clases teóricas expositivas participativas, demostraciones de código programado en Python y talleres de resolución de problemas de procesamiento de lenguaje natural en salud utilizando programas en Python.

Contenidos:

- Semántica vectorial - Word embeddings

Unidad: Métodos de modelamiento predictivo con datos de texto

Encargado: Fabián Bernardo Villena Rodríguez

Logros parciales de aprendizajes:

Aplicar métodos de modelamiento predictivo con datos de texto

Acciones Asociadas:

Clases teóricas expositivas participativas, demostraciones de código programado en Python y talleres de resolución de problemas de procesamiento de lenguaje natural en salud utilizando programas en Python.

Contenidos:

- Redes neuronales recurrentes - Transformers - Grandes modelos de lenguaje

Bibliografía							
Caracter	Título	Autor	Edición	Idioma	Formato	Vínculo(Url)	Fecha de consulta
Obligatorio	Clinical Text Mining	H. Dalianis	1ra	Inglés	Libro digital	https://www.spr...	16/01/2021
Complementario	Speech and Language Processing	D. Juraksky y J. H. Martin	3ra	Inglés	Libro digital	https://doi.org...	16/01/2021

Plan de clases					
Fecha	Horario	Actividad	Condición	Tema	Profesor(es)
2025-06-02,Lun	18:00 - 21:20	Clase 1	Libre	Introducción al procesamiento de lenguaje natural clínico	Fabián Bernardo Villena Rodríguez
2025-06-06,Vier	18:00 - 21:20	Clase 2	Libre	Preprocesamiento y exploración de texto	Fabián Bernardo Villena Rodríguez
2025-06-09,Lun	18:00 - 21:20	Clase 3	Libre	Semántica vectorial	Fabián Bernardo Villena Rodríguez
2025-06-13,Vier	18:00 - 21:20	Clase 4	Libre	Word embeddings	Fabián Bernardo Villena Rodríguez
2025-06-16,Lun	18:00 - 21:20	Clase 5	Libre	Redes neuronales recurrentes	Fabián Bernardo Villena Rodríguez
2025-06-18,Mie	18:00 - 21:20	Clase 6	Libre	Transformers	Fabián Bernardo Villena Rodríguez
2025-06-23,Lun	18:00 - 21:20	Clase 7	Libre	Grandes modelos de lenguaje y sus habilidades emergentes	Fabián Bernardo Villena Rodríguez
2025-06-25,Mie	18:00 - 21:20	Clase 8	Libre	Cuantización y fine-tuning eficiente para grandes modelos de lenguaje	Fabián Bernardo Villena Rodríguez
2025-07-02,Mie	18:00 - 21:20	Clase 9	Libre	Agentes y Retrieval Augmented Generation	Fabián Bernardo Villena Rodríguez
2025-07-09,Mie	18:00 - 21:20	Examen	Libre	Examen	Fabián Bernardo Villena Rodríguez