

### Programa de curso

|                            |   |
|----------------------------|---|
| Unidad Académica           | :Centro de Informática Médica Y Telemedicina<br>Departamento de Medicina Interna Sur<br>Centro de Informática Médica Y Telemedicina<br>Departamento de Medicina Interna Sur |
| 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             | :CAMAELI  |
| Versión                    | :v. 1   |
| Modalidad                  | :A distancia  |
| Semestre                   | :1  |
| Año                        | :2021   |
| Días/Horario               | :Lun 18:00-21:20, Mar 18:00-21:20, Mie 18:00-21:20, Jue 18:00-21:20   |
| Fecha inicio               | :08/06/2021   |
| Fecha de término           | :13/07/2021   |
| Lugar                      | :Zoom   |
| Cupos mínimos              | :4  |
| Cupos máximo               | :35   |
| Créditos                   | :3  |

#### Tipo de curso

AVANZADO

#### Datos de contacto

|          |                      |
|----------|----------------------|
| Nombre   | : Jocelyn Dunstan    |
| Teléfono | : 0944772030         |
| Email    | : jdunstan@uchile.cl |
| Anexo    | : 89641              |

#### Horas cronológicas

|               |       |
|---------------|-------|
| Presenciales: | : 0   |
| A distancia:  | : 120 |

#### Tipos de actividades(Horas directas estudiante)

|                         |      |
|-------------------------|------|
| Clases(horas)           | : 30 |
| Seminarios (horas):     | : 0  |
| Evaluaciones (horas)    | : 0  |
| taller/trabajo práctico | : 0  |
| Trabajo/proyecto        | : 2  |
| investigación:          | : 3  |
| Créditos                | : 3  |

**PROFESOR ENCARGADO/A DEL CURSO (PEC)**

Dunstan Escudero Jocelyn Mariel

| Docente Participantes | Unidad Académica                               | Función               |
|-----------------------|--|-----------------------|
| Pablo Baez            | Centro de Informática Médica Y<br>Telemedicina | Profesor Participante |
| Fabián Villena        | Centro de Informática Médica Y<br>Telemedicina | Profesor Participante |

**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**

Alumnos de magíster en informática médica Alumnos libres

**Requisitos**

Conocimientos básicos de Python

**Resultado de aprendizaje**

- Poder preprocesar, visualizar, y hacer clasificación básica de textos clínicos
- Entender la importancia de crear recursos lingüísticos para analizar textos clínicos producidos en Chile y cuáles son las técnicas para ello
- Comprender de donde emerge el uso de redes neuronales para el análisis de texto
- Comprender y calcular word embeddings
- Tener una visión global del área de procesamiento del lenguaje natural en medicina, conociendo los problemas clásicos así como los desafíos actuales

**Metodologías de enseñanza y aprendizaje**

Clase teórica

Cantidad

30

| Metodologías de evaluación                   | Cantidad | Duración horas                              | Ponderación |
|--|----------|---|-------------|
| Informe, trabajo o proyecto de investigación | 1        | 2   | 100.0 %     |
|  |          | <b>Suma (Para nota presentación examen)</b> | 100.0 %     |
|  |          | <b>Total %</b>                              | %           |

**Requisitos de aprobación y asistencia.**

Sin requisitos de asistencia. Aprobación mayor o igual a 4.0 en el trabajo final.

## Unidades

### Unidad: Preprocesamiento de texto

Encargado: Dunstan Escudero Jocelyn Mariel

Logros parciales de aprendizajes:

- Conocer y manejar las herramientas básicas del preprocesamiento de texto
- Poder realizar visualizaciones básicas de texto
- Aprender a usar Python para el preprocesamiento de texto

Acciones Asociadas:

- Clases expositivas
- Sesiones prácticas de programación

Contenidos:

### Unidad: Anotación de textos clínicos

Encargado: Pablo Baez

Logros parciales de aprendizajes:

- Entender la necesidad de crear recursos lingüísticos para el texto clínico
- Aprender cómo se comparan distintos etiquetados
- Conocer a BRAT Rapid Annotation Tool

Acciones Asociadas:

- Clases expositivas
- Sesiones prácticas de programación

Contenidos:

### Unidad: Redes neuronales para analizar texto

Encargado: Dunstan Escudero Jocelyn Mariel

Logros parciales de aprendizajes:

- Entender el concepto de word embedding y saber como construirlos y testarlos
- Entender por qué las redes recurrentes son usadas para analizar texto
- Aprender a evaluar modelos de deep learning
- Conocer cuales son los problemas actuales en el campo de procesamiento del lenguaje natural en medicina

Acciones Asociadas:

- Clases expositivas
- Sesiones prácticas de programación

Contenidos:

| Bibliografía   |                                |                            |         |        |               |   |                   |
|----------------|--------------------------------|----------------------------|---------|--------|---------------|---|-------------------|
| Caracter       | Título                         | Autor                      | Edición | Idioma | Formato       | Vínculo(Url)  | Fecha de consulta |
| Obligatorio    | Clinical Text Mining           | H. Dalianis                | Primera | Inglés | Libro digital | <a href="https://www.spr...">https://www.spr...</a> | 16/01/2021        |
| Complementario | Speech and Language Processing | D. Juraksky y J. H. Martin | Tercera | Inglés | Libro digital | <a href="https://doi.org...">https://doi.org...</a> | 16/01/2021        |

| Plan de clases |               |           |             |   |   |
|----------------|---------------|-----------|-------------|---|---|
| Fecha          | Horario       | Actividad | Condición   | Tema  | Profesor(es)                                      |
| 2021-06-08,Mar | 18:00 - 21:20 | Clase 1   | Libre       | Introducción al texto clínico en medicina                               | Dunstan Escudero<br>Jocelyn Mariel                |
| 2021-06-10,Jue | 18:00 - 21:20 | Clase 2   | Libre       | Preprocesamiento de texto   | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-06-15,Mar | 18:00 - 21:20 | Clase 3   | Libre       | Visualización y clasificación de texto                                  | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-06-17,Jue | 18:00 - 21:20 | Clase 4   | Libre       | Word embedding  | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-06-21,Lun | 18:00 - 21:20 | Clase 5   | Libre       | Anotación de textos   | Pablo Baez  |
| 2021-06-23,Mie | 18:00 - 21:20 | Clase 6   | Libre       | Introducción a redes recurrentes  | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-06-29,Mar | 18:00 - 21:20 | Clase 7   | Libre       | Transformers y métricas de evaluación                                   | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-07-01,Jue | 18:00 - 21:20 | Clase 8   | Libre       | Problemas en Chile de texto clínico y consultas de proyectos personales | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-07-05,Lun | 18:00 - 21:20 | Clase 9   | Libre       | Desafíos actuales en área de procesamiento del lenguaje natural clínico | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |
| 2021-07-13,Mar | 18:00 - 20:00 | Examen    | Obligatoria | Examen  | Dunstan Escudero<br>Jocelyn Mariel;Fabián Villena |