

Nombre de la unidad curricular: Análisis de datos para ciencias

Forma parte de la Oferta Estable: No

Licenciaturas: Astronomía, Biología humana, Bioquímica, Ciencias Biológicas, Ciencias de la atmósfera, Estadística, Física, Física médica, Geología, Geografía, Matemática, Licenciatura en Recursos Naturales y Desarrollo Sustentable, Cartografía, Tecnicatura en Recursos Naturales y Desarrollo Sustentable

Frecuencia y semestre de la formación al que pertenece: Bienal, semestre par

Créditos asignados: 12

Nombre del/la docente responsable: Gabriel Illanes

E-mail: gillanes@cmat.edu.uy

Requisitos previos: - Conocimientos básicos de probabilidad

- Conocimiento básico de álgebra lineal

- Conocimiento básico de cálculo

Ejemplos de unidades curriculares de Facultad de Ciencias u otros que aportan dichos

conocimientos: - Probabilidad: Probabilidad (matemática y física), Bioestadística (otras carreras de FCIEN), Probabilidad y Estadística (FING), Probabilidad 1 (FCEA)

- Álgebra lineal: Álgebra lineal 1 (matemática), Álgebra Lineal y Geometría I (física), Matemática 2 (otras carreras de FCIEN), Geometría y álgebra lineal 1 (FING), Álgebra lineal (FCEA)

- Cálculo: Cálculo 1 (matemática), Cálculo diferencial e integral 1 (física), Matemática 2 (otras carreras de FCIEN), Cálculo Diferencial e Integral en una Variable (FING), Cálculo 1 (FCEA)

Conocimientos adicionales sugeridos:

Conocimientos avanzados de cálculo y álgebra lineal. Conceptos básicos de estadística. Conocimientos de computación (en particular usando el software R).

Objetivos de la unidad curricular:

a) Herramientas, conceptos y habilidades que se pretenden desarrollar

La idea del curso es aprender conceptos básicos de modelos de regresión y clasificación. Se pretende, principalmente, dar un enfoque intuitivo y aplicado.

También se enseñará a usar herramientas avanzadas del software R, tanto para correr los modelos, como para la parte de tratamiento de datos y visualización.

b) En el marco del plan de estudios

Temario sintético de la unidad curricular:

Parte 1: Introducción y herramientas avanzadas del software R.

Parte 2: Modelos de regresión.

Parte 3: Modelos de clasificación.

Temario desarrollado:

Parte 1: Introducción y herramientas avanzadas del software R.

- Conceptos básicos sobre ciencia de datos
- Repaso de conceptos básicos de probabilidad, álgebra lineal y software R (de ser necesario)
- Estudio de herramientas avanzadas del software R (entre ellas, el paquete Tidyverse)

Parte 2: Modelos de regresión.

- Regresión lineal simple y múltiple. Limitaciones del modelo
- Validación cruzada
- Modelos lineales sesgados
- Si da el tiempo: análisis de componentes principales y máxima verosimilitud

Parte 3: Modelos de clasificación.

- Motivación, consistencia y error de Bayes.
- Regresión logística, Análisis de varianza
- Análisis discriminante lineal
- Vecinos más cercanos
- Modelos basados en árboles

Bibliografía

a) Básica:

An Introduction to Statistical Learning (with applications in R): James, Witten, Hastie, Tibshirani.
R for Data Science: Wickham, Golemund

b) Complementaria:

Modalidad cursada: Híbrida

Metodología de enseñanza:

Duración en semanas: 15

Carga horaria total: 180

Carga horaria detallada:

a) Horas aula de clases teóricas: 45

b) Horas aulas de clases prácticas: 22.5

c) Horas de seminarios:

d) Horas de talleres:

e) Horas de salida de campo:

f) Horas sugeridas de estudio domiciliario durante el período de clase: 112.5

Mantiene horarios 2021: No

Sistema de APROBACIÓN final

Tiene examen final: Si

Se exonera el examen final: No

Sistema de GANANCIA

a) Características de las evaluaciones:

La ganancia del curso es automática. Durante el curso, los estudiantes (en grupos de hasta 2 personas) desarrollarán un proyecto de análisis de datos, usando herramientas enseñadas durante el curso (puede incluir la incorporación de herramientas más avanzadas, si el proyecto lo requiere).

El examen del curso es obligatorio, y consiste en la defensa del proyecto, una vez esté terminado y listo para presentar. La defensa es grupal y dura aproximadamente una hora.

b) Porcentaje de asistencia requerido para ganar la unidad curricular: 0

c) Puntaje mínimo individual de cada evaluación y total: 0 (ver descripción de la ganancia del curso)

d) Modo de devolución o corrección de pruebas:

Habilitada a rendir en calidad de examen libre: No

COMENTARIOS o ACLARACIONES: