

Nombre: Ciencia de Datos y Biología: entre la Explicación y la Predicción

Puntaje: 2 puntos. **Doctorado** Arancel: categoría 3

Correlatividades: no consigna

Carácter: no consigna

Carga horaria: 4hs semanales (2 de teoría+2 problem) **40 hs totales**

Carreras: Doctorado

Programa analítico del curso

Objetivos : El principal objetivo de este curso es que los alumnos y las alumnas se familiaricen con diferentes técnicas de modelado estadístico (no abarcadas por otras materias del área) y comprendan que la elección de la misma dependerá de la naturaleza del fenómeno bajo análisis, de los objetivos del estudio y de las particularidades del set de datos con el que se cuenta. Cada técnica será presentada mediante el planteo de un problema biológico, simulando una situación real en cuanto a preguntas de investigación y obtención de los datos. De esta manera, los conceptos teóricos involucrados detrás de cada análisis se entrelazan con la aplicación e interpretación de los resultados en contexto.

- 1- Adecuación y formato de base de datos. Limpieza y curado. Exploración de datos. Datos faltantes. Tratamientos posibles.
- 2- Métodos de validación cruzada.
- 3- Concepto de Robustez. Regresión robusta.
- 4- Métodos de regularización. Regresión Lasso.
- 5- Modelos aditivos generales y generalizados.
- 6- Métodos basados en árboles.

Bibliografía:

Alberti, J., & Massone, O. (2022). Tired of losing valuable data? Build your lab ecological database as a cornerstone for long-term approaches. *Ecología Austral*, 32(1), 151–157.

Donders, A. R. T., Van Der Heijden, G. J. M. G., Stijnen, T., & Moons, K. G. M. (2006). A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, 59(10), 1087–1091.

Gareth, J., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. In *An Introduction to Statistical Learning*. Springer.

Holmes, S. H., & Huber, W. (2018). Modern statistics for modern biology. Cambridge University Press.

(<http://web.stanford.edu/class/bios221/book/Chap-Supervised.html>)

Maronna, R. A., Martin, R. D., Yohai, V. J., & Salibián-Barrera, M. (2019). Robust statistics: theory and methods (with R). John Wiley & Sons.