
Asignatura: “Técnicas de Aprendizaje Automático”

(Esta asignatura es parte del curso de posgrado Estadística Multivariada Computacional (PM60))

Profesor de la asignatura ¹: Dr. Mathias Bourel (Gr. 3, IMERL)

Instituto ó Unidad: Instituto de Matemática y Estadística “Prof. Ing. Rafael Laguardia”.

Departamento ó Area: Laboratorio de Probabilidad y Estadística.

Horas presenciales: 27

Nº de Créditos: 5

Público objetivo y Cupos:

Estudiantes de la Licenciatura en Estadística, Maestría en Ingeniería Matemática y otros estudiantes de posgrado interesados.

Objetivos: Introducir los aspectos metodológicos de algunas técnicas modernas de estadística multivariada.

Aplicar dichas técnicas a conjuntos de datos reales e interpretar los resultados obtenidos.

Acercar el estudiante al empleo de los paquetes para el análisis estadístico de datos disponibles en el ambiente de desarrollo de software libre R (<http://www.r-project.org/>).

Conocimientos previos exigidos: Un curso de Análisis Multivariado.

Conocimientos previos recomendados: Uso del software estadístico R.

Metodología de enseñanza:

(comprende una descripción de las horas dedicadas por el estudiante a la asignatura y su distribución en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)

Horas clase (teórico): 15

•Horas clase (práctico): 5

•Horas clase (laboratorio):

•Horas consulta: 5

•Horas evaluación: 2

•Subtotal horas presenciales: 27

•Horas estudio: 15

•Horas resolución ejercicios/prácticos: 10

•Horas proyecto final: 25

•Total de horas de dedicación del estudiante: 77

Forma de evaluación: carpeta de ejercicios y trabajo final a entregar con defensa oral.

Temario:

1. Repaso: Análisis de datos. Problemas de clasificación. Clasificación Bayesiana.
2. Introducción a las técnicas de Aprendizaje Automático. Principios básicos.
3. Árboles de clasificación y regresión (CART).
4. k vecinos más cercanos.
5. Support Vector Machines.
6. Métodos de agregación de modelos: Bagging, Boosting, Random Forest.
7. Aprendizaje no supervisado.

Bibliografía:

- **Bourel, M. (2012)** Model aggregation methods and applications. *Memorias de trabajos de difusión científica y técnica*, Vol. 10, p. 19-32, 2012.
 - **Breiman L. (2001)**. Random forests. *Machine Learning* 45 (1): pp 5–32.
 - **Breiman L., Friedman J, Stone CJ & RA Olshen (1984)** *Classification and Regression Trees*. Wadsworth Internacional Group, Belmont, CA.
 - **Hastie T., Tibshirani R and Friedman J (2011)**. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. (5th. Edition). Springer Seris in Statistics.
 - **James, G., Witten, D., Hastie, T., Tibshirani, R. (2013)** An introduction to Statistical Learning with Applications in R, Springer.
 - **Peña, D. (2002)**. Análisis de datos multivariantes, Daniel Peña, Mac Graw Hill
-

Fecha de inicio y finalización: 24/10/17 al 30/11/17 (6 semanas)

Horario: Martes y Jueves de 8 a 10.

Salón: Instituto de Matemática y Estadística, Facultad de Ingeniería.