
Formulario de Aprobación Curso de Posgrado
Asignatura: ESTADISTICA AVANZADA Y APLICACIONES
(Si el nombre contiene siglas deberán ser aclaradas)

Profesor de la asignatura ¹: Dr. Mathias Bourel, Gr. 3, IMERL
Dra. Carolina Crisci, Gr. 3, PDU MAREN (Servicio Referencia: FING), CURE.
Dr. Angel Segura, Gr. 3, PDU MAREN (Servicio Referencia: FING), CURE.

Profesor Responsable Local ¹: Dr. Mathias Bourel, Gr. 3, IMERL
(título, nombre, grado, Instituto)

Otros docentes de la Facultad:
(título, nombre, grado, Instituto)

Docentes fuera de Facultad:
(título, nombre, cargo, Institución, país)

Instituto ó Unidad: PDU Modelización y Análisis de Recursos Naturales, CURE -
Laboratorio de Probabilidad y Estadística, IMERL, Facultad de Ingeniería
Departamento ó Area:

Fecha de inicio y finalización: 28/03/2016-08/07/2016
Horario y Salón: Centro Universitario Regional Este, sede Rocha

Horas Presenciales: 60
(se deberán discriminar las mismas en el ítem de Metodología de enseñanza)

Créditos: 10

Público objetivo y Cupos: Estudiantes de Posgrado y docentes de Disciplinas diversas con interés en "Machine Learning" y sus aplicaciones.

Objetivos: El curso de Estadística Avanzada y Aplicaciones está destinado a docentes y estudiantes de los diversos grupos de trabajo y carreras, en particular de posgrado, del CURE. Se tratarán temas que pretenden dar herramientas para fortalecer el análisis estadístico en trabajos de grupos de investigación, así como de estudiantes en formación.

El curso constará mayormente de exposiciones por parte de los docentes sobre técnicas estadísticas de interés para el análisis de datos. Se dará un espacio importante a la ilustración de las técnicas mediante la presentación de aplicaciones y se promoverá la participación activa de los participantes. Habrá instancias prácticas en donde se aplicarán las técnicas expuestas utilizando el software R.

En caso de que surja la necesidad por parte de los asistentes, se dictarán clases introductorias de R de modo de brindar herramientas para la realización de los trabajos prácticos y del trabajo final.

Por otro lado se contará con exposiciones por parte de investigadores nacionales sobre algunas aplicaciones de las técnicas vistas en el curso en distintas áreas de interés de los participantes.

Conocimientos previos exigidos: PROBABILIDAD Y ESTADISTICA (nivel de grado), experiencia de trabajo con bases de datos

Conocimientos previos recomendados:

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 estudiantes)

- Horas clase (teórico): 45 horas de teórico-práctico
- Horas clase (práctico): no corresponde
- Horas clase (laboratorio): no corresponde
- Horas consulta: 13
- Horas evaluación: 2
- Subtotal horas presenciales: 60

- Horas estudio: 25
- Horas resolución ejercicios/prácticos: 10
- Horas proyecto final/monografía: 20
- Total de horas de dedicación del estudiante: 115

Forma de evaluación: Entrega de ejercicios y trabajo final con defensa oral obligatoria.

Temario:

Módulo 1: Análisis de ordenación

Repaso de Algebra Lineal
Distribución Normal Multivariada
Análisis de Componentes Principales
Clustering
Análisis de Correspondencia

Módulo 2: Análisis clásicos de Regresión y Clasificación

Regresión lineal simple y múltiple
Modelos Lineales Generalizados - GLM
Análisis Discriminante

Módulo 3: Técnicas avanzadas de Regresión y Clasificación

Introducción al Aprendizaje Automático
Modelos Aditivos Generalizados (GAM)
Árboles de Clasificación y Regresión (CART)
Métodos de agregación de modelos: Bagging, Boosting, Stacking
Bosques Aleatorios- (Random Forest)
Máquinas de Vectores de Soporte (SVM)

Bibliografía:

Bourel, M. (2012) Model aggregation methods and applications. *Memorias de trabajos de difusión científica y técnica*, Vol. 10, p. 19-32, 2012.

Crisci C, Ghattas B, Perera G (2012). A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling* 240: 113-122.

De'Ath, G. et al (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis.

Everitt y Torsten (2010). A Handbook of Statistical Analyses Using R, 2nd edition, Brian S., Chapman & Hall/CRC.

Hastie T., Tibshirani R and Friedman J (2011). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. (5th. Edition). Springer Seris in Statistics.

Izenman, AJ. (2008). *Modern Multivariate Statistical Techniques. Regression, Classification and Manifold Learning*. Springer Texts in Statistics.

James, G., Witten, D., Hastie, T., Tibshirani, R. (2013) An introduction to Statistical Learning with Applications in R, Springer.

Legendre, P., Legendre, L. (1998). Numerical Ecology. Elsevier Science

Peña, D. (2002). Análisis de datos multivariantes, Daniel Peña, Mac Graw Hill