



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Ref. Expte. N° 1188/2020

Ciudad Autónoma de Buenos Aires, 28 de septiembre de 2020

VISTO:

La nota presentada por la Dirección del Instituto de Cálculo, mediante la cual eleva la información del curso de posgrado **Taller Avanzado de Consultoría Estadística** para el año 2020,

CONSIDERANDO:

lo actuado por la Comisión de Doctorado,
lo actuado por la Comisión de Posgrado,
lo actuado por este Cuerpo en la sesión realizada en el día de la fecha,
en uso de las atribuciones que le confiere el Artículo 113° del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD
DE CIENCIAS EXACTAS Y NATURALES
RESUELVE:**

ARTÍCULO 1°: Aprobar el dictado del curso de posgrado **Taller Avanzado de Consultoría Estadística** de 56 horas de duración, que será dictado por los Dres. María Eugenia Szretter Noste y Ricardo Maronna.

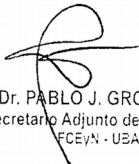
ARTÍCULO 2°: Aprobar el programa del curso de posgrado **Taller Avanzado de Consultoría Estadística** para su dictado en el segundo cuatrimestre de 2020.

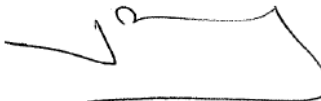
ARTÍCULO 3°: Aprobar un puntaje máximo de dos (2) puntos para la Carrera del Doctorado.

ARTÍCULO 4°: Disponer que de no mediar modificaciones en el programa y la carga horaria, el presente Curso de Posgrado tendrá una vigencia de cinco (5) años a partir de la fecha de la presente Resolución.

ARTÍCULO 5°: Comuníquese a todos los Departamentos Docentes, a la Dirección de Estudiantes y Graduados, a la Biblioteca de la FCEyN y a la Secretaría de Posgrado con copia del programa incluida. Cumplido, archívese.

RESOLUCIÓN CD N° 0819


Dr. PABLO J. GROISMAN
Secretario Adjunto de Posgrado
FCEyN - USA


Dr. JUAN CARLOS REBOREDA
DECANO

Taller Avanzado de Consultoría Estadística

Objetivos:

El objetivo del Taller es poner a los alumnos en situaciones que semejan las que se encuentran en consultoría. Cada situación se basa en un caso real. Se describen los datos y las condiciones en que fueron obtenidos, y los objetivos de un supuesto “comitente”, expresados en términos del área del problema (no en términos estadísticos). Los alumnos deben convertir el problema original en un problema estadístico, lo que implica plantear modelos de la situación; y luego realizar análisis de datos e inferencia. Por último deben redactar un informe, expresando su solución en términos comprensibles y útiles para el “comitente”. Este puede -representado por el profesor- suministrar eventualmente aclaraciones e información adicional. Una vez entregado el informe, el profesor -junto con los alumnos- realiza la crítica del mismo, pudiendo pedir eventualmente un replanteo parcial o total del trabajo.

Es de notar que en estas situaciones no hay una “solución verdadera”. La corrección del trabajo consiste en procurar que éste sea racional y coherente, y adecuado para los objetivos del comitente.

Programa:

1. Introducción al estudio de conjunto de datos con gran número de variables o de alta dimension.
2. Análisis discriminante clásico; regularización para datos de alta dimension. Análisis discriminante diagonal y método de centroides encogidos.
3. Regresión logística. Se introducirán diferentes métodos como: métodos basados en árboles: CART, Bosques aleatorios, Vecinos más cercanos (NN).
4. Regresión lineal para datos de alta dimension en particular se introducirán los estimadores Ridge y Lasso.
5. Introducción al modelo complejo, modelos aditivos y modelos aditivos generalizados
6. Presentación de análisis no supervisado, componentes principales y estudio de métodos adaptados a datos de alta dimension: Componentes “ralas”. Curvas principales.

Cantidad de horas de clase: 56.

Cantidad de semanas: 14.

Requisitos: Familiaridad con los métodos de inferencia y análisis de datos en Modelo Lineal, y conocimientos básicos de Análisis Multivariado (Lineal).

Bibliografía

1. Hastie-Tibshirani-Friedman. *The Elements of Statistical Learning_ Data Mining, Inference, and Prediction*. Springer, 2013.
2. James-Witten-Hastie-Tibshirani. *An Introduction to Statistical Learning with Applications in R*. Springer, 2014.
3. Matloff. *Statistical Regression and Classification_ From Linear Models to Machine Learning*. Chapman & Hall, 2017.

4. Johnson, R. A. y Wichern, D.W. (2007). Applied Multivariate Statistical Analysis. Prentice Hall.
5. Cook, D. and Swayne, D.F. (2007). Interactive and Dynamic Graphics for Data Analysis. Springer UseR! Series.
6. Izenman, A. (2008). Modern Multivariate Statistical Techniques, Regression, Classification and Manifold Learning. Springer New York.
7. Murray, S. (2013). Interactive Data Visualization for the Web. O'Reilly.
8. Ward, M.O., Grinstein, G. and Keim, D. (2015). Interactive Data Visualization, 2nd edition. Chapman & Hall/CRC.