

1 Información del equipo pedagógico y horario atención a estudiantes

Profesor: Ignacio Sarmiento-Barbieri (i.sarmiento@uniandes.edu.co)

- Horario Clase: Sábados 8:00 a.m. – 10:50 a.m. (cada dos semanas)
- Web del Curso: Bloque Neón
- Horario de atención a estudiantes: Via Slack o hacer cita en [este link](#)

Profesor Complementario: Gustavo Adolfo Castillo Álvarez (ga.castillo@uniandes.edu.co)

- Horario Clase: Miércoles, 18:00 p.m. – 19:20 p.m. Virtual
- Horario de atención a estudiantes: Via Slack o hacer cita en [este link](#)

2 Descripción del curso

Este es un curso con un enfoque especial en herramientas relevantes para economistas y ciencias sociales. Está destinado a estudiantes interesados en investigación aplicada y/o análisis de datos grandes y no estructurados. Problemas de predicción e inferencia, con especial énfasis en inferencia causal, atraviesan transversalmente al curso.

Mediante una combinación de contenido asincrónico, clases sincrónicas magistrales y complementarias, talleres grupales, y quices los estudiantes adquirirán las herramientas estadísticas y computacionales necesarias para responder varias preguntas en economía y en una gran cantidad de subcampos en investigación aplicada. Se hará énfasis especial en el análisis de datos reales, y la aplicación de metodologías específicas; ejemplos incluyen encuestas de hogares, precios de propiedades, datos de internet y redes sociales.

3 Resultados de aprendizaje

- Aplicar las técnicas provenientes de la ciencia de datos, la ciencia computacional, y la estadística usando la visión de la economía para resolver problemas puntuales identificados en el contexto.
- Contrastar los distintos algoritmos y su conveniencia para contestar preguntas económicas y sociales con base en criterios relacionados con la naturaleza de problemas económicos y sociales.
- Implementar procesos técnicos para el manejo cuantitativo de datos que surgen de distintas fuentes: páginas web, encuestas, geoespaciales, texto, etc, para resolver problemas económicos y sociales

- Generar conclusiones y recomendaciones sobre preguntas relevantes a las ciencias sociales por medio del manejo, análisis y síntesis de bases de datos con gran número de observaciones y variables.
- Aplicar el software R y su ecosistema para análisis estadístico, de big data y machine learning.

4 Cronograma

	Tema
Modulo 1 (2 semanas)	Introducción a big data y machine learning
Modulo 2 (2 semanas)	Sobreajuste, validación cruzada y bootstrap. Adquisición de datos web
Modulo 3 (2 semanas)	Selección de modelos y regularización
Modulo 4 (2 semanas)	Clasificación
Modulo 5 (2 semanas)	Árboles, bosques y boosting
Modulo 6 (2 semanas)	Redes neuronales y aprendizaje profundo.
Modulo 7 (2 semanas)	Datos espaciales y Superlearners
Modulo 8 (2 semanas)	Texto como datos y reducción de dimensión no supervisada

5 Referencias

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An introduction to statistical learning (ISLR)
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

6 Metodología

La metodología del curso combina lecturas, clases sincrónicas magistrales bisemanales, clases complementarias semanales, talleres grupales, y quices.

Se espera que los estudiantes realicen las actividades asincrónicas antes de cada clase sincrónica magistral, asistan a las mismas, estudien las lecturas disponibles en la página del curso y repliquen

las aplicaciones presentadas por el profesor. Todas las actividades asincrónicas están pensadas para facilitar el aprendizaje y posterior desarrollo de las actividades individuales y grupales.

La evaluación del aprendizaje se realizará a través de quices individuales y tres talleres prácticos grupales que servirán para que los estudiantes pongan a prueba los conocimientos adquiridos y desarrollen las habilidades necesarias para realizar trabajo en equipo.

7 Evaluaciones

- Quices. Los estudiantes tendrán 8 quices individuales al finalizar los módulos en Bloque Neón que evaluarán el aprendizaje individual. El peso de los quices será de 5% para un peso total de 40%.
- Talleres. Los estudiantes realizarán trabajos prácticos grupales donde los grupos no podrán superar los 4 miembros. Habrá 3 talleres durante el cursado. Los talleres serán entregados vía Bloque Neón y deberán contar con un repositorio en GitHub. Se espera que todos los miembros hagan contribuciones al repositorio del taller. La calificación del taller se verá reducida si no hay evidencia de contribución de todos los miembros. El peso de los talleres será del 20% para un peso total de 60%.

Sistema de aproximación de notas definitiva

Las calificaciones definitivas de las materias serán numéricas de uno cinco (1,50) a cinco (5,00), en unidades, décimas y centésimas. La calificación aprobatoria mínima será de tres (3,0). En este curso se aproximará la nota a la centésima más cercana. Por ejemplo, si el cálculo del cómputo es 3.245, la nota final se aproximará a 3.25; si el resultado del cálculo es 2.994 la nota final será de 2.99

8 Asistencias

Se espera que los estudiantes asistan a todas las clases. Si los asistentes faltan a más del 20% sin excusa válida de las clases en las que nos reuniremos se penalizará hasta un 10% la nota final del curso.

9 Políticas generales de los cursos de Economía y fechas importantes

Los estudiantes deben consultar [este enlace](#), donde se encuentran las reglas sobre asistencia a clase, excusas válidas, fraude académico y faltas disciplinarias, reclamos, políticas de bienestar y fechas importantes del semestre.