

1. Horario atención a estudiantes, correos electrónicos y nombres de los profesores complementarios

Clase magistral

Profesor:	Ignacio Sarmiento-Barbieri
Correo:	i.sarmiento@uniandes.edu.co
Sitio web:	https://github.com/ECON-4676-UNIANDES
Horario y lugar de atención:	Via Slack y a concretar con el profesor
Profesor asistente:	Jacob Muñoz Castro
Correo:	j.munoz14@uniandes.edu.co
Horario y lugar de atención:	Via Slack y a concretar con el profesor

2. Introducción y descripción general del curso

Se buscará introducir a los alumnos a un conjunto de herramientas estadísticas, matemáticas, y computacionales para abordar problemas de gran cantidad/tipos/calidad de datos ("large n"), y cantidad de variables ("large p"). Mediante una combinación de conjuntos de talleres, presentaciones, exámenes, y un trabajo final grupal, los estudiantes adquirirán las herramientas estadísticas y computacionales necesarias para hacer uso de big data y machine learning en investigación empírica. Es pre-requisito haber cursado Microeconomía 3 y Econometría 1. Se recomienda experiencia con programación en R, aunque no es requisito. Si bien no es requisito tener experiencia con R, si es requisito tener mucha voluntad de aprender y experimentar. ¡Este programa (y todos) se aprende utilizándolos!

3. Objetivos de la materia

El objetivo de este curso es introducir a los alumnos a un conjunto de herramientas estadísticas, matemáticas, y computacionales para abordar problemas de gran cantidad/tipos/calidad de datos ("large n"), y cantidad de variables ("large p"). Problemas de predicción e inferencia, con especial énfasis en inferencia causal, atravesarán transversalmente al curso. Se buscará también familiarizar a los alumnos con la literatura reciente que utiliza estas herramientas.

4. Organización del curso

Temario (sujeto a cambios)

1. Introduction to ML: prediction and inference. Supervised and unsupervised learning. MCO revision. Goodness of fit. Introduction to R, Jupyter Lab, Github, and AWS

2. New Economic Observation. Search and computer-mediated behavior. Text Data: news media and social media. Large N Problems: compute and processing. Web scrapers and APIs.
3. Observing from above: Introduction to spatial econometrics. Modeling spatial dependence. Processing big spatial/satellite data, raster data.
4. Intro to non parametric econometrics. Kernels, densities, and non parametric regressions. The curse of dimensionality.
5. Classification: Bayes Risk, Logit Models, ROC analysis.
6. Non lineal methods: Clusters, PCA, K-means, Tress, Boosting and Random Forests, Support vector machines.
7. Bonus Track: Machine Learning for Causal Inference and Deep Learning.

5. Metodología

La metodología del curso combina clases virtuales, talleres, una propuesta de trabajo, y un examen final. La participación de los estudiantes es fundamental para sacar el mayor provecho del curso. La virtualidad impone nuevos desafíos y es importante mantenerse conectados para crear las sinergias que surgen de las interacciones humanas. A su vez, los estudiantes realizarán trabajos prácticos grupales para evaluar su aprendizaje.

El producto final de este curso es un plan de trabajo con una propuesta de cómo implementar los conceptos y herramientas aprendidas a un problema concreto. Finalmente, habrá un examen final donde se evaluarán los conceptos y habilidades aprendidas a lo largo del cursado.

6. Competencias

- Comprender las técnicas provenientes de la ciencia de datos, la ciencia computacional, y la estadística desde una visión de economistas
- Ser capaz de contrastar distintas técnicas econométricas y su conveniencia para contestar preguntas económicas y sociales.
- Desarrollar habilidades técnicas para el manejo cuantitativo de datos que surgen de distintas fuentes: texto, geoespaciales, encuestas, paginas web, etc.
- Desarrollar la capacidad manejar, analizar y sintetizar bases de datos con gran numero de observaciones y variables para generar conclusiones y recomendaciones sobre preguntas relevantes a las ciencias sociales.
- Aprender a manejar con fluidez distintas herramientas computacionales.
- Capacidad de analizar críticamente los datos presentados, ya sean provenientes de textos académicos o de los ejercicios presentados en el cursado
- Trabajar y resolver situaciones en grupo
- Habilidad de exponer y defender (en forma escrita y oral) el trabajo realizado

7. Criterios de evaluación (Porcentajes de cada evaluación)

	Puntaje Individual	Puntaje Total
Participación		10%
Talleres	10%	40%
Propuesta de Trabajo		25%
Primer Entrega	5%	
Segunda Entrega	5%	
Tercer Entrega	15%	
Examen Final		25%
Total		100%

Participación. Si bien participación es la actividad con menos peso en la composición final, será el "tiebreaker" por el cual decidire la nota final. Participación no incluye solamente la asistencia a clases, sino también actividades fuera de clase. Una vez registrados en el curso los estudiantes recibirán invitación al canal de Slack, al aula virtual de AWS y a github. La participación será juzgada en función a la participación en las discusiones, en los trabajos grupales, de las interacciones en el canal de Slack, el aprovechamiento de AWS y se espera que estudiantes encuentren al menos un error de tipeo o cualquier otro tipo y los arreglen a través de pull requests en github.

Talleres. Los estudiantes realizarán trabajos prácticos grupales para evaluar su aprendizaje. Los grupos no podrán superar los 3 miembros. Habrá 4 talleres durante el semestre. Se dedicarán al menos 4 clases para la discusión y presentación de los talleres. Los talleres serán submitidos vía github y parte de nota de la participación saldrá de la evaluación de la historia del repositorio donde se verá la contribución de cada estudiante.

Propuesta de trabajo.

La actividad estará dividida en 3 entregas. En la primera entrega los grupos se reunirán conmigo y presentarán brevemente (máximo 5 slides) la idea y como planean llevarla a cabo. En una segunda entrega donde se expondrán los datos propuestos. La entrega final será al concluir el curso que consolida todo el trabajo. Se otorgarán bonos a los estudiantes que además de presentar el plan de trabajo o propuesta, entreguen resultados concretos.

Examen final. Este examen pretende evaluar los conceptos y habilidades aprendidas en el curso. Va a ser un examen domiciliario, de tiempo fijo, entre 48-72 horas.

8. Sistema de aproximación de notas definitiva

Reclamos. Los estudiantes podrán presentar reclamos sobre sus evaluaciones por medio de una carta (o email) dirigida al profesor. Dicha carta debe explicar clara y en detalladamente los motivos del reclamo. Luego del análisis de la petición por parte del profesor, se concertará una reunión con el estudiante para revisar la evaluación. De acuerdo a los artículos 64, 65 y 66 del Reglamento general de estudiantes de pregrado, el estudiante tendrá **cuatro** días hábiles después de la entrega de la evaluación calificada para presentar un reclamo. El profesor responderá al reclamo en los **cinco** días hábiles siguientes. Si el estudiante considera que la respuesta no concuerda con los criterios de evaluación podrá solicitar un segundo calificador al Consejo de la Facultad en los **cuatro** días hábiles posteriores a la recepción de la decisión del profesor.

Notas finales. Las calificaciones definitivas de las materias serán numéricas de uno punto cinco (1.50) a cinco punto cero (5.00), en unidades, décimas y centésimas.

9. Bibliografía (*sujeta a cambios*)

- Anselin, Luc. 2003. "An Introduction to Spatial Regression Analysis in R." Available at: https://geodacenter.asu.edu/drupal_files/spdepintro.pdf.
- Anselin, Luc, and Anil K Bera. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics." *Statistics Textbooks and Monographs* 155. MARCEL DEKKER AG: 237–90.
- Arbia, Giuseppe. 2014. *A Primer for Spatial Econometrics: With Applications in R*. Palgrave Macmillan.
- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. High-dimensional methods and inference on structural and treatment effects. *The Journal of Economic Perspectives* 28.2 (2014): 29-50.
- Bivand, Roger S, Edzer J Pebesma, and Virgilio Gomez-Rubio. 2008. *Applied Spatial Data Analysis with R*. Springer. Springer.
- Boehmke, B., & Greenwell, B. M. (2019). *Hands-on machine learning with R*. CRC Press.
- Cárdenas Rubio, J. A., Chaux Guzmán, F. J., & Otero, J. (2019). Una base de datos de precios y características de vivienda en Colombia con información de Internet. *Revista de Economía del Rosario*, 22(1), 25.
- Davidson, R., & MacKinnon, J. G. (2004). *Econometric theory and methods* (Vol. 5). New York: Oxford University Press.
- Einav, Liran, and Jonathan D. Levin. The data revolution and economic analysis. No. w19035. National Bureau of Economic Research, 2013.
- Farrell, D., Greig, F., and Deadman, E. (2020). Estimating family income from administrative banking data: A machine learning approach. *AEA Papers and Proceedings*, 110:36-41.
- Gentzkow, Matthew, and Jesse M. Shapiro. What drives media slant? Evidence from US daily newspapers. *Econometrica* 78.1 (2010): 35-71.
- Gentzkow, M., and J. Shapiro. Nuts and Bolts: Computing with Large Data. NBER Lecture and Videos.
- Glaeser, E.L., Hillis, A., Kominers, S.D. and Luca, M., 2016. Predictive Cities Crowdsourcing City Government: Using Tournaments to Improve Inspection Accuracy. *The American Economic Review*, 106(5), pp.114-118.
- Glaeser, E. L., Kominers, S. D., Luca, M., and Naik, N. (2018). Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, 56(1):114-137.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and generalizations*. CRC press.
- Kuhn, M. (2012). *The caret package*. R Foundation for Statistical Computing, Vienna, Austria. URL [https://cran.r-project.org/package= caret](https://cran.r-project.org/package=caret).
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.
- Mullainathan, S. and Spiess, J., 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), pp.87-106.
- Sosa Escudero, W. (2019). *Big Data*. Siglo Veintiuno Editores.
- Temperature and Temperament: Evidence from a Billion Tweets. Patrick Baylis. Job Market Paper. 2015.
- Varian, H.R., 2010. Computer Mediated Transactions. *The American Economic Review*, 100(2), pp.1-10.
- Varian, Hal R. Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives* 28, no. 2 (2014): 3-28.
- What will the Impact of Machine Learning be on Economics? Forbes Magazine [#14906ea67ad5l](http://www.forbes.com/sites/quora/2016/01/27/what-will-the-impact-of-machine-learning-be-on-economic).

10. Fechas importantes:

Inicio de clases: **10 de agosto.**

Semana de receso: **del 5 a 9 de octubre.**

Fecha de entrega del 30% de las notas: **23 de octubre.**

Último día de clases: **5 de diciembre.**

Exámenes finales: **7 al 17 de diciembre.**

Último día para solicitar retiros: **24 de diciembre (6:00 p.m.).**

Último día para subir notas finales en banner: **23 de diciembre.**

Último día para que los estudiantes retiren cursos: **24 de diciembre.**

11. Cláusula de ajustes razonables

Si lo considera pertinente, siéntase en libertad de informar al profesor lo antes posible **si usted tiene alguna condición, visible o invisible**, por la cual requiera algún ajuste para estar en igualdad de condiciones con los y las demás estudiantes. Debido a las actuales circunstancias, barreras de conectividad o acceso a los recursos tecnológicos indispensables para la clase son parte de las condiciones que pueden requerir ajustes. Por la misma razón, no necesitará presentar documentación para solicitar esos ajustes.

También lo invitamos a buscar asesoría y apoyo en la Dirección de su programa, en la Decanatura de Estudiantes (<http://centrodeconsejeria.uniandes.edu.co>, Bloque Ñf, ext. 2207, 2230 y 4967, horario de atención L-V 8:00 a.m. a 5:00 p.m.) o en el Programa de Acción por la Igualdad y la Inclusión Social (PAIIS) de la Facultad de Derecho (pais@uniandes.edu.co). Si su solicitud se basa en dificultades de acceso a conectividad o tecnología, es particularmente importante que haga este contacto adicional para que pueda acceder a los recursos de apoyo que brinda la Universidad.

Se entiende por ajustes razonables todas "las modificaciones y adaptaciones necesarias y adecuadas que no impongan una carga desproporcionada o indebida, cuando se requieran en un caso particular, para garantizar a las personas con discapacidad el goce o ejercicio, en igualdad de condiciones con las demás, de todos los derechos humanos y libertades fundamentales" Convención sobre los Derechos de las personas con discapacidad, art.2.

12. Cláusula de respeto por la diversidad

Todos debemos respetar los derechos de quienes hacemos parte de esta comunidad académica. En esta comunidad consideramos inaceptable cualquier situación de acoso, acoso sexual, discriminación, matoneo, y/o amenaza. La persona que se sienta en alguna de estas situaciones puede denunciar su ocurrencia y buscar orientación y apoyo ante alguna de las siguientes instancias: el equipo pedagógico del curso, la Coordinación o la Dirección del programa, la Decanatura de Estudiantes (DECA, Ed. Ñf-Casita amarilla), la Ombudsperson (ombudsperson@uniandes.edu.co, Edificio RGA–Pedro Navas, Of. 201, ext. 5300 y 3933) o el Comité MAAD (lineamaad@uniandes.edu.co, <https://uniandes.edu.co/MAAD> o a la ext. 2707 o 2230). Si quieren mayor información, guía o necesitan activar el protocolo MAAD pueden acudir a Nancy García (n.garcia@uniandes.edu.co) en la Facultad. También puede acudir a los grupos estudiantiles que pueden ofrecerle apoyo y acompañamiento: No Es Normal (derechoygenero@uniandes.edu.co o <https://www.facebook.com/noesnormaluniandes/?fref=ts>); Pares de Acompañamiento Contra el Acoso-PACA (paca@uniandes.edu.co o <https://www.facebook.com/PACA-1475960596003814/?fref=ts>).

Para mayor información sobre el protocolo MAAD, puede visitar esta página: <https://decanaturadeestudiantes.uniandes.edu.co/index.php/es/sobre-la-decanatura/827>

13. Política de Momentos difíciles

Todas las personas pueden pasar por un momento difícil que de alguna manera pueda afectar nuestra vida en la Universidad. Pueden ser problemas en casa, con la pareja, incluso estrés por esta u otra materia. Si usted siente que está pasando por un momento complicado, **sin importar el motivo**, siéntase con la tranquilidad de hablar con la profesora para pedir tiempo o apoyo. Ningún trabajo o entrega puede sobrepasar su salud mental y física. Su bienestar es lo más importante.