



MACHINE LEARNING AND PUBLIC POLICY

ECON 4555

Álvaro Riascos

ariascos@uniandes.edu.co

Universidad de los Andes

Edward McFowland

emcfowla@umn.edu

Universidad de Minnesota

ESCUELA DE VERANO 2018

Salón: PTE

Fechas: 25 de junio al 7 de julio. Clase sábado 7 de julio

Horario:

Álvaro Riascos: 25-29 junio 2:00 a 5:00 pm

Monitoría: 5:00 a 6:00 pm

Edward McFowland: 3-6 julio 2:00 a 5:00 pm

Monitoría: 5:00 a 6:00 pm

Edward McFowland: 7 julio (sábado): 8:00 a 11:00 am

Monitoría: 11:00 a 12:00 pm

Programa sujeto a cambios

OBJECTIVES

The age of Big Data has catalyzed the maturation and widespread adoption of machine learning. It is now more critical than before that practitioners across the social science disciplines become familiar with machine learning methods, understanding both the potential and limits of these methods in the social sciences. Therefore, the first objective of this course is to introduce participants to modern machine learning methods. This will provide an understanding of the fundamental building blocks upon which these methods are built; enabling participants to be insightful practitioners of modern machine learning. The second objective of this course is to introduce some of the recent and advanced developments of machine learning with a specific focus on public policy. This will provide an understanding of the role of machine learning in improving public policy, and address challenges across the social science problem spectrum. Participants should leave this course empowered to think critically about the intersection of machine learning and social science, and develop machine learning solutions that inform and solve real-world policy problems.

Prerequisite Knowledge: Familiarity with any computer programming language and basic econometrics for causal inference.

SYLLABUS

DATE	TOPIC	PRESENTER	REFERENCES
June 25	<ul style="list-style-type: none"> • Introduction to Data Mining • Foundations of statistical learning: Concepts, methods and results 	Álvaro Riascos	[JWHT]: Chapter 1. [HTF]: Chapter 1. [LS] [JWHT]: Sections 2.1, 2.2
June 26	Linear methods of clasification and regression Regularitzation Application: Prevention on unnessary hospitalization	Álvaro Riascos	[JWHT]: Chapter 3,4. [HTF]: Chapter 3,4.
June 27	Validation, model selection. Cross validation. Bagging, bootstrapping. Ensembles Trees, random forests, boosting. Application: Risk adjustment in competitive health insurance system	Álvaro Riascos	[HTF]: Capítulos 7 y 8 [JWHT]: Capítulo 5
June 28	Unsupervised learning: KDE, clustering, associative rules. Application: Crime prediction, anomaly detection in claims data	Álvaro Riascos	
June 29	Neural networks Deep neural networks (advanced) Application: TBA Handout 1	Álvaro Riascos	
July 3	Classification and Prediction for Policy and Development	Edward McFowland III	
July 4	Text Based Modeling	Edward McFowland III	
July 5	Anomaly Detection, Event Detection, and Surveillance Systems	Edward McFowland III	
July 6	Machine Learning for Causal Inference	Edward McFowland III	
July 7	Machine Learning for Casual Inference (Advanced) Handout II Proposal final project	Edward McFowland III	

METHODOLOGY:

This is an applied course and requieres students to learn basic programing in R or Phytion. They are required to make groups of 2 to 4 people and turn in the following:

- Handout 1 (30% of the final grade)
- Handout 2 (30% of the final grade)

- Final project proposal (10% of the final grade).
- Final project (30% of the grade): Maximum 7 pages including tables, graphs, references, etc.
To turn in July 15th.

GRADING SYSTEM:

Grades will be approximated to the nearest decimal: .0 or .5.

REFERENCES

The main references for this course are:

[LS]: Luxburg, U., B. Scholkopf. 2008. Statistical Learning Theory: Models, Concepts and Results.

<http://arxiv.org/abs/0810.4752>

[JWHT]: Introduction to Statistical Learning with Applications in R.

<http://www-bcf.usc.edu/~gareth/ISL/>

[HTF]: Hastie, T., Tibshirani, R. y J. Hastie. 2009. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Segunda Edición. Springer

http://web.stanford.edu/~hastie/local.ftp/Springer/OLD/ESLII_print4.pdf

[AP]: Joshua Angrist and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics. Princeton University press.

FECHA DE RETIRO:

The student may withdraw the course, without refund, up to one business day before the date of the final test stipulated by the teacher. The University will not return the money for tuition paid for these summer courses.