Universidad de los Andes
Facultad de Economía

Dates: 25 of June - 7 of July
Hour: 2PM - 5PM (theory lecture), 5PM – 6PM (practical session)
Language: English and Spanish

Professor Edward McFowland III
University of Minnesota
E-mail: emcfolwla@umn.edu
Webpage: https://carlsonschool.umn.edu/faculty/edward-mcfowland-iii

Profesor: Alvaro J. Riascos Villegas
University of los Andes
E-mail: ariascos@uniandes.edu.co
Webpage: www.alvaroriascos.com

1. 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: Familarity with any computer programing language and basic econometrics for causal inference.

2. Contents

Date	Topic	Presenter	References
June 25	Introduction to Data Mining. Foundations of statistical learning: Concepts, methods and results. Applications: ML & economics.	Alvaro Riascos Hamadys Benavides (75 minutes)	[JWHT]: Chapter 1. [HTF]: Chapter 1. [LS] [JWHT]: Sections 2.1, 2.2.
June 26	Linear methods of classification and regression regularization. Validation,	Alvaro Riascos	[JWHT]: Chapter 3,4. [HTF]:

	model selection, cross validation, bagging, bootstrapping. Ensembles.	Hamadys Benavides (45 minutes)	Chapter 3,4.
June 27	Trees, random forests, boosting. Applications: Public health policy (risk adjustment, hospitalizations, renal chronic disease).	Alvaro Riascos Hamadys Benavides (45 minutes)	[HTF]: Capítulos 7 y 8 [JWHT]: Capítulo 5
June 28	Unsupervised learning: KDE, clustering, associative rules. Application: Crime prediction, anomaly detection in claims data.	Alvaro Riascos Hamadys Benavides (30 minutes) Simon Ramirez (60 minutes)	
June 29	Neural networks Deep neural networks (advanced) Application: Image classification **Handout 1**	Alvaro Riascos Simon Ramirez (75 minutes)	
July 3	Classification and Prediction for Policy and Development	Edward McFowland III Simon Ramirez (60 minutes)	
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	

3. Methodology

This is an applied course and requieres students to learn basic programing in R or Phyton. 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 proyect proposal (10% of the final grade).
- Final project (30% of the grade): Maximun 7 pages including tables, graphs, references, etc. **To turn in July 15th.**

4. Grading system

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

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 Minning, 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.