TRAINING PROGRAM: MACHINE LEARNING FOR ECONOMIC ANALYSIS AND FORECASTING WITH SAS & PYTHON

Tutor: Rodolfo Méndez Marcano (PhD in Statistics, MSc in Economics)

PART I: USER-FRIENDLY & PRACTICAL INTRO (8 sessions)

- Introduction: 1 session
    - Clearing Up the terminology Morass: Data Mining, Big Data, DSc, ML, DL, AI
    - What's is and isn't useful for an economist to know?
    - Bias Versus Variance Trade-Off
    - Machine Learning Project Life Cycle
- SAS Overview: 1 session
    - Procedure Approach
    - Task Approach
    - **Visual Pipeline Approach (SAS-VDMML)**

- Data Preparation by Example via SAS-VDMML: 1 Session
    - Basics of Data Preparation
    - Examples

- Decision Trees by Example via SAS-VDMML: 1 Session
    - Basics of Decision Trees
    - Examples

- Random Forest by Example via SAS-VDMML: 1 Session
    - Basics of Random Forest
    - Examples

- Logit/Probit Regression by Example via SAS-VDMML: 1 Session
    - Basics of Logit/Progit
    - Examples

- Regularized Linear Regression by Example via SAS-VDMML: 2 Sessions
    - Basics of Shrinkage
    - Examples

PART II: ADVANCED MACHINE LEARNING (8 sessions)
- Replicating and refining previous examples using procedures
- Combing SAS and Python

METHODOLOGY:

One on-line session per week (teams, slides), plus weekly readings and homeworks

READINGS:

Hyndman, Rob (2014): "Measuring Forecast Accuracy", published as Chapter 3 of Gilliland, Tashman and Sglavo (2015): "BUSINESS FORECASTING: Practical Problems and Solutions", Wiley. Link: https://robjhyndman.com/papers/forecast-accuracy.pdf

James, G., Witten, D. ,Hastie, T. and Tibshirani, R. (2017): "An Introduction to Statistical Learning with Applications in R". Springer.
Link: https://hastie.su.domains/ISLR2/ISLRv2_corrected_June_2023.pdf

Kostenko, A. and Hyndman, R. (2008): "Forecasting without significance tests?".
Link: https://www.researchgate.net/publication/222105750_Forecasting_without_significance_tests

SAS (2018): "Visual Data Mining and Machine Learning 8.2: User's Guide".
Link: http://documentation.sas.com/doc/en/vdmmlcdc/8.2/vdmmlug/p0psvjbnx1y57mn1taay7h8d4ve0.htm

Shmueli, Galit (2010): "To Explain or to predict?", Statistical Science 25(3).
Link: https://www.stat.berkeley.edu/~aldous/157/Papers/shmueli.pdf

Varian, Hal (2014): "Big Data: New Tricks for Econometrics", Journal of Economic Perspectives 28(2). Link: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.28.2.3