INDIVIDUAL WORK: Deadline: 28/04/2021

Sent the pdf file with the work AND source codes by email to leandro.coelho@pucpr.br **SUBJECT**: Computational work 01

Computational work 01: Regularization approaches

Leandro dos Santos Coelho

Pontifícia Universidade Católica do Paraná (PUCPR), Escola Politécnica

Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS), Graduação em Engenharia de Controle e Automação (Mecatrônica)

Rua Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR, Brasil

Universidade Federal do Paraná (UFPR), Graduação e Pós-Graduação em Engenharia Elétrica, Campus Centro Politécnico Av. Cel. Francisco H. dos Santos, 100, CEP 81530-000, Curitiba, PR, Brasil

e-mail: leandro.coelho@pucpr.br; lscoelho2009@gmail.com

 $Curr\'iculo\ Lattes: \ \textit{http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id} = K4792095Y4$

Google Scholar: https://scholar.google.com/citations?user=0X7VkC4AAAAJ&hl=pt-PT

Linkedin: https://www.linkedin.com/in/leandro-dos-santos-coelho-07a08893/

Computational work

INDIVIDUAL WORK: Deadline: 28/04/2021

Sent the pdf file with the work AND source codes by email to leandro.coelho@pucpr.br **SUBJECT**: Computational work 01

- 1) Choice three case studies with focus in regression.
- 2) Adapt the linear regression (with OLS), ridge, LASSO and elastic net approaches to the three case studies and after run it (in Matlab, R or Python).
- Analysis (tabels, figures, statistical analysis, box-plot, histograms, train-validation performance, etc) of the results for forecasting (MSE or MAPE or RMSE or R² metrics using **cross-validation**, **5-fold**) to **diferent values** of regularization parameter and without/with a feature selection method.

Dataset (adopt case study with 1 output)

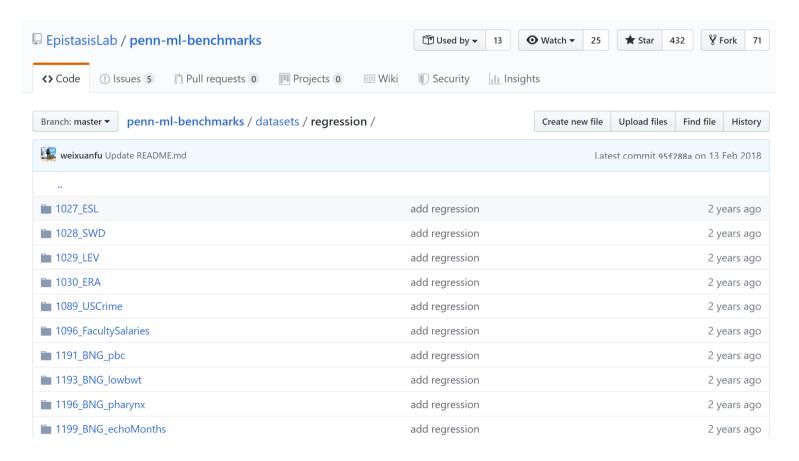


98 Data Sets

<u>Name</u>	<u>Data Types</u>	Default Task	Attribute Types	# Instances	# Attributes	<u>Year</u>
3D Road Network (North Jutland, Denmark)	Sequential, Text	Regression, Clustering	Real	434874	4	2013
Air Quality	Multivariate, Time-Series	Regression	Real	9358	15	2016
Air quality	Multivariate, Time-Series	Regression	Real	9358	15	2016
Airfoil Self-Noise	Multivariate	Regression	Real	1503	6	2014
Amazon Access Samples	Time-Series, Domain- Theory	Regression, Clustering, Causal-Discovery		30000	20000	2011

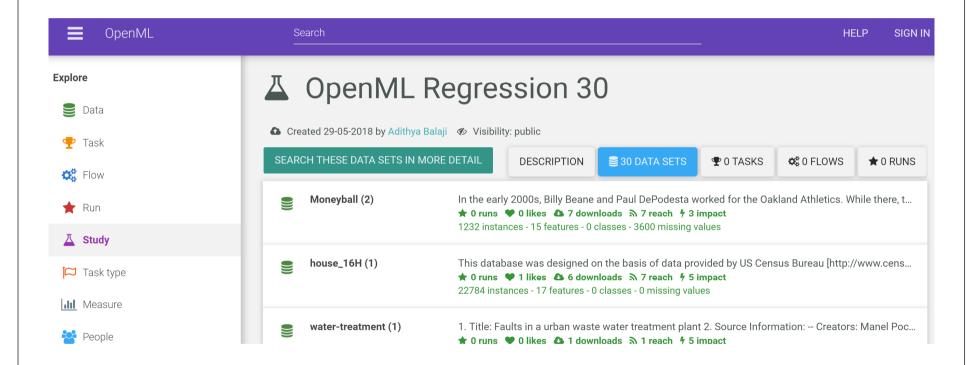
Dataset (adopt case study with 1 output)

Penn Machine Learning Benchmark (PMLB) repository



Dataset (adopt case study with 1 output)

OpenML



Quote

Absence of evidence is **not** evidence of absence.

Carl Sagan (1934-1996)

American astronomer, cosmologist, astrophysicist, astrobiologist.

Sagan published more than 600 scientific papers and articles and was author, co-author or editor of more than 20 books.

Sagan advocated scientific skeptical inquiry and the scientific method, pioneered exobiology and promoted the Search for Extra-Terrestrial Intelligence (SETI).

