Ε Θ Ν Ι Κ Ο Μ Ε Τ Σ Ο Β Ι Ο Π Ο Λ Υ Τ Ε Χ Ν Ε Ι Ο
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ & ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
Τομέας Μαθηματικών
Πολυτεχνειούπολη – Ζωγράφου ΑΘΗΝΑ - 157 80
ΤΗΛ. : 772 1774
FAX : 772 1775

ΠΡΟΣΚΛΗΣΗ

Σας προσκαλούμε στην ομιλία του **Daniel Schmidt, Associate Professor of Computer Science at the Department of Data Science and AI, Monash University, Australia**, η οποία θα πραγματοποιηθεί την Παρασκευή 29 Μαρτίου 2024, ώρα 13:00, στην Αίθουσα Σεμιναρίων του Τομέα Μαθηματικών, Κτήριο Ε, 2ος όροφος, Πολυτεχνειούπολη Ζωγράφου, με θέμα:

Prevalidated ridge regression as a highly-efficient drop-in replacement for logistic regression for high-dimensional data

Abstract/Περίληψη: Linear models are widely used in classification and are particularly effective for high-dimensional data where linear decision boundaries/separating hyperplanes are often effective for separating classes, even for complex data. A recent example of a technique effectively utilising linear classifiers is the ROCKET family of classifiers for time series classification. One reason that the ROCKET family is so fast is due to its use of a linear classifier based around standard squared-error ridge regression. Fitting a linear model based on squared-error is significantly faster and more stable than fitting a standard regularised multinomial logistic regression based on logarithmic-loss (i.e., regularised maximum likelihood), as in the latter case the solutions can only be found via a numerical search. While fast, one drawback of using squared-error ridge-regression is that it is unable to produce probabilistic predictions. I will demonstrate some very recent work on how to use regular ridge-regression to train L2-regularized multinomial logistic regression models for very large numbers of features, including choosing a suitable degree of regularization, with a time complexity that is no greater than single ordinary least-squares fit. This in contrast to logistic regression, which requires a full refit for every value of regularisation parameter considered, and every fold used for cross-validation. Using our new approach allows for models based on linear classifier technology to provide well calibrated probabilistic predictions with minimal additional computational overhead. If time permits, I will also discuss some thoughts on when such linear classifiers would be expected to perform well.

Short Bio: Daniel Schmidt is an Associate Professor of Computer Science at the Department of Data Science and AI, Monash University, Australia. He obtained his PhD in the area of information theoretic statistical inference in 2008 from Monash University. From 2009 to 2018 he was employed at the University of Melbourne, working in the field

of statistical genomics (GWAS, epigenetics and cancer). Since 2018 he has been employed at Monash University in a teaching and research position. His research interests are primarily time series classification and forecasting, particularly at scale, and Bayesian inference and MCMC, with an emphasis on sparsity and shrinkage, Bayesian optimisation and Bayesian function approximation. He also has a keen interest in the best ways to provide statistical/machine learning education.