

LTCC Advanced Course

Title: Kernel methods in machine learning and statistics

Basic Details:

- Core Audience: Second/Third year Stats and Applied
- Course Format: Extended; 5 x 2hr lectures

Course Description: This course will cover the mathematical foundations of reproducing kernel Hilbert spaces (RKHSs) and Gaussian processes, emphasising the connections and equivalences between those and with a view towards applications in machine learning and statistics. Regression will serve as a running example and will be discussed from the RKHS, Gaussian process, kernel trick and deep learning perspective. The module will conclude with an overview of kernel-based discrepancies between probability measures.

- Keywords: reproducing kernel Hilbert spaces, Gaussian processes, neural tangent kernel, maximum mean discrepancy, kernelised Stein discrepancy
- Syllabus:
 - Positive definite kernels and reproducing kernel Hilbert spaces: construction, basic properties and examples
 - Kernel ridge regression (representer theorem, generalisation)
 - The kernel trick (kernel PCA, support vector machines)
 - Gaussian processes (GP regression, Bayesian optimisation)
 - The neural tangent kernel regime in the training of neural networks
 - divergences based on kernels: maximum mean discrepancy and kernelised Stein discrepancy
- Recommended reading:
 - Kanagawa, K., Hennig, P., Sejdinovic, D., Sriperumbudur, B. K.: *Gaussian processes and kernel methods: A review on connections and equivalences*. CoRR abs/1807.02582 (2018)
 - Schölkopf, B., Smola, A. J. (2018). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. Adaptive Computation and Machine Learning series (2018)
- Additional or optional reading:
 - Fasshauer, G. and McCourt, M.: *Kernel-based Approximation Methods using MATLAB*. Interdisciplinary Mathematical Sciences (2015)
 - Rasmussen, C. E. and Williams, C.K.I.: *Gaussian processes for machine learning*. I The MIT Press (2006)
 - Jacot, A., Gabriel, F., Hongler, C: *Neural Tangent Kernel: Convergence and Generalization in Neural Networks*. Advances in Neural Information Processing Systems (2018)

- Muandet, K., Fukumizu, K., Sriperumbudur, B., Schölkopf, B.: *Kernel Mean Embedding of Distributions: A Review and Beyond*. Found. Trends Mach. Learn.10(1-2): 1-141 (2017)
- Prerequisites: Familiarity or some exposure to basic functional analysis, probability and statistics will be useful

Format:

- Electronic lecture notes will be provided
- 10 hours of lectures

Lecturer Details:

- Lecturer: Dr Nikolas Nüsken
- Lecturer home institution: King's College London
- Lecturer e-mail: nikolas.nusken@kcl.ac.uk