

**PiAI Seminar Series: Physics informed AI in Plasma Science**  
**10:00-11:00, 12 June 2023 (CEST)**  
**17:00-18:00, 12 June 2023 (JST)**  
**Web Seminar**

**Nonlinear Statistical Modeling for High-Dimensional and Small Sample  
Data**

**Makoto Yamada**

*Okinawa Institute of Science and Technology Graduate University (OIST), Japan*

Feature selection/variable selection is an important machine learning problem, and it is widely used for various types of applications such as gene selection from microarray data, document categorization, and prosthesis control, to name a few. The feature selection problem is a fundamental and traditional machine learning problem, and thus there exist many methods, including the least absolute shrinkage and selection operator (Lasso). The Lasso method is highly used in various applications and widely studied in machine learning, statistics, and biology communities; However, it cannot handle nonlinearly related data. In this talk, I first introduce a simple yet effective nonlinear feature selection method called HSIC Lasso [1,2,3,4]. Then, I further present the selective inference variant of HSIC Lasso [5]. Finally, I introduce the applications of the proposed nonlinear feature selection methods.

Reference:

- [1] Makoto Yamada, Wittawat Jitkrittum, Leonid Sigal, Eric P. Xing, Masashi Sugiyama: High-Dimensional Feature Selection by Feature-Wise Kernelized Lasso. *Neural Comput.* 26(1): 185-207 (2014)
- [2] Héctor Climente-González, Chloé-Agathe Azencott, Samuel Kaski, Makoto Yamada: Block HSIC Lasso: model-free biomarker detection for ultra-high dimensional data. *Bioinform.* 35(14): i427-i435 (2019)
- [3] Benjamin Poignard, Makoto Yamada: Sparse Hilbert-Schmidt Independence Criterion Regression. *AISTATS 2020*: 538-548
- [4] Makoto Yamada, Yuta Umezu, Kenji Fukumizu, Ichiro Takeuchi: Post Selection Inference with Kernels. *AISTATS 2018*: 152-160
- [5] Tobias Freidling, Benjamin Poignard, Héctor Climente-González, Makoto Yamada: Post-selection inference with HSIC-Lasso. *ICML 2021*: 3439-3448