Machine Learning under Distribution Shifts

Masashi Sugiyama

RIKEN and The University of Tokyo, Japan.

https://www.ms.k.u-tokyo.ac.jp/sugi/ sugi@k.u-tokyo.ac.jp

A common assumption in standard machine learning methods is that the data used for training a predictor follow the same probability distribution as the data used for testing the prediction performance in the inference phase. However, in many real-world applications, this common assumption is often violated, e.g., due to changing environments over time or sample selection bias caused by privacy concerns. Such a situation is called distribution shift, and how to overcome the distribution shift is an urgent challenge in the machine learning community.

In this talk, I will first give an overview of the classical importance weighting approach to distribution shift adaptation, which consists of an importance estimation step and an importance-weighted training step [1,2]. Then, I will present a more recent approach that simultaneously estimates the importance weight and trains a predictor. I will also discuss a more practical scenario of sequential distribution shifts, where the data distributions change sequentially over time. Finally, I will discuss ongoing challenges such as joint distribution shift, out-of-distribution adaptation, and more.

[1] Sugiyama, M. & Kawanabe, M. Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation, MIT Press, Cambridge, Massachusetts, USA, 2012.

[2] Quiñonero-Candela, J., Sugiyama, M., Schwaighofer, A., & Lawrence, N. D. (Eds.), Dataset Shift in Machine Learning, MIT Press, Cambridge, Massachusetts, USA, 2009.