ID de Contribution: **173** Type: **Invited talk**

Score Learning under the Manifold Hypothesis: Theory and Implications for Data Science

jeudi 31 juillet 2025 10:45 (30 minutes)

The score function plays a central role in modern generative modeling, particularly in diffusion models and related score-based methods. Despite its theoretical appeal, learning the score function in practice is notoriously difficult: it is sensitive to hyperparameter choices and prone to various forms of instability that often require ad hoc corrections.

In this talk, we study score learning under the popular **manifold hypothesis**, which posits that high-dimensional data concentrates near a low-dimensional manifold. A key and perhaps surprising insight is that, under this hypothesis, it is significantly easier to learn the underlying manifold structure than the full data distribution. We provide theoretical support for this claim and explore its consequences for score estimation and generative modeling. In particular, we show how this perspective can clarify challenges in training diffusion models and guide the development of more stable and efficient algorithms.

Author: HSIEH, Ya-Ping (ETH Zürich)

Orateur: HSIEH, Ya-Ping (ETH Zürich)

Classification de Session: ML

Classification de thématique: Machine learning