

Explainable Machine Learning

Thursday 2 October 2025 10:00 (1h 30m)

Explainable machine learning involves two complementary types of questions: (1) Method designers and theorists ask: “Why do deep neural networks work so well? How can we further improve them? How can we provide formal performance guarantees?” (2) Instead, method users ask, “How did the network arrive at its conclusion? What variables determine the result, and in what way? Can the result be trusted?” The talk will introduce both areas and review the current state of our answers.

Presenter: KOETHE, Ullrich (Heidelberg University)