

Luca Saglietti:

Investigating the limits of active learning in the perceptron model

In a large class of machine learning problems unlabeled data is abundant yet obtaining labels can be expensive.

An active learning algorithm can be employed for selecting a limited number of samples and querying the corresponding labels, subsequently used for supervised learning. In this talk I will frame this problem in the case of the teacher-student perceptron model, where a large deviation analysis provides bounds for the best (and worst) possible active learning performance.

I will then show how, in our simple setting, AMP can be employed for efficiently approaching the obtained bounds.