



Contribution ID: 296

Type: **Talk**

Robustness and over-parameterization

Monday 3 November 2025 18:00 (20 minutes)

In many modern machine learning applications, models are often trained to near zero “training loss” (in other words, to interpolate the training data), while also having far more training parameters than the “number of data points”. This appears to violate traditional rules-of-thumb for avoiding overfitting, and considerable work has thus been devoted to gain a better understanding of such over-parameterization. A more recent development is an interesting direction suggested by Bubeck and Sellke, who postulated that, in certain settings, a much larger number of parameters may indeed be required to interpolate the training data if the model being trained is also required to be “robust”: small modifications to the model input should not lead to very large changes in the model output.

In this talk, we will survey this formulation of the connections between robustness and over-parameterization. We will also present the conceptual view that a bias-variance type decomposition for the loss function lies at the heart of the results of Bubeck and Sellke, and then use this idea to show that the setting of losses corresponding to Bregman divergences is the natural setting for understanding the connection between robustness and over-parameterization in this formulation.

Parallel Session (for talks only)

Algorithms and artificial intelligence

Authors: BATRA, Jatin (TIFR Mumbai); SRIVASTAVA, Piyush (Tata Institute of Fundamental Research); DAS, Santanu (TIFR Mumbai)

Presenter: SRIVASTAVA, Piyush (Tata Institute of Fundamental Research)

Session Classification: Algorithms and artificial intelligence