

Stephane d'Ascoli

Title: Reconciling the double descent curve with older ideas

Abstract: A recent line of research has highlighted the existence of a double descent curve in deep learning, whereby the generalisation error of neural networks decreases, peaks, then decreases again as we increase the number of parameters or number of training examples.

In this talk, we shed light on this behaviour by reconciling it with two older ideas :

(i) The bias-variance decomposition, according to which the generalisation error should instead follow a U-curve ;

(ii) The double descent observed in linear models, which corresponds to a different kind of overfitting. In nonlinear models, the latter is implicitly regularized, but can still be observed under certain conditions.

Joint work with Maria Refinetti, Levent Sagun, Giulio Biroli, Florent Krzakala.