



Contribution ID: 103

Type: **not specified**

Learning Dynamics of Neural Networks

Nowadays, the increasing amount of data requires more and more sophisticated tools that are capable of processing it, without the need of human assistance. Fields such as computer vision, speech recognition, and self-driving cars rely on artificial neural networks and their interconnected structure to find common patterns in data. This is due to their structure which is inherently adaptable in the aim of a specific goal given by the creator of the network. However, there still lacks a solid theoretical foundation to explain many observed behaviours, such as the double descent effect in large networks, the effectiveness of stochastic gradient descent for nonlinear problems, and the link between overparameterization and overfitting. Addressing these knowledge gaps is challenging but important. As a way to better understand how these tools learn, statistical treatments will be applied, as large neural networks can be viewed as dynamically interacting degrees of freedom. The setup will consist of the study of the Hessian matrix of the loss function of the student network in the teacher-student framework, which allows for an intimate control of the learning process, through easy accessibility to the dataset and the student network. Moreover, it is expected that chaotic behaviour will be observed, which will be quantified through various metrics, such as Lyapunov Exponents, Kolmogorov Entropy and Attractor Dimension.

Author: MOREIRA, Miguel

Presenter: MOREIRA, Miguel