Ph.D. in Information Technology: Thesis Defense

February 12th, 2021

online by Teams – at 10.00

Matteo SANGIORGIO – XXXIII Cycle

Deep learning in multi-step forecasting of chaotic dynamics

Supervisor: Prof. **Giorgio Guariso**

Abstract:

Chaotic systems are difficult to predict due to the well-known sensitivity to initial conditions: an infinitesimal uncertainty on the state leads to a substantially different evolution. In the last few decades, many attempts have been done to discover how far into the future such systems can be predicted, adopting a wide range of models. The topic became more and more debated in recent years, due to the development of many machine learning techniques for time series forecasting.
We focus on LSTM (long short-term memory) neural networks, which represent the state-of-the-art approach in many sequential tasks (e.g., natural language processing). LSTM nets are traditionally trained making use of the so-called "teacher forcing", i.e. the ground truth data are used as input for each time step ahead. This does not allow to correct small errors because, during the training phase, the prediction at a particular time step does not affect future predictions. Conversely, we propose to feed the previous predictions back into the recurrent neurons, as it happens when the network is used in forecasting, also during the training phase.
LSTM predictors are first tested on noise-free data generated by canonical chaotic oscillators. We then extend the analysis to artificially-generated noisy data and slow-fast analytical systems. Lastly, after assessing their degree of chaoticity, solar irradiance and ozone concentration time series are examined to assess the forecasting accuracy of the neural predictors on real-world cases.
The results demonstrate that LSTM predictors outperform the feed-forward benchmarks in all the considered tasks. In particular, they may still show high (R2-score greater than 0.8) accuracy after 6-7 Lyapunov times when forecasting the dynamics of famous chaotic attractors.
The generalization capability of the neural architectures in terms of domain adaptation is also analyzed. We show that the predictors trained on data sampled in a specific location (source domain) maintain analogue performances when used, without retraining, in different locations (target domains).

PhD Committee
Prof. **Carlo Piccardi**, DEIB
Prof. **Mattia Frasca**, Universita' di Catania
Prof**. Rene' Alquezar**, Universitat Politecnica de Catalunya