Training Methods for LSTM Networks for Complex High-Dimensional Dynamics

Investigating temporal embeddings and training methods for Long-Short Term Memory (LSTM) Recurrent Neural Networks.

Recurrent Neural Networks (RNNs) contain states with self-loops enabling them to capture complex temporal dependencies. A class of RNNs namely the LSTM experienced a widespread success in sequence modelling, e.g. natural language processing, speech recognition, etc. However, successful applications of RNNs in high dimensional systems are scarce. Capturing the complex dynamics and different time scales in these systems is challenging. Two main questions need to be answered. The first one is how to discretise the high dimensional time series data (embedding space) to guarantee absence of correlations in the input of the LSTM that leads to suboptimal performance. The second question is what is the optimal way to train a LSTM in practice.

The student will learn the theoretical concepts behind RNNs and implement an LSTM architecture that efficiently copes with the challenges. The student will also gain experience in machine learning software (tensorflow) and Python.

This project can be adapted to BSc/MSc/Project.

PREREQUISITES
- Basic programming skills
- C/C++/Python
- Ability to work Independently
- Basic knowledge of ML

CONTACT
Prof. Petros Koumoutsakos
Dr. Georgios Arampatzis
MSc. Pantelis Vlachas
e-mail: petros@ethz.ch
garampat@ethz.ch
pvlachas@ethz.ch

In the CSE Lab, we combine computational methods, computer science tools and domain specific knowledge to solve scientific and engineering problems in areas such as Fluid Mechanics, Nanotechnology and Life Sciences. The core computational competences of our group are in particle methods and in stochastic optimization techniques. Motivated by challenges in application fields, we focus on identifying the common elements among computational techniques and on formulating common methodological, algorithmic and software structures that facilitate their further development.