Karlsruhe Institute of Technology

Autonome Lernende Roboter (ALR)

Prof. Gerhard Neumann

Project Type

- ☐ Master Thesis
- ☑ Bachelor Thesis
- ☑ Praktikum
- ☑ Seminar

Supervisors

Saleh GHOLAM ZADEH

@ salehgholamzadeh17@gmail.com

vaisakh Shaj

Difficulty

Algorithmic

Math

Application

Long range time-series forecasting with sequence to sequence models

Description

A longstanding challenge in the field of machine learning is efficiently modelling sequential data longer than a few thousand-time steps. The usual paradigms for designing sequence models involve recurrence (e.g., RNNs) or convolutions (e.g., CNNs), each of which come with tradeoffs. For example, RNNs suffer from a "vanishing gradient" which empirically limits their ability to handle long sequences. CNNs encode local context and enjoy fast, parallelizable training, but are not sequential, resulting in more expensive inference and an inherent limitation on the context length.

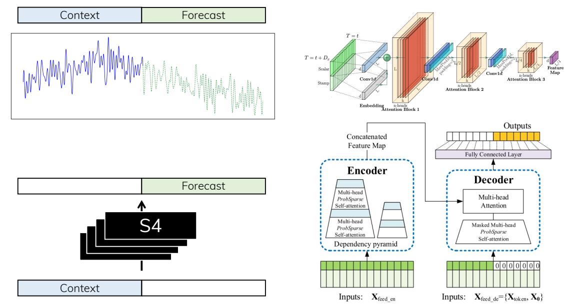

Figure 1: Some recent methods for sequence modeling.

The recently proposed Structured State Space for Sequence Modeling (S4) [1] architecture showed that simple linear models can capture very long-range dependencies over tens of thousands of steps. This widely acclaimed work was able to outperform state-of-the-art models including Transformers [3] on the challenging Long Range Arena benchmark. However, this model is not specifically designed for long-term forward prediction

On the other hand, a recent work MTS3 [2], introduces a probabilistic approach for learning multi-time scale sequential models. This model demonstrates superior performance by enabling predictions at multiple levels of temporal granularity/abstractions.

In this project, we want to benchmark these methods on several datasets in order to understand their specific strengths. This will also give the student hands-on experience on the latest deep learning and deep probabilistic methods for solving sequence modelling tasks.

Tasks

Depending on the scope of the project the tasks in this project will involve:

- Review state-of-the-art: Literature review on the latest time-series prediction methods, namely S4, MTS3 and Transformers
- Applying the code: Understand, adapt where necessary, and apply the above algorithms on the given time-series datasets
- Evaluation: Evaluate the performance across different settings
- Report and presentation: Deliver the final report and present the results
- Documentation: Document and deliver the code

References

[1] Albert Gu, Karan Goel, and Christopher Re. Efficiently modeling long sequences with structured state spaces. In *International Conference on Learning Representations*, 2021.

[2] Vaisakh Shaj, Saleh Gholam Zadeh, Demir Ozan, Louis Riccardo, and Gerhard Neumann. Multi time scale state space models. *Publication Under Review*, 2023.

[3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.