Contribution ID: **93**   Type: **Tutorial**

Transformer Tutorial

Tuesday, March 5, 2024 9:00 AM (1h 30m)

The Transformer is a deep learning architecture introduced in 2017, that has since then taken over the natural language processing field and has recently gained public popularity thanks to large language models like ChatGPT. The self-attention mechanism introduced with the Transformer allows it to learn complex patterns and relationships in data without explicitly using recurrent mechanisms like classic RNN-style architectures. While the Transformer was developed for sequence-to-sequence language modeling like translation tasks, the usefulness for time series prediction has been less explored in the machine learning community. Particularly, the lack of beginner-friendly tutorials and guides for using transformers with uni- and multivariate continuous input and outputs are not easily found online, as opposed to for natural language tasks. Therefore, this tutorial aims to introduce the Transformer architecture and how to use standard deep-learning library Transformer building blocks to construct a simple time series prediction model and explain the inputs and outputs of the transformer model. As an appendix, we will give a quick outlook of current state-of-the-art time series prediction architectures based on the basic Transformer as well as alternative modern time series forecasting methods.

Primary Keyword

transformers

Secondary Keyword

Tertiary Keyword

Primary author: LU, Anton (Technical University of Vienna (AT))

Presenter: LU, Anton (Technical University of Vienna (AT))

Session Classification: Tutorials

Track Classification: Methods