Title of project:

Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting

Goals:

- Discover interplay between classical approaches of time series prediction and modern deep learning techniques.

- Reproduce the model in paper "Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting" by Bryan Lim, Sercan Arik, Nicolas Loeff and Tomas Pfister.

- Test the model on generated data and compare my results on datasets used by the authors in the paper.

Justification of the proposed project:

The focus of this project is to better understand how we can deal with high-dimensional time series with multiple inputs, missing values, and irregular timestamps. In this situation, the performance of the classical approaches is not satisfying, and naïve applications of deep learning also fail. Even the deep learning models that have shown promising results tend to be "black boxes" with little insight into how to interpret the results.

Lim et al. propose a modern attention-based architecture to combine multi-horizon forecasting functionality with interpretable insights related to temporal dynamics.

Classical time series approaches have generally performed better on one-step-ahead predictions. Still, decision-making in many real situations require forecasting to be done multiple steps into the future. For example, monthly revenue forecast for the next semester in order to plan hiring, inventory, production, etc.

Accurate forecasts of this type usually require access to multiple inputs including invariant or static features (e.g. country), information about the future (e.g. weather forecasts), or parallel and exogenous time series (e.g. website traffic).

The attention-based architecture proposed by the authors of the analyzed paper has been designed specifically to leverage prior knowledge with suitable inductive biases for multi-horizon forecasting. More specifically, the Temporal Fusion Transformer includes the following novel components:

- Static covariate encoders that produce context vectors

- Gating mechanisms and variable selection to reduce the impact of irrelevant inputs.

- Local processing of known and observed inputs through a sequence-to-sequence layer.

- Long-term dependency learning through a temporal self-attention decoder.

Interpretation of the results is centered around the identification of globally important variables, persistent temporal patterns and significant events.

The main objective of this project is to reproduce the architecture proposed by the authors and evaluate its performance on generated toy datasets, as well as the datasets used by the authors in the original paper.

1. Bryan Lim, Sercan Arik, Nicolas Loeff and Tomas Pfister. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. https://arxiv.org/pdf/1912.09363.pdf

2. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, I. Polosukhin, Attention is all you need, in: NIPS, 2017.

3. S. S. Rangapuram, et al., Deep state space models for time series forecasting, in: NIPS, 2018.

4. S. Li, et al., Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting, in: NeurIPS, 2019.