ENHANCING TRANSFORMER-BASED ARCHITECTURES WITH TEMPORAL HIERARCHIES FOR TIME-SERIES FORECASTING

Emma Bachyrycz, Systems Science, MS Candidate

Time 11:00 am - 12:30 pm on 4/26/2024

Room: ITC WISE Room, BI 1610

Abstract

This research proposes an enhanced transformer model for time-series forecasting to align the various levels of decision-making. Transformer models were initially proposed in 2017 for natural language processing and have gained popularity since, notably with the release of ChatGPT. This architecture is unique with its parallel multi-head attention mechanism that addresses the concerns of the popular deep learning model recurrent neural networks. Transformer mechanisms can handle large datasets without experiencing the vanishing gradient problem, aligning with the rise of the Industry 4.0 Revolution and Big Data. In this research we explore the opportunities of transformer models in time-series forecasting, offering enhancements using temporal hierarchies. This involves aggregating the data at different temporal scales, such as daily, monthly, and yearly, and incorporating the hierarchical levels into the model's framework. By training the model on recurrent structures at different frequencies, our preliminary results encourage accurate, dynamic, and robust forecasting of the dataset that outstands alternatives. The various temporal hierarchies represent the levels of decision-making within an organization. The monthly dataset aligns with operational decisions frequently made by lower management. The quarterly dataset is the tactical level, which incorporates the medium-term goals of a company and is made by middle management. Finally, the yearly dataset is for the strategic level decisions made by executives. The results show the model combining the monthly, quarterly, and yearly temporal hierarchies performed with the most accuracy, excelling above the benchmark models. With three attention heads and a forecasting horizon of two years, the model performed with an MSE of 1.741, sMAPE of 0.296%, and RMSE of 1.319. Under the same forecasting parameters, the second-highest-performing model is the forecast combining operational and tactical levels or monthly and quarterly data. This model performed with an MSE of 1.885, sMAPE of 0.329%, and RSME of 1.373