Azul Garza

Nixtla

The Scaling Laws of Time (or The Unexpected Virtue of Diversity)

Recent research into the scaling laws of large language models has demonstrated how performance improves with increased data volume and model complexity. This paper explores these scaling laws in the context of time series forecasting, analyzing more than 100 billion data points using transformer-based architectures. We define diversity in time series data as richness in frequencies, patterns, and use cases, spanning domains such as energy, finance, and retail. Our findings reveal that transformer-based models effectively extrapolate scaling laws when trained on such diverse datasets, achieving significant improvements in forecasting accuracy. Conversely, we observe diminishing returns when data diversity is limited, highlighting its critical role in leveraging scaling for generalization. These results provide key insights into the design of foundational time series forecasting models and emphasize the importance of dataset diversity for optimal performance in heterogeneous environments.