THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

Master's Thesis Presentation

Rahul Kukreja
Department of Statistics
The University of Chicago

"Interpreting Transformers for Time-Series Forecasting"

November 13, 2025, at 9:30 AM
Jones 111, 5747 S. Ellis Avenue

Abstract

Transformer models have often demonstrated improved performance across time-series forecasting tasks. Models like AutoFormer, Informer, and PatchTST have shown to outperform other non-transformer baselines on real-world datasets. However, their internal interpretability remains underexplored. In this work, we evaluate whether time-series transformers encode meaningful structure aligned with the underlying data-generating processes. Using synthetic datasets generated from known linear and non-linear models—including ARMA, SARIMA, STAR, FAR, and bilinear processes—we investigate internal representations using probing classifiers and clustering analyses. We also introduce a non-attention (MLP-based) baseline to contrast the inductive biases of attention-based architectures.

Our findings show that while both transformers and MLP models achieve near-perfect accuracy on simple series, transformers demonstrate slightly superior representation quality in more complex non-linear settings, as evidenced by layer-wise probe accuracy and cluster purity.

Importantly, the probing results suggest that structural information becomes increasingly disentangled in deeper layers. Overall, our analysis suggests that transformers, and even to some extent non-attention models can encode generative distinctions across multiple independent time-series.

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@uchicago.edu. If you wish to subscribe to our email list, please visit the following website: https://lists.uchicago.edu/web/info/statseminars.