Integrasi Seasonal-Trend Decomposition Dengan Cross-Variable Attention Pada Arsitektur Transformer
Untuk Peramalan Multivariate Time Series
Nasra Pratama Putra, Prof. Dr. Suprapto, M.I.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

INTEGRASI *SEASONAL-TREND DECOMPOSITION* DENGAN *CROSS-VARIABLE ATTENTION* PADA ARSITEKTUR *TRANSFORMER* UNTUK PERAMALAN *MULTIVARIATE TIME SERIES*

oleh

Nasra Pratama Putra
21/484563/SPA/00796

Transformer berbasis dekomposisi telah banyak digunakan dalam peramalan *multivariate time series* karena kemampuannya dalam memodelkan dependensi jangka panjang. Namun, dua persoalan mendasar masih dihadapi oleh pendekatan ini. Pertama, dependensi temporal dalam mekanisme *attention* membentuk pola perhatian yang berorientasi sekuensial sepanjang dimensi waktu, sehingga berpotensi menyebabkan penurunan performa peramalan, terutama pada horizon jangka panjang. Kedua, penggunaan arsitektur *encoder–decoder* menuntut pengulangan komputasi *attention*, yang berdampak pada meningkatnya kebutuhan memori dan waktu komputasi selama proses pelatihan maupun inferensi.

Untuk mengatasi permasalahan tersebut, pada penelitian ini diusulkan sebuah model *Transformer* berbasis *single encoder* yang mengintegrasikan dekomposisi *seasonal–trend* berbasis *moving average* sebagai mekanisme ekstraksi fitur, kemudian representasi hasil dekomposisi diproses melalui mekanisme *cross-variable attention*. Melalui integrasi tersebut, interaksi antarvariabel dapat dimodelkan secara langsung tanpa bergantung pada pemodelan berulang dalam skema *encoder–decoder*, sehingga beban komputasi dapat dikurangi sekaligus representasi dependensi global dapat diperkaya.

Model usulan kemudian dievaluasi pada delapan *dataset benchmark*. Hasil pengujian menunjukkan bahwa model usulan mampu mengungguli *iTransformer* pada enam dari delapan *dataset benchmark*, memberikan performa yang sebanding pada *dataset* Weather, dan hanya sedikit lebih rendah pada *dataset* Traffic. Selain itu, efisiensi komputasi yang sebanding dengan *iTransformer* tetap dapat dipertahankan, dengan perbedaan jumlah parameter kurang dari 2%, konsumsi memori GPU kurang dari 1%, dan waktu komputasi per iterasi kurang dari 3%. Pada saat yang sama, jumlah parameter berhasil dikurangi hingga 97% dibandingkan *Crossformer*, sedangkan waktu komputasi berhasil dipercepat hingga 89.1% dibandingkan *FEDformer*. Lebih lanjut, hasil uji *paired t-test* dua arah pada tingkat signifikansi $\alpha = 0.05$ menunjukkan bahwa perbedaan performa antara model usulan dan *iTransformer* signifikan secara statistik.

Kata kunci: *Transformer, Encoder, Dekomposisi, Moving Average, Attention.*

UNIVERSITAS
GADJAH MADA
**Integrasi Seasonal-Trend Decomposition Dengan Cross-Variable Attention Pada Arsitektur Transformer
Untuk Peramalan Multivariate Time Series**
Nasra Pratama Putra, Prof. Dr. Suprapto, M.I.Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

INTEGRATED SEASONAL–TREND DECOMPOSITION AND CROSS-VARIABLE ATTENTION IN TRANSFORMER ARCHITECTURE FOR MULTIVARIATE TIME SERIES FORECASTING

by

Nasra Pratama Putra
21/484563/SPA/00796

Decomposition-based Transformers have been widely adopted for multivariate time series forecasting due to their ability to capture long-term dependencies. However, two fundamental challenges remain. First, temporal dependencies within the attention mechanism produce sequential attention patterns along the time dimension, potentially leading to degraded forecasting performance, particularly for long-term forecasting horizons. Second, the encoder–decoder architecture requires repeated attention computations, resulting in increased memory consumption and computational time during both training and inference.

To address these limitations, a single-encoder Transformer integrating moving average-based seasonal–trend decomposition for feature extraction and cross-variable attention for representation learning is proposed. The decomposed representations are processed through the cross-variable attention mechanism, enabling inter-variable interactions to be modeled directly without repeated dependency modeling in an encoder–decoder framework. Consequently, computational overhead is reduced while the representation of global dependencies is enhanced.

The proposed model was evaluated on eight benchmark datasets. The experimental results demonstrate that the proposed model outperformed iTransformer on six of the eight benchmark datasets, achieved comparable performance on the Weather dataset, and performed only slightly worse on the Traffic dataset. Furthermore, computational efficiency comparable to that of iTransformer was maintained, with less than a 2% difference in the number of parameters, less than a 1% difference in GPU memory consumption, and less than a 3% difference in computation time per iteration. At the same time, the number of parameters was reduced by up to 97% compared with Crossformer, while computation time was accelerated by up to 89.1% compared with FEDformer. Moreover, the results of a two-tailed paired t-test at a significance level of $\alpha = 0.05$ indicated that the performance difference between the proposed model and iTransformer was statistically significant.

Keywords: *Transformer, Encoder, Decomposition, Moving Average, Attention*.