**Foundations of Process/Product Analytics and Machine learning**
**FOPAM 2023, University of California, Davis CA, July 30-Aug 3, 2023**

# MODELING COMPLEX COMBUSTION WITH MODIFIED TRANSFORMER NEURAL NETWORKS

Ethan Gallup, Jacob Tuttle, Blake Billings, Jacob Immonen, and Kody M. Powell*
University of Utah Department of Chemical Engineering
50 Central Campus Dr, Salt Lake City UT 84112

*Keywords: Transformer neural networks, Dynamic optimization, Model Predictive Control*

The growing popularity of renewable energy sources is adding to the variability of the already fluctuating load on the grid. This makes load following in coal-fired power plants increasingly difficult and opens opportunities to improve dynamic control systems.

Recent research has shown that transformer neural networks used as multi-step ahead time-series prediction models significantly enhance both control performance and computation efficiency of model predictive control (MPC) algorithms when compared with long-short-term memory (LSTM) neural networks (Park 2022). Popular transformer models lack the ability to efficiently correlate variables individually. This project designs a transformer with modified positional embedding layers from vision models to increase understanding of the relationships between variables (Ramachandran et al. 2019).

## Transformer Neural Networks

The transformer network used requires the data to be structured as shown in Figure 1.

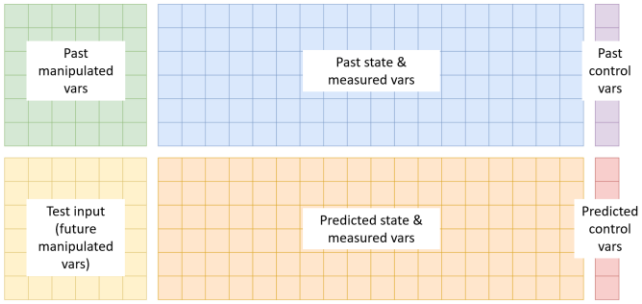


*Figure 1. Data structure used by the transformer model.*

The encoder of the transformer takes the past variables as input and passes a hidden state to the decoder which takes the test inputs and predicts control variables.

## Modifications for Combustion Modeling

The generic transformer takes its own outputs as an input to the decoder. The new model takes manipulated variables as input to the decoder allowing it to draw stronger correlations between them and the control variable. The structure and inputs of the model are in Figure 2. The attention mechanism has a global positional embedding layer based on visual attention models.

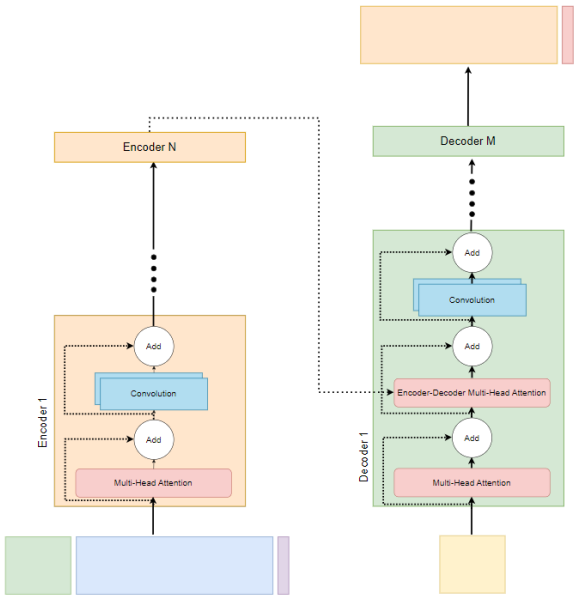


*Figure 2. general transformer model with manipulated variables as the decoder inputs.*

## Results

The new model was tested against GRU, LSTM, bidirectional LSTM, and previous transformer models. The newly developed model showed at least a 15% increase in computational efficiency over all recurrent models and a 5% increase in response to manipulated variables over the previous transformer model.

## References

Park, J. (2022). Hybrid Machine Learning and Physics-Based Modeling Approaches for Process Control and Optimization (dissertation).

Ramachandran P., Parmar N., Vaswani A., Bello I., Levskaya A, Shlens J. (2019). Stand-Alone Self-Attention in Vision Models, *arXiv*.

* kody.powell@ utah.edu.