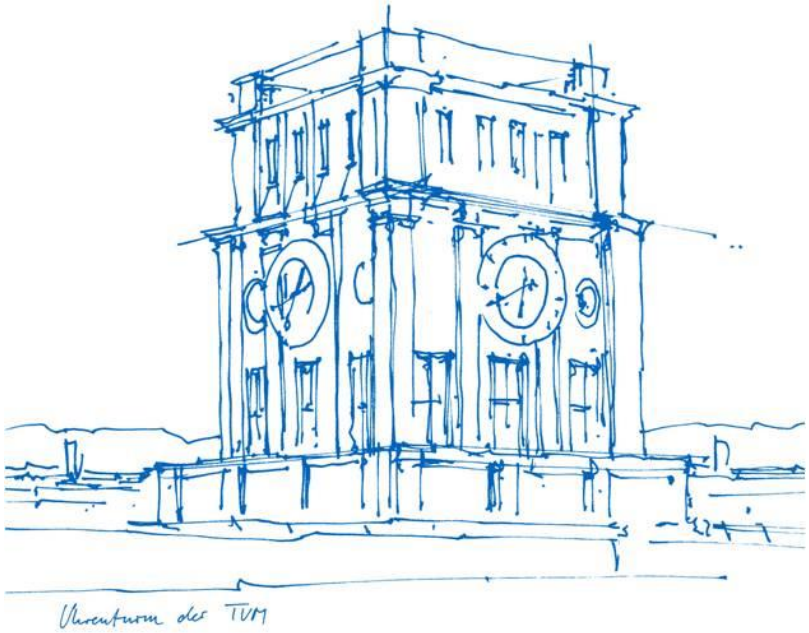
Third Update – Connectome Informed Attention

Andres Zapata | Mohamed Said Derbel | Niklas Bühler

Munich, November 2022



Third Update – Connectome Informed Attention

Overview

1. Recap
   - Transformer for Forecasting

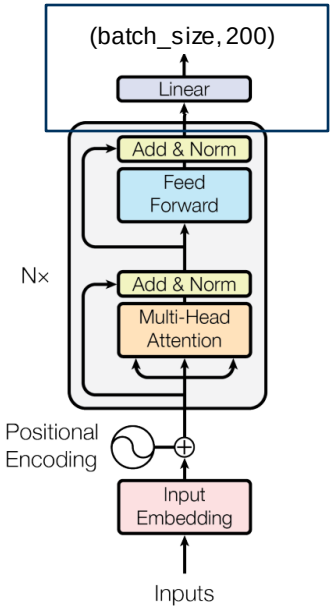
2. Progress and Findings
   - Updated Transformer Encoder
   - Connectivity Data
   - Different incorporation methods
   - Long Range Spatiotemporal Transformer

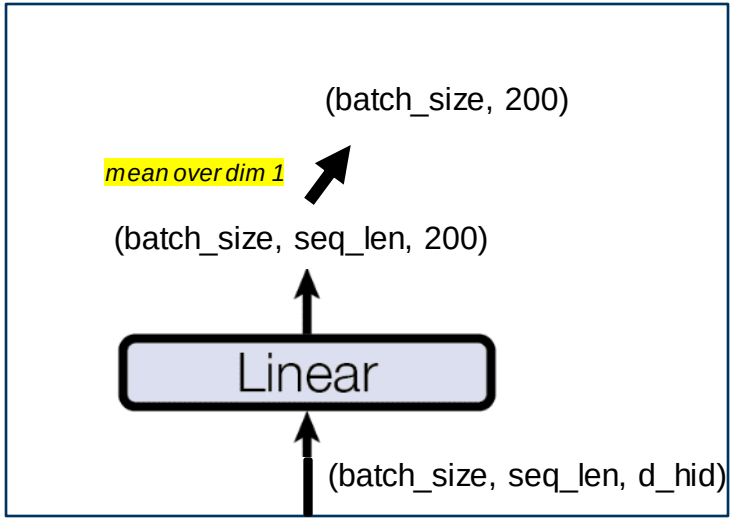
3. Next steps
   - Connectivity Informed Attention
   - Long-Range Spatiotemporal Transformer

1. Recap

- **Transformer for Forecasting**

Transformer Model: Basic Architecture



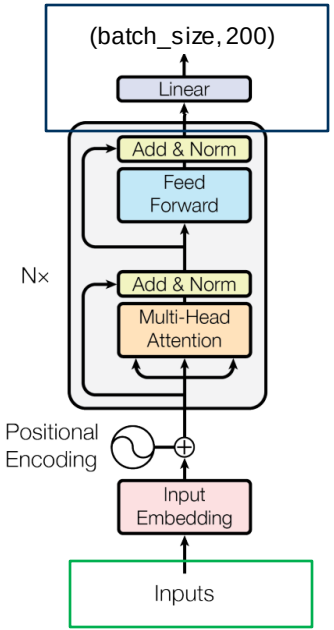


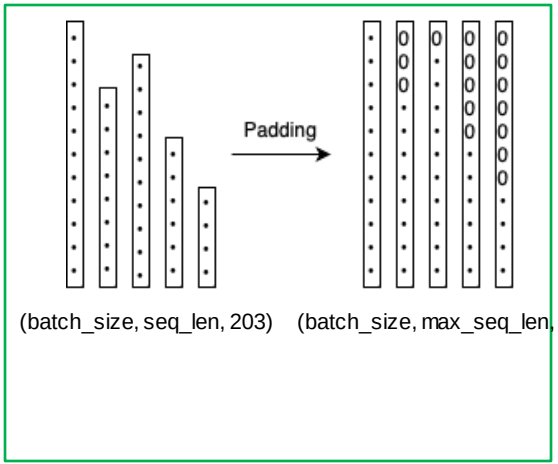
Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

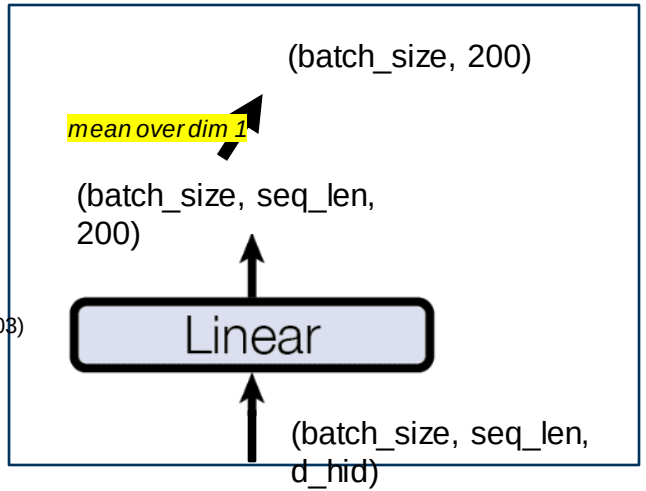
2. Progress and Findings

- **Updated Transformer Encoder**
- Connectivity Data
- Different incorporation methods
- Long Range Spatiotemporal Transformer

Updated Transformer Architecture





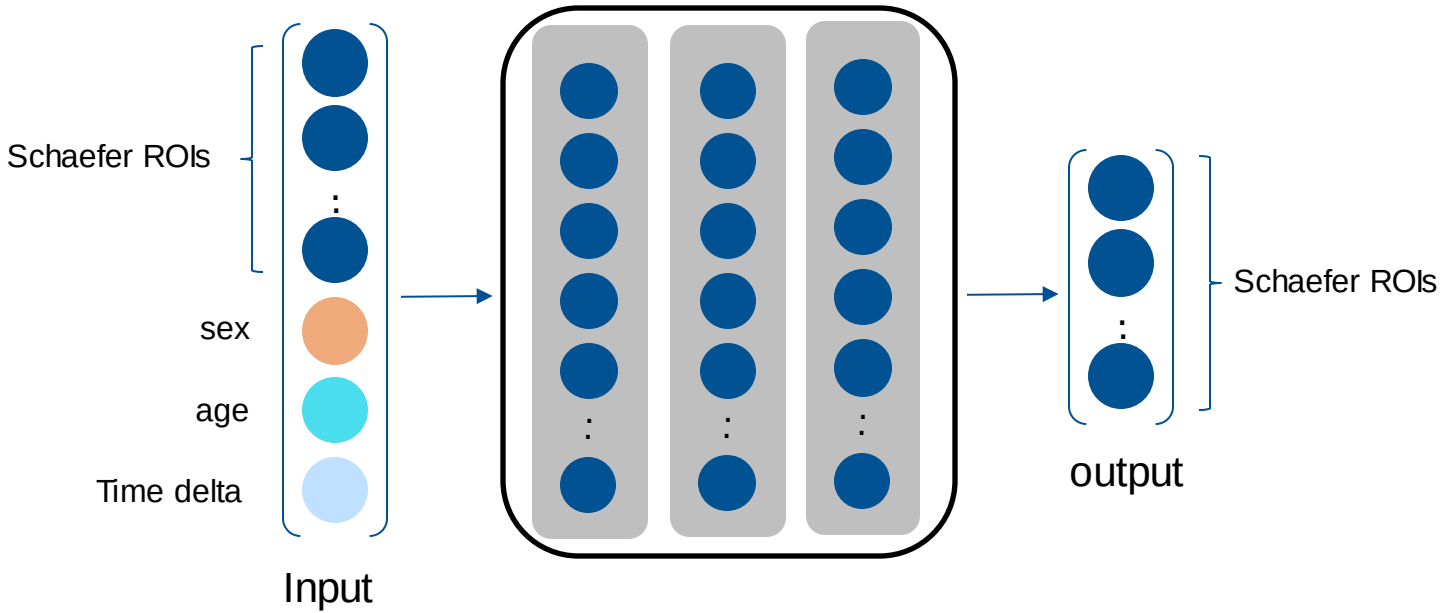


Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

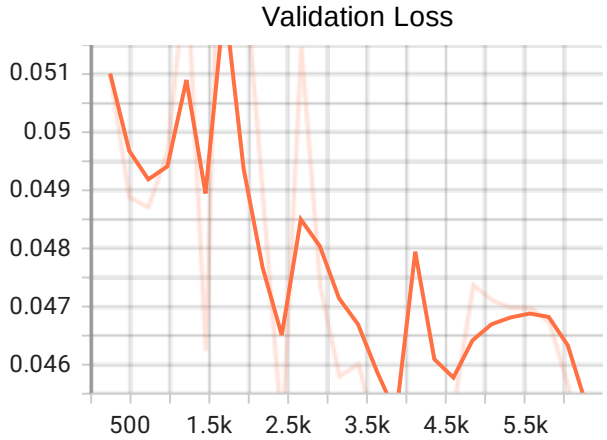
2. Progress and Findings

- Forecasting
  - **MLP**
  - LSTM
  - Transformer
- Overview of the Training Results

MLP Architecture

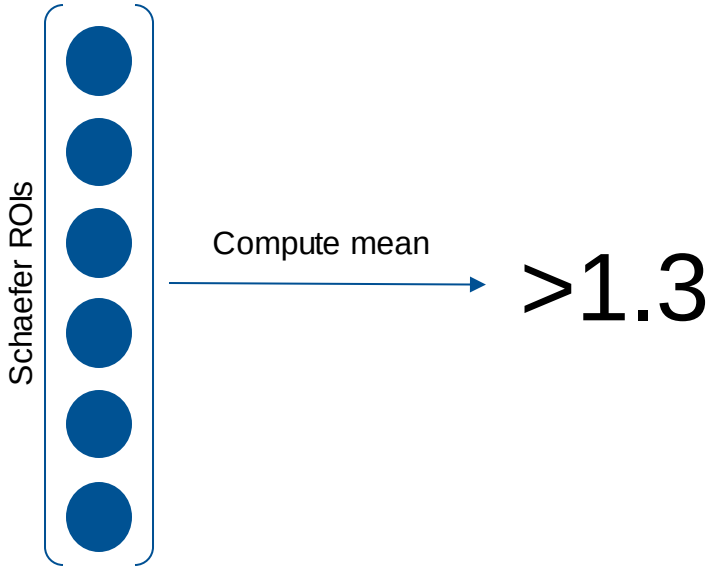


MLP: Results



Test loss	Test Accuracy
0.036	**0.898**

Biological Accuracy: Tau Positivity

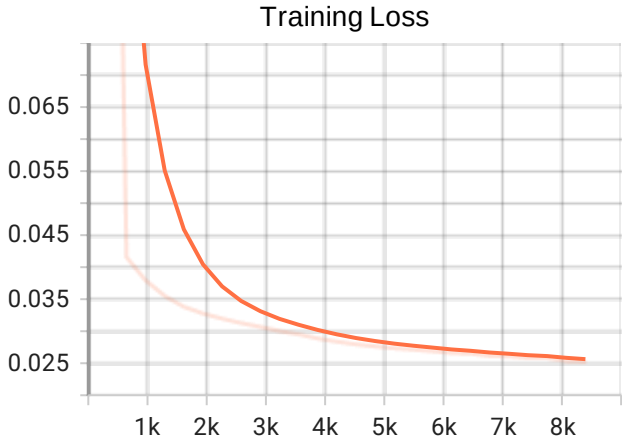


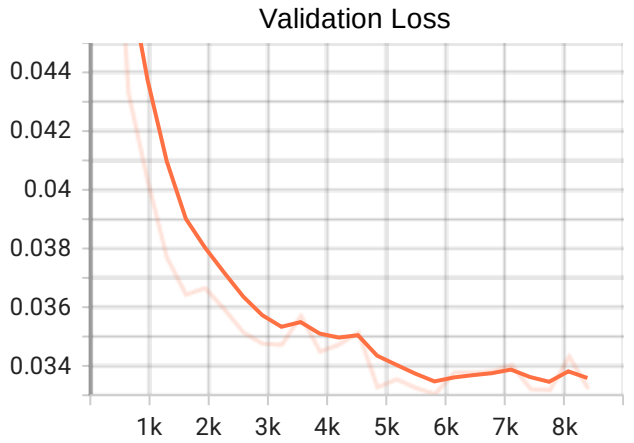
2. Progress and Findings

- Forecasting
  - MLP
  - **LSTM**
  - Transformer
- Overview of the Training Results

LSTM Model: Architecture

LSTM Model: Results





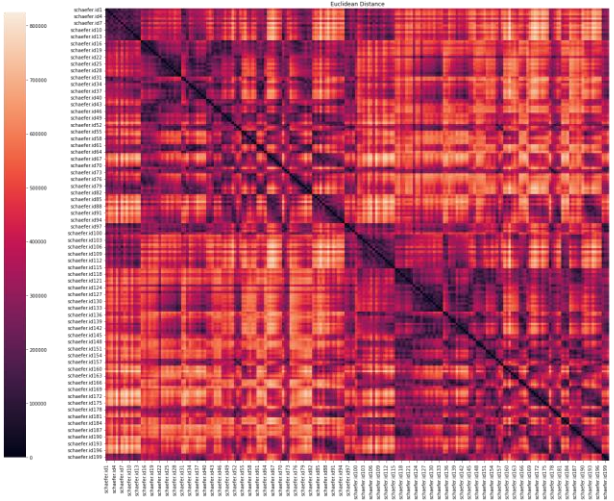
Test loss	Test Accuracy
0.0297	**0.939**

2. Progress and Findings

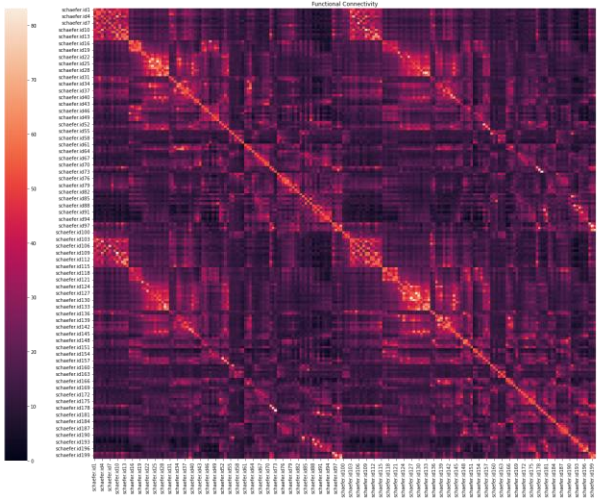
- Forecasting
  - MLP
  - LSTM
  - **Transformer**
- Overview of the Training Results

Connectivity Data

Connectome



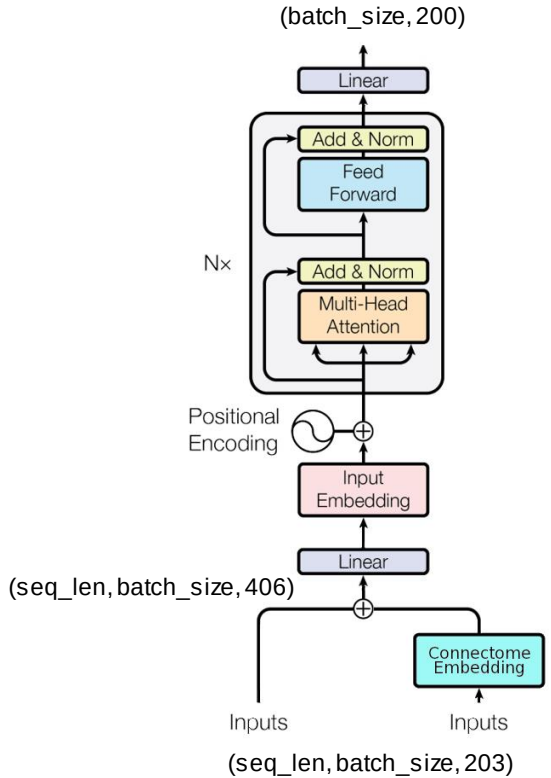
Euclidean Distance



Functional Connectivity

Transformer with Early Fusion of Connectivity Data

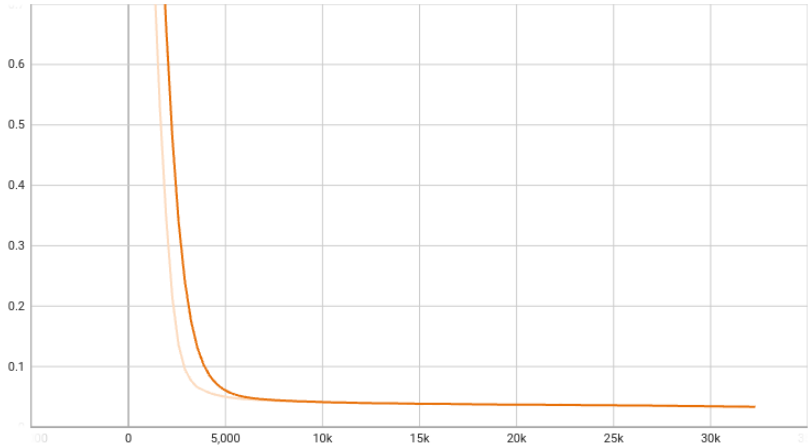
- Connectome Embedding Layer
  - Initialized with normalized functional connectivity data
  - Weights are frozen
  - In parallel for every entry in sequence
- Concatenation of original input and embedding is fed into the Transformer Encoder



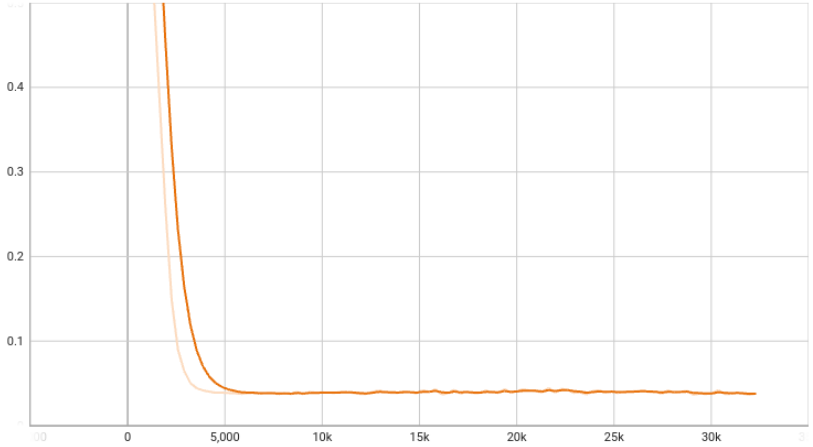
Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

Early Fusion Model: Results

Training Loss



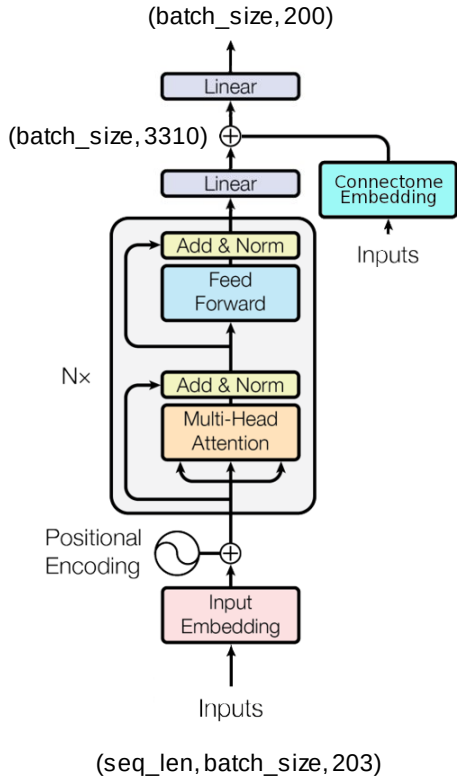
Validation Loss



Test Loss	Test Accuracy
0.0307	**0.9536**

Transformer with Late Fusion of Connectivity Data

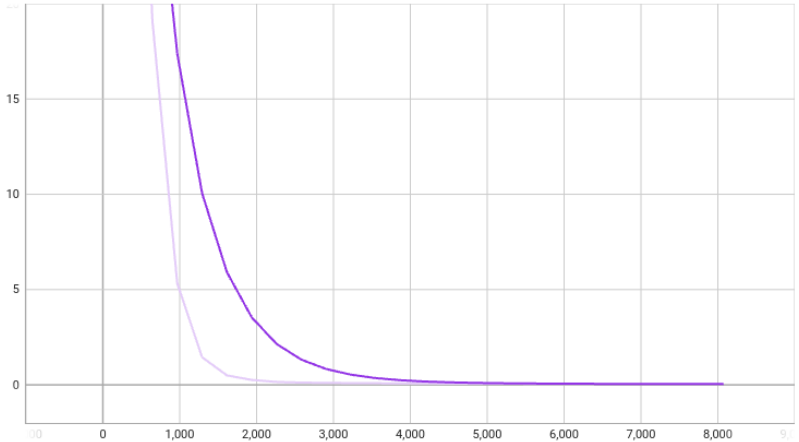
- Connectome Embedding Layer
  - Same procedure, but sequence flattened before concatenation
- Input dimension of final linear layer:
  (batch_size, (d_model + 203) * max_seq_len)
- Additional activation function



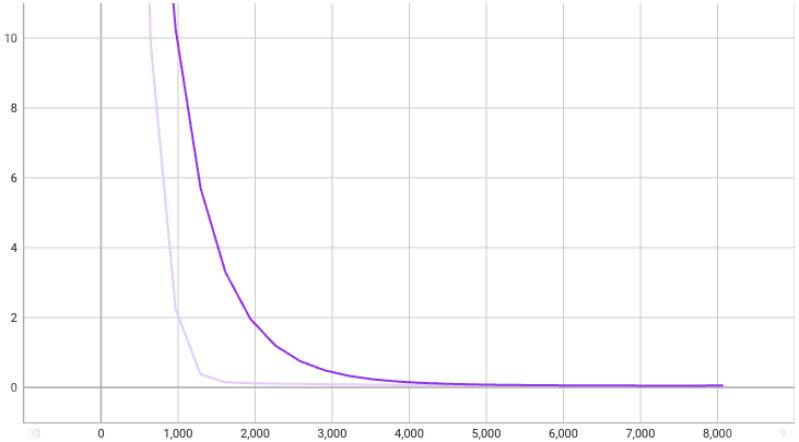
Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

Late Fusion Model: Results

Training Loss



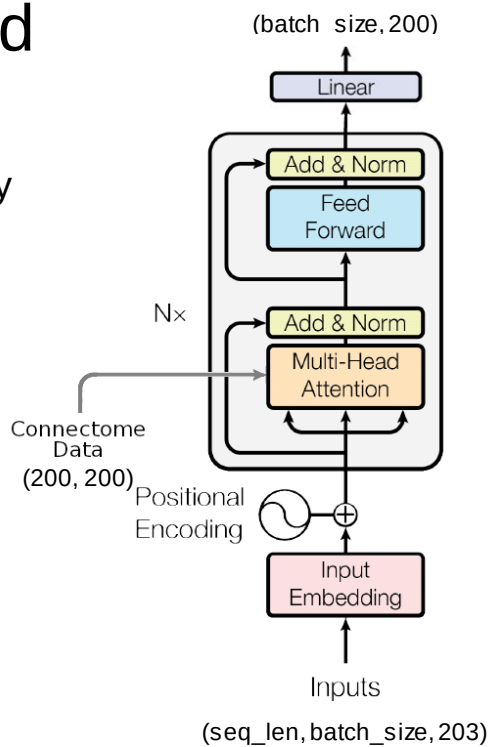
Validation Loss



Test Loss	Test Accuracy
0.0441	**0.9120**

Transformer with Connectivity Initialized Attention

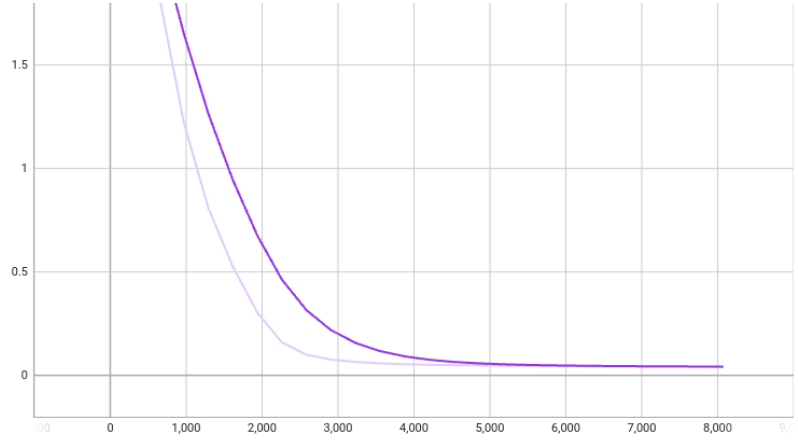
- Extension with identity entries omitted, as attention dimensionality needs to be multiple of `n_attn_heads`
- Instead: raw `200x200` Schaefer functional connectivity matrix is used for initialization of Q matrix
- Weights not frozen



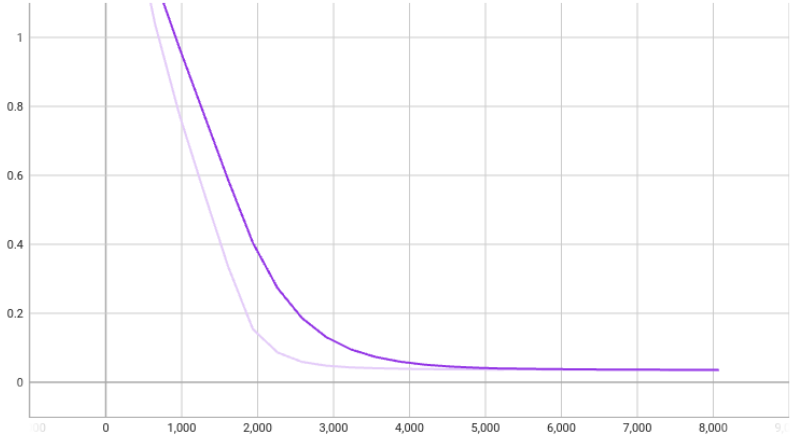
Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

Connectivity Initialized Attention Model: Results

Training Loss



Validation Loss



Test Loss	Test Accuracy
0.0306	**0.9445**

2. Progress and Findings

- Forecasting
  - MLP
  - LSTM
  - Transformer
- **Overview of the Training Results**

Overview of the Training Results

	Test loss	Test Accuracy
MLP	0.036	0.898
LSTM	**0.0297**	0.939
Transformer	0.0319	**0.954**
Early Fusion	0.0307	0.9536
Late Fusion	0.0441	0.9120
Initialized Attention	0.0306	0.9445

3. Next Steps

- Connectivity Informed Attention
- Long term Spatio-temporal Transformer Architecture

3. Next Steps

- **Connectivity Informed Attention**
- Long term Spatio-temporal Transformer Architecture

3. Next Steps

- Connectivity Informed Attention
- **Long-Range Spatiotemporal Transformer**

Long-Range Spatiotemporal Transformer Architecture

Long-Range Transformers for Dynamic Spatiotemporal Forecasting

Jake Grigsby
University of Virginia
jcg6dn@virginia.edu

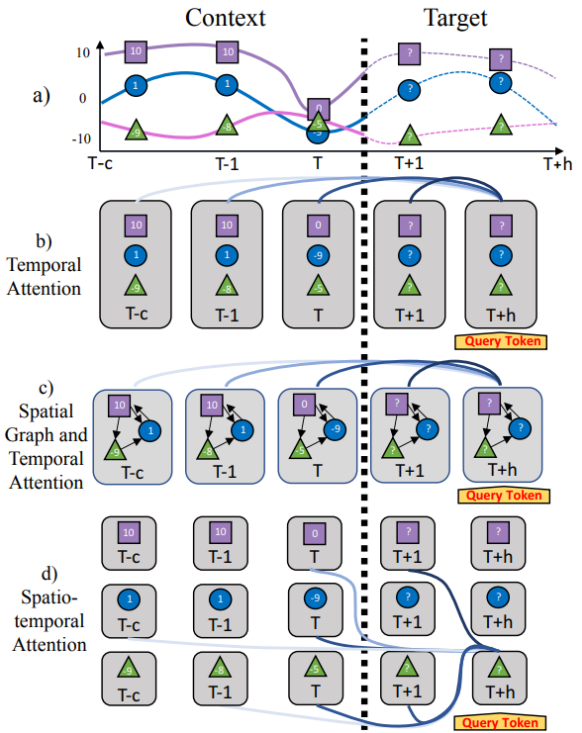
Zhe Wang
University of Virginia
zw6sg@virginia.edu

Yanjun Qi
University of Virginia
yanjun@virginia.edu

	Linear AR	LSTNet	LSTM	MTGNN	Temporal	Spacetimeformer
MSE	14.3	15.09	14.35	11.40	14.3	8.96
MAE	2.29	2.08	2.02	1.76	1.85	1.43

Table 5: AL Solar Results.

arXiv:2109.12218v2



Thank you for your attention!