Neural Network Structures for Traffic Forecasting

William Zhuk

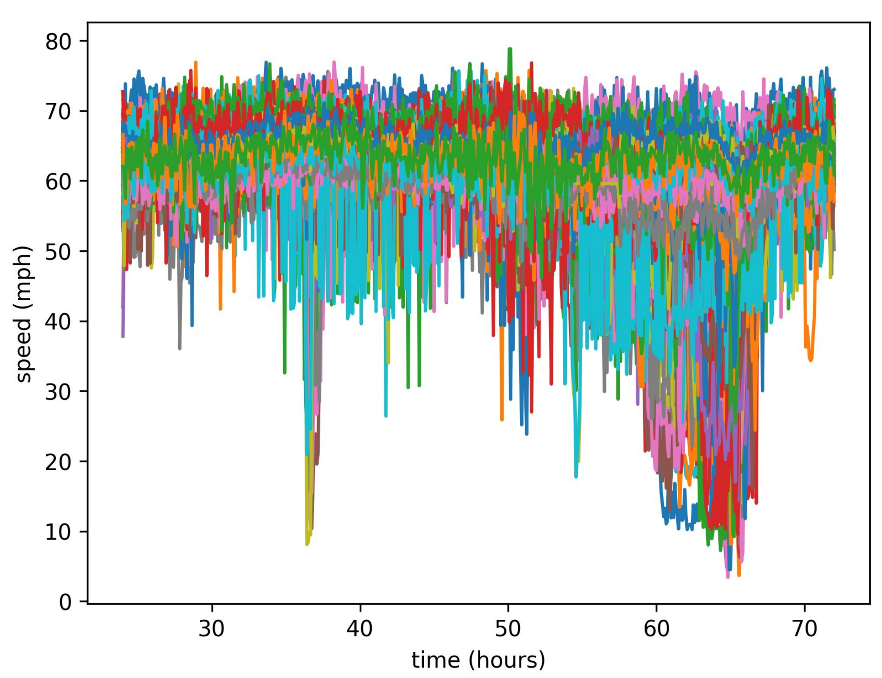
Introduction

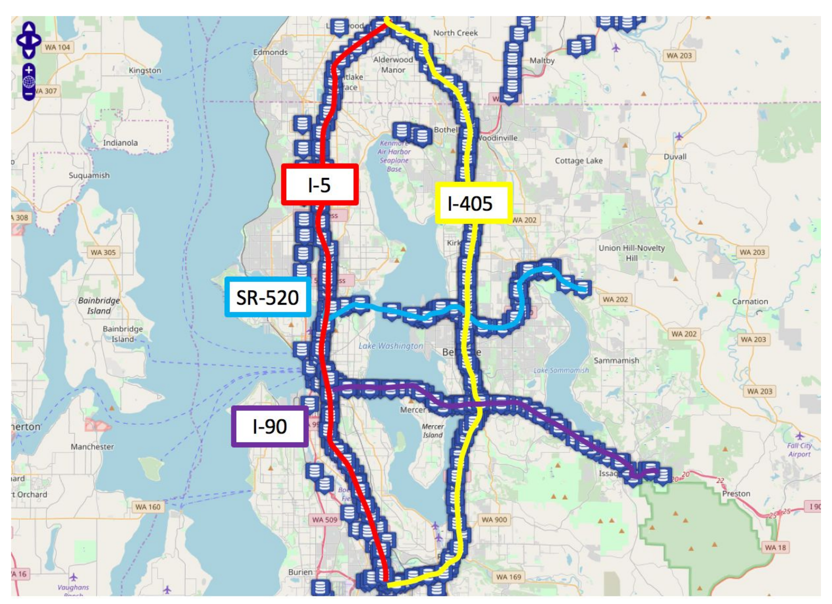
Traffic Prediction is an important field for customer-facing (Maps & Guidance) and control applications. Here we examine the importance of architecture decisions in the more recent graph convolutional models springing up in this field.

Dataset Characteristics and Acquisition

LOOP is a set of time-series speed data from road installed metal loop detectors in Seattle.

The set is given in 5 minute segments over a 1 year time period. An adjacency matrix is also provided.

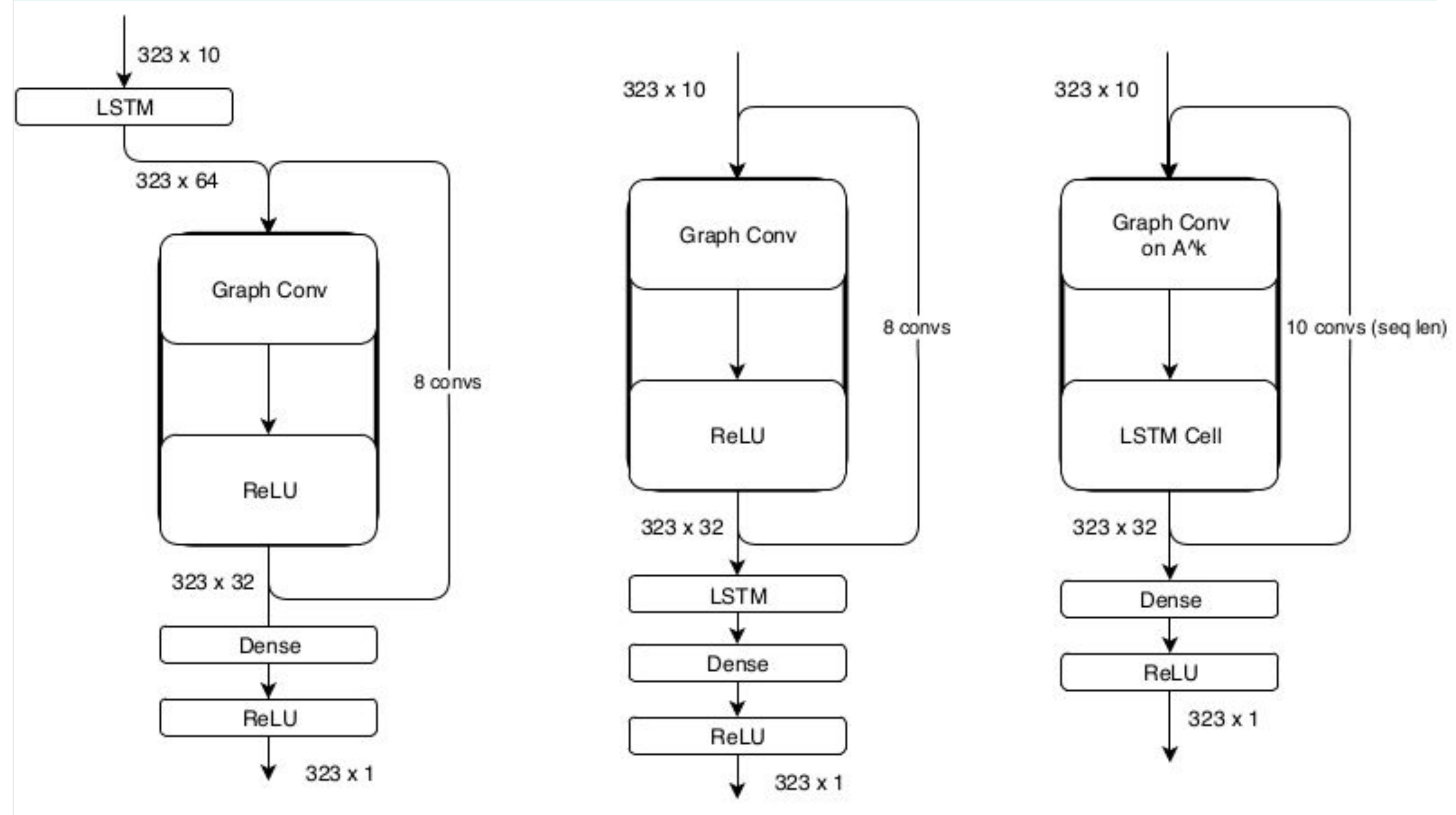




Lack of Features

The dataset passed in has only 5-minute averaged speed as a feature. The rest of the features are to be extracted via the graph and the sequence nature of the data. This is where LSTMs and GCNs shine.

Neural Networks Architecture



LSTM-Convolution

The LSTM obtains the sequence of input speeds, encodes them, and passes them to graph convolutions that encode graph topology into the dataset before feeding to some deep layers.

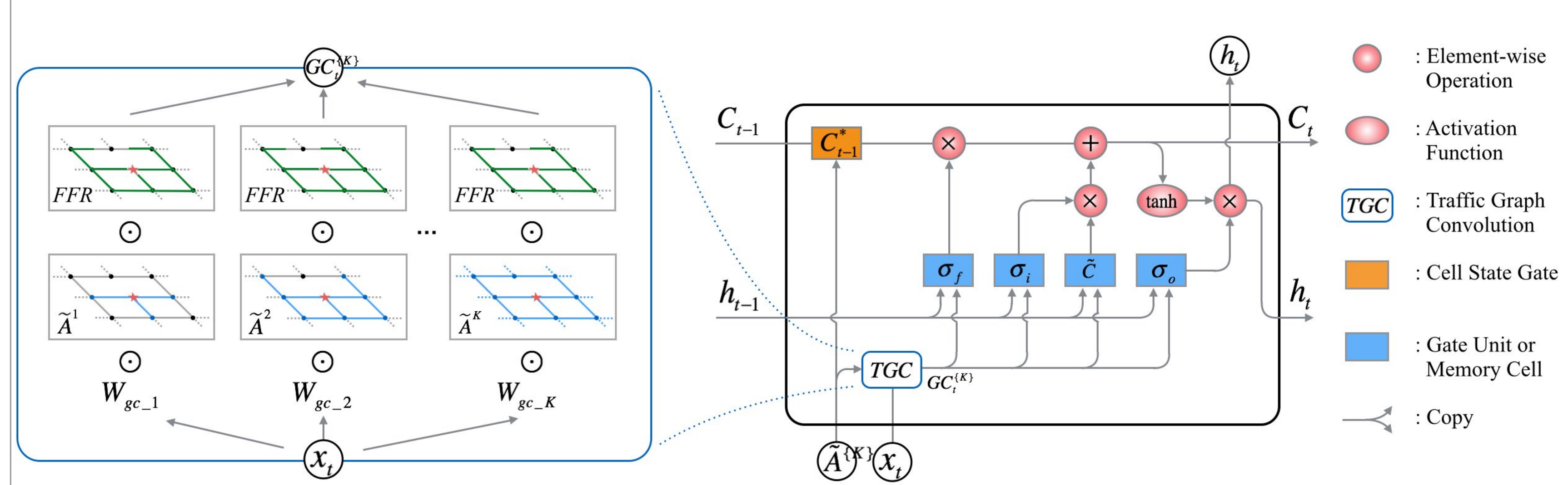
Convolution-LSTM

First the sequenced data is sent through graph convolutions, then the encoded results are given the the LSTM as a sequence. Finally, passed to deep layers.

LSTM-in-Conv

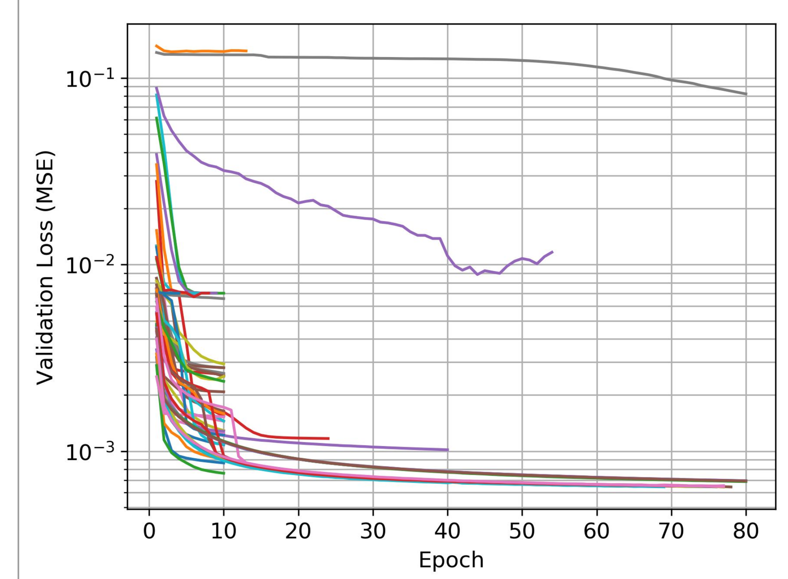
Every convolution feeds the results of aggregation into a LSTM-Cell before sending to the next convolution.

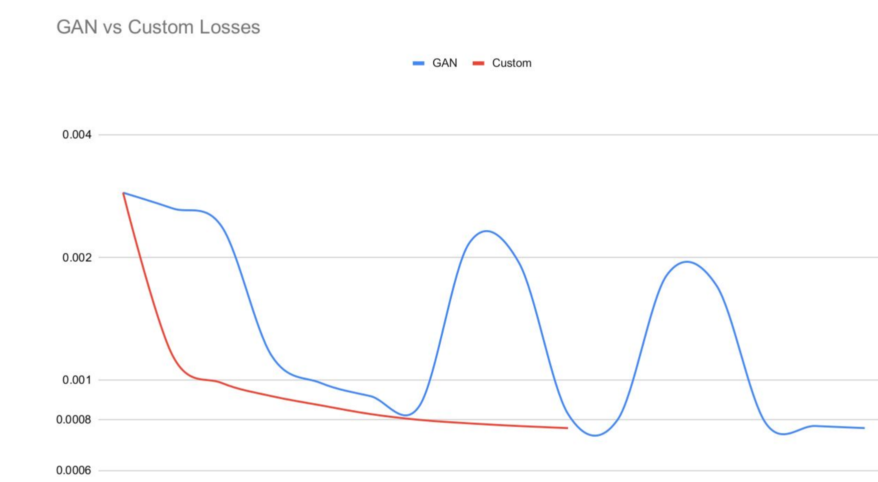
Existing Models



Hyperparameters

A major part of this project was encoding and searching hyperparameters to verify models. Here is an example run on the models with the LSTM Cells inside of the convolutions.





The best models performed roughly equivalent, although the best LSTM-GAT model seemed to be fluctuating and brittle, unlike the LSTM-in-Conv together (or the original paper)

Discussion / Conclusion

The models performed equally even though the expressiveness of the LSTM-in-Conv model seemed far more powerful than the series models! Even more surprising is that using multi-hop graphs to increase convolutional sizes did not impact the results heavily.

In the end, it seems that repeated small convolutions on the original modified adjacency matrix with a series approach seems to get the most bang for processing time.

Future Work

- Performance on larger and more feature-rich datasets
- Performance for longer time horizons
- Susceptibility to noise of LSTM-in-Conv
- Enhancing LSTM-cell in Convolution with deep layers.