The Impact of Covid-19 on Air Traffic: Spatiotemporal Forecasting and Benchmarking

Adelle Driker, Bo Yan, Yuan Hu

Advisor: Professor Rose Yu

Modeling

Analytic Approach

The data points that are to be ingested by the model currently contain the following attributes: location (country), time stamp (day), and number of flights. The first two columns will be the inputs that the model will use to calculate the target number of flights. Please note that we will be adding more inputs as we work to improve our model's performance. Currently, we are utilizing the three core models: two traditional time series models - AR and ARIMA - and a deep learning model - LSTM. In addition, we are implementing and tuning a Seq2Seq model and DCRNN. For additional information, we have also implemented LSTM with Mean Interval Score Regression and LSTM with Quantile Regression models.

Another type of model that we have experimented with is the Sequence to Sequence (Seq2Seq) model. This model is part of a class of RNNs that is typically used to solve complex language problems like Question Answering, Text Summarization, etc. The algorithm, originally developed by google for Machine Translation, relies on an Encoder-Decoder LSTM (or GRU) model, which helps avoid the problem of the vanishing gradient. Each component is turned into a corresponding hidden vector by the Encoder, which is then decoded into an output component that is used as input in the next layer.

One more model that we have begun to research and experiment with is the DCRNN - Diffusion Convolutional Recurrent Neural Network. This model specifically helps capture spatial dependencies using bidirectional random walks and temporal dependencies using the encoder-decoder architecture with scheduled sampling. When tested on two real-world large scale traffic datasets, this model consistently performed 12%-15% better than the state-of-the-art models.

Model Description

The LSTM with Mean Interval Score Regression model uses a mean interval score loss function with a LSTM network. And the LSTM with Quantile Regression modeling uses a quantile loss function with a LSTM network. We continued using different deep learning models to train against our whole dataset and tuned performance.

Deep Learning Models:

- **lmst-mis-regression model**
    - Training: LSTM
    - Scoring: Mean interval Score
    - Learner Parameterization:
        - Input size: 1
        - Output size: 1, 2
        - Hidden size: 16
        - Interval: 0.95
        - Optimizer: optim.Adam
        - lr: 0.005

        - Epochs: 100
        - Batch size: 10

- **lmst-quantile-regression model**
    - Training: LSTM
    - Scoring: quantile loss function
    - Learner Parameterization:
        - Input Dim: 1
        - Output Dim: 1
        - lr: 0.01
        - Quantiles: 0.025, 0.5, 0.975

Model Performance

Notebook: [lmst-mis-regression.ipynb](), [lstm-quantile-regression.ipynb]()

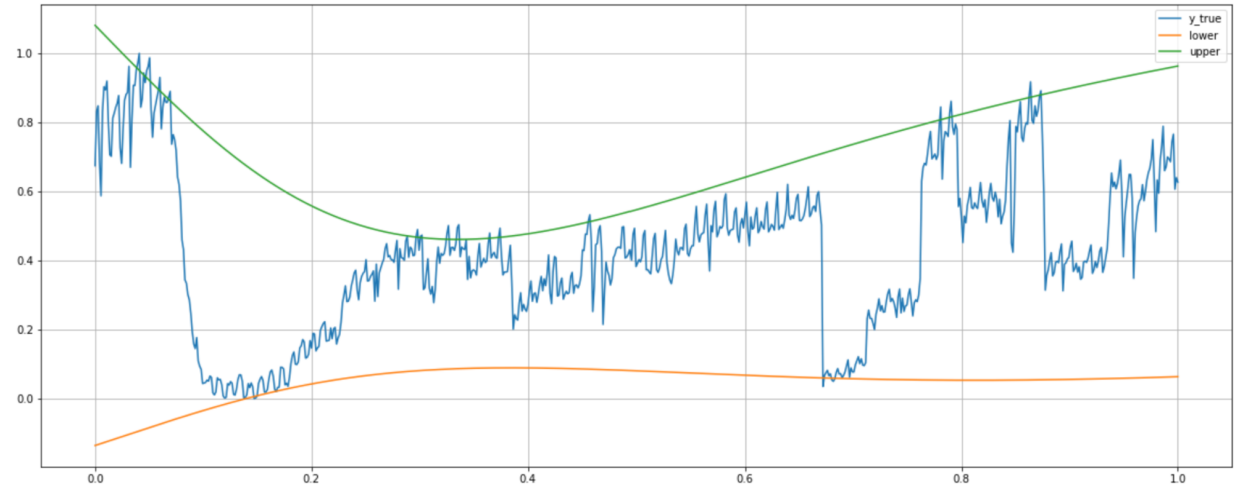

Fig. 1 lstm-mis-regression model to predict one confidence interval

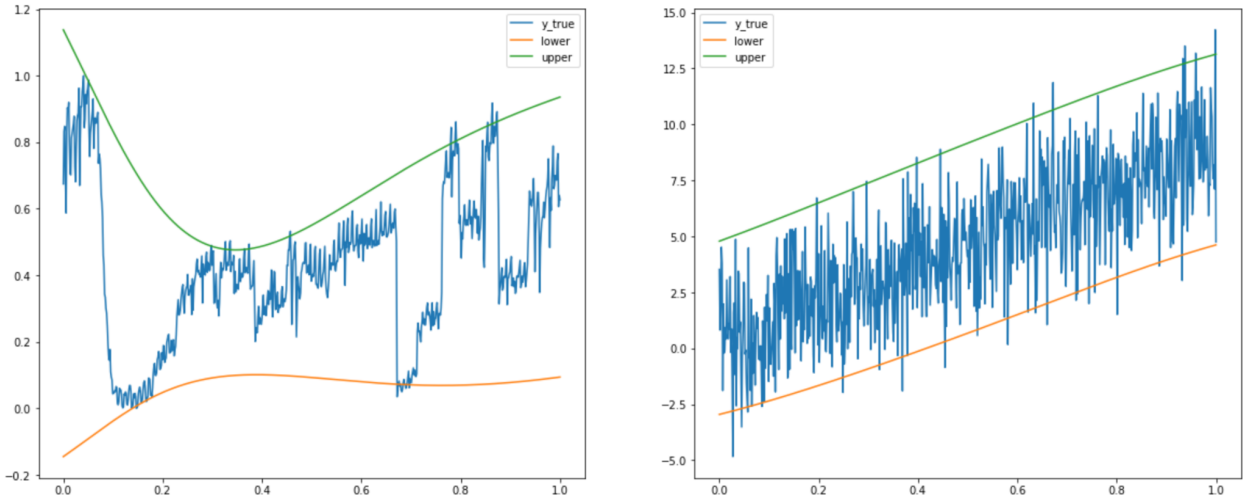

Fig. 2 lstm-mis-regression model to predict two confidence intervals

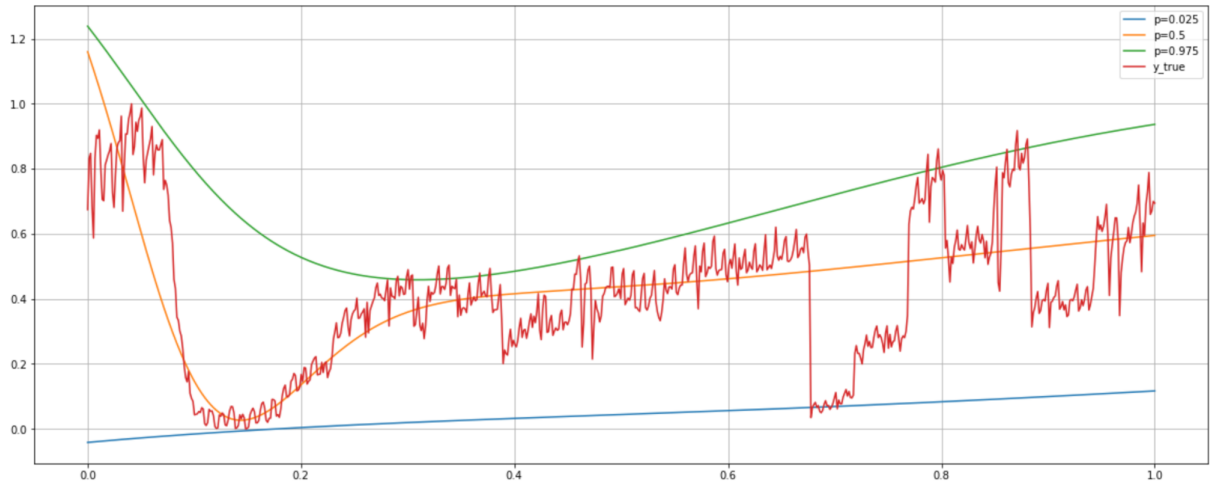

Fig. 3 lstm-quantile-regression modeling results

Model Interpretation

- The figures above depict the predictions based on two types of models - the LSTM-MIS- and LSTM-Quantile-Regression models.
- The LSTM-MIS-Regression model results consistently tend to stay within the estimation corridor, and wherever the plot travels beyond the confidence intervals shown in blue and green indicates that the confidence interval becomes narrower. A larger sample will reduce the sampling error, give more precise estimates and thus smaller intervals.
- The LSTM-Quantile-Regression modeling results show that the required quantiles, 0.025, 0.5 and 0.975, have 3 output nodes, with each node having a different loss function. This ensures that the structure of the data is shared in the first few layers.

Conclusion and Discussions for Next Steps

- The LSTM models with confidence intervals confirm that the LSTM creates accurate predictions (with a small number of points ranging outside of the interval)
- Having trained our model on the full set of data, we can already see a definite improvement in the results using the new deep learning models (lstm-mis-regression and lstm-quantile-regression).
- The LSTM DL model performed sufficiently well, but some improvements can still be made.
- Further updates to the dataset include modifying the location column to include an airport name along with its latitude and longitude coordinates so that a more granular analysis may be conducted.
- Further performance tuning is needed.
- We also need to rank models based on the performance of their predicting results to better serve the creation of the benchmarks.

Team Member Contributions

Bo:

- Wrote scripts for lstm-mis-regression, lstm-quantile-regression and Seq2Seq models, performed debugging and performance tuning for all these models
- Further research on Seq2Seq and DCRNN model
- Researched on Open Graph Benchmark and performed setup and testing
- Maintained Capstone Project Planning spreadsheet, GitHub, and Rose's Documentation

Yuan:

- Managed AWS EC2 instances for hosting database(influxDB) and dashboard(grafana)
- Wrote scripts to ingest processed data into database
- Initiated grafana dashboard from querying time-series data from database

Adelle:

- Coordinated meetings between our group and advisor
- Contributed to Analytic Approach, Model Description, Model Performance, Model Interpretation, Conclusion. and Major Updates sections

Major Updates to Steps 1-6

- Since our previous milestone, we have continued to tune both the LSTM and Seq2Seq models that we previously implemented.
- We have also begun to research and implement a DCRNN model as advised by Professor Yu.
- Besides the model, we have also begun to put together a dashboard using Grafana which will be available to the end user with our modeling results. Although this portion is not assigned for this milestone, we believe it would be helpful to begin the planning stages of the final product. We aim to have this dashboard available with working functionality by the next presentation.
- We also began researching the leaderboard implementation for creating benchmarks against different models.