

Forecasting Emergency Surgery and Inpatient Volumes

Yuan Cao, Isabel Chalfant, Joseph Evans, Yash Handa, Stanley Kim, Shreya Laxmi Nagendra, Tony Lin, Yuanfei Song



The Data Mine

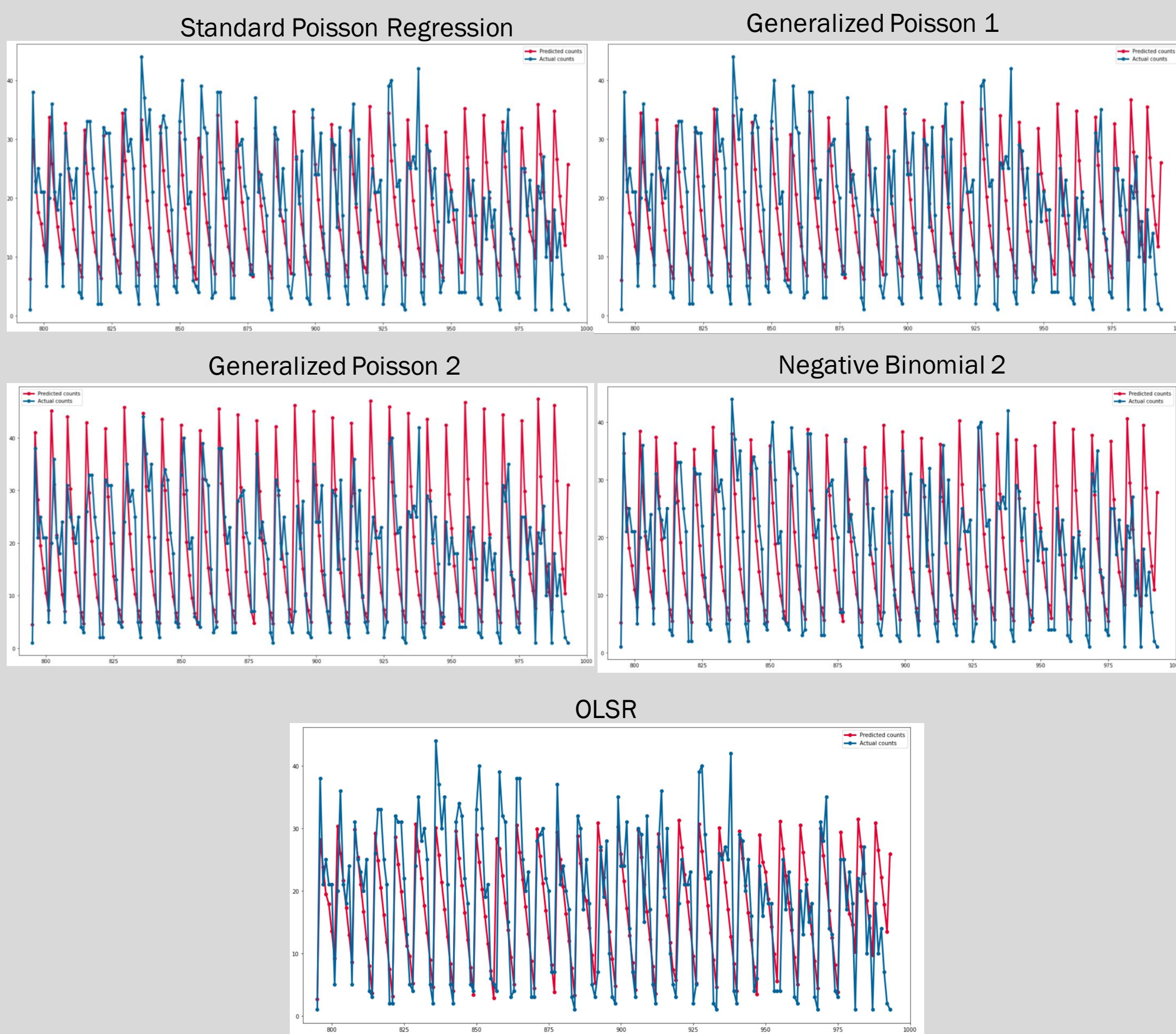


Introduction: Indiana University Health provides personalized, high-quality care in partnership with the Indiana University School of Medicine. The goal of the project was to develop a forecasting model to predict different volumes that affect nurse scheduling with a focus on the effect of the COVID-19 pandemic.

Methodology:

- Data Engineering
 - Perform basic analysis on surgery and inpatient volumes
 - Split data by hospital location
 - Organize and clean data for modeling
- Data Science
 - Research various forecasting models for short-term forecasting
 - Focus on generalized linear models and LSTM
 - Implement model in Python and train on given data
 - Analyze and compare performance of each model

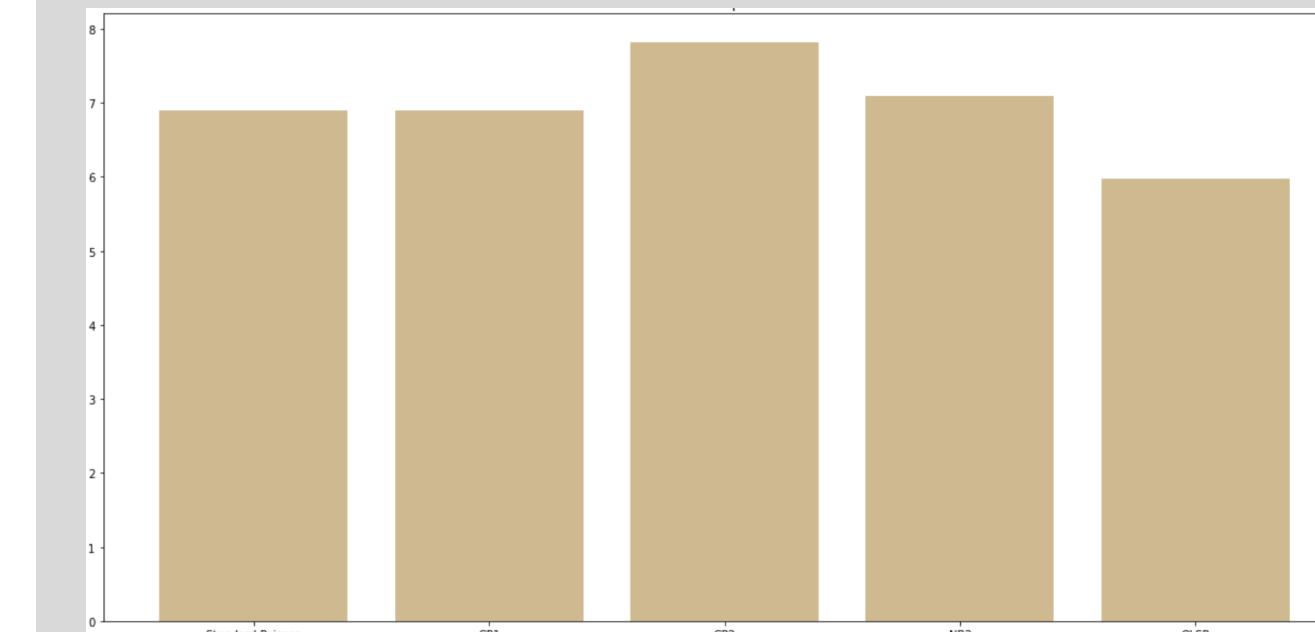
Generalized Linear Models: Used different GLMs to model emergency surgeries across the entire IU Health system. Predictions for test dates compared below. Test values range from March 2021 to December 2021, using the previous two years to train the models. The ordinary least squares regression model resulting the lowest RMSE and MAE



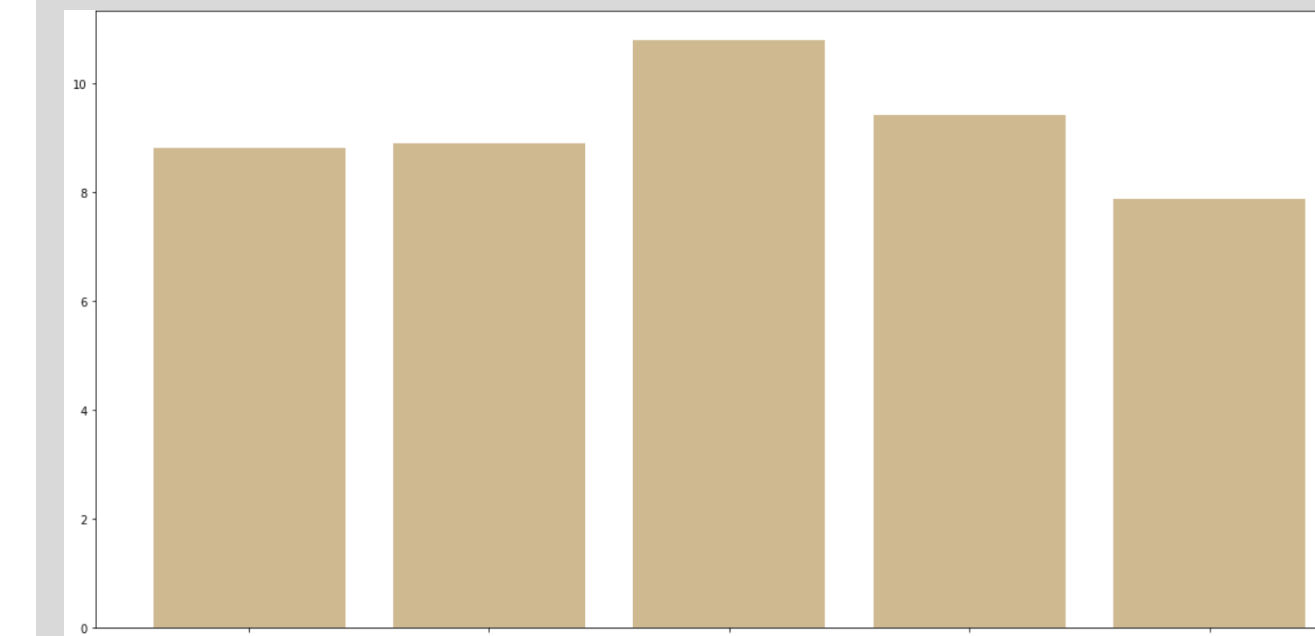
Data Visualization: We focused on two sets of data in the IU health system; surgery volumes and inpatient volumes. Pictured below.



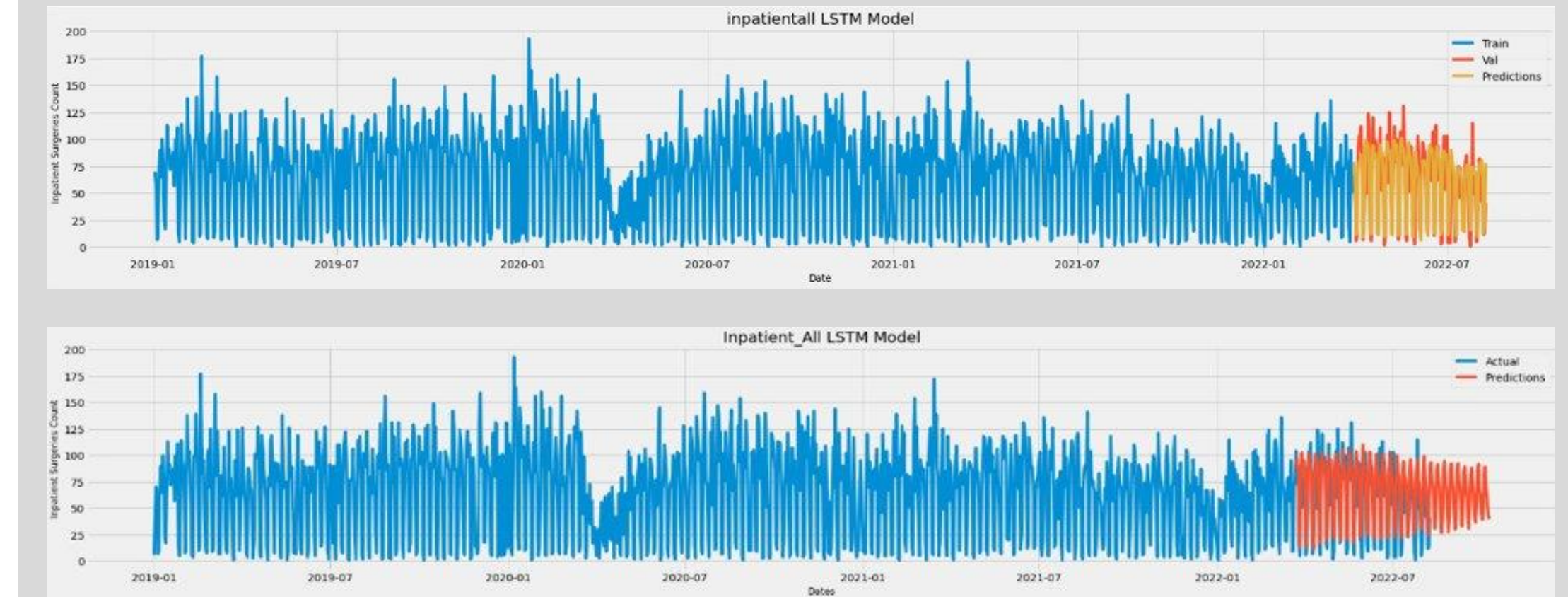
MAE Comparison



RMSE Comparison

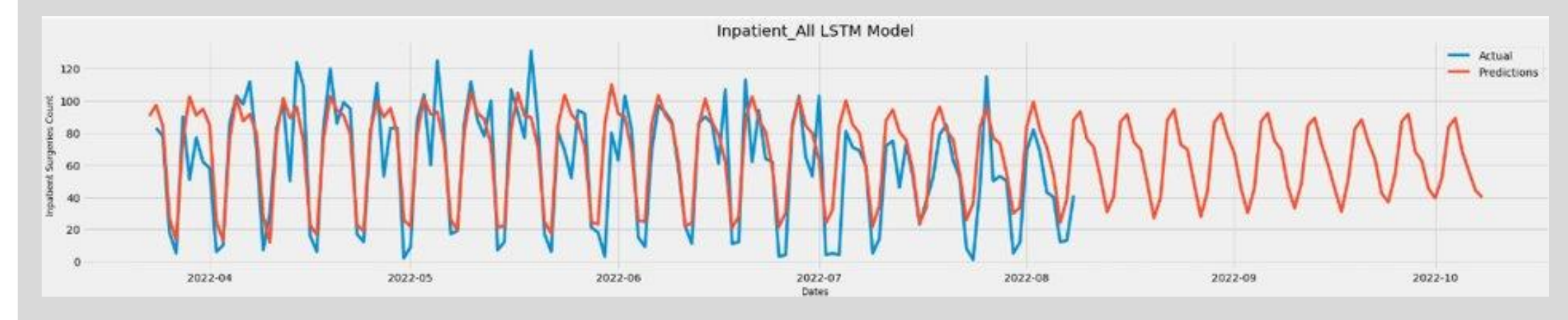


Results of GLMs: All the models provide decent predictions for surgery volumes however, OLSR provides the best predictions based on MAE and RMSE values. It is important to note that towards the end of the predictions the values are less and less accurate. These models perform well for short-term predictions but are less accurate the further in the future the predictions are made.



LSTM (Long Short-Term Memory) Neural Network: LSTM is a recurrent neural network used for forecasting, especially powerful when there is a longer-term trend in the data. We decided to choose LSTM since it is designed to capture the complex relationships within the data to make an accurate predictions about future outcomes.

To optimize the LSTM model, we used grid search to tune the hyperparameters, which involved testing many hyperparameter combinations to find the best combination for the model. After tuning the hyperparameters, we achieved a 97% accuracy with an RMSE value of 20 and MAE value of 15 when compared to the test data. This model can then be used to forecast unseen future data for IU Health for about 2 months in advance.



Challenges:

- Grid search involved testing many hyperparameter combinations, and therefore, was very time-consuming and computationally demanding.
- Forecasting unseen data is different process than training the model, because it required the model to generalize well to new and unseen scenarios.
- Each of the models developed were unable to forecast as far into the future as desired without a large decrease in accuracy.

Future Goals:

- Develop long-term forecasting for both surgery and inpatient volumes
- Further improve predicting surgery volumes for each location in IU Health system
- Create simple and interactive way to predict future values

Acknowledgements: Nate Dotzlaf, Tiffany Rapp, Jamee Malan, Nan Kong, Kevin Cusimano, Gautam Choudhary

