Predicting Call Volumes for the Emergency Medical Services

Project Goals and Stakeholders

Our goal is to **predict future call volumes** for the Emergency Medical Services (EMS). Using data provided by the National EMS Information System, which includes specifications about what time calls occurred as well as anonymized location information, we want to **train models that are tuned for a given county or state** that will outperform baseline forecasts.

Accurate forecasts about call volumes would help **maximize efficiency for the EMS** by giving guidance about how to allocate their limited personnel and equipment, which could save lives. For instance, timely responses can significantly increase the odds of survival from cardiac arrest, so sufficient staffing is critical. On the other hand, EMS personnel on uneventful shifts may not receive the experience they need to provide effective treatment, yielding adverse patient outcomes.

Proposed Models

Our first model is the **SARIMA model**, which combines auto regression with a moving average forecast and takes seasonality into account. We observed weekly seasonality patterns both in our exploratory data analysis and autocorrelation analysis, so we set the seasonality parameter to 7 and used the Akaike information criterion to select the parameters for the auto regression and moving average parameters. We trained a separate model for each county (state?).

Jessica’s model info here

Our other model uses a standard neural network for time series analysis, the long short term memory cell, or LSTM. The LSTM takes in a sequence of observations and attempts to predict the next observation. For us, an observation consists of all the number of calls on a day, and for a more sophisticated model we separate out some of the more common call types for a county and lump all the other call types together. The number of separated call types, the length of the input sequence, i.e. the ‘window length’, and some details of the LSTM architecture are all hyperparameters which we tuned to find the most accurate models.

Model Performance

Using cross validation against the baseline naive and average forecasts, we found that our models beat the baseline forecasts by x% margins on average

For most counties, a window length of 7-10 days and 3 or more separated call types produced an LSTM which could perform significantly better than the average or naive forecasts, measured by MSE. Visually, the LSTM was able to learn the weekly cycles of different call types very well. Inspecting the data, we saw that the number of calls and quality of the data varied year to year dramatically in a way that could only be explained by data collection methods. The LSTM models were able to adapt to inconsistencies easily, quickly finding moving averages and short cycles. Some counties however resisted all attempts at modelling and the LSTMs performed the same as average or naive models. Inspecting these counties, we could see that there were no weekly cycles for the models to discover and the quality of the data was too bad to be useful.

Conclusions and Next Steps

We were able to achieve our goal of obtaining models trained on the state and county level that significantly beat baseline models. For future work, we would like to create an applet that summarizes our findings and allows others to use our models to make predictions based on their local data. We are also interested in implementing harmonic regressions models that could take multiple types of seasonality into account for even more accurate predictions.