Predicting Hate Crime Trends: A Time Series Model

To identify hate crime trends, we used time series modeling to generate forecasts through 2019, training on hate crime data from the FBI's Crime Data Explorer.[12] Specifically, we modeled a 24-step ahead (two-year) forecast for the U.S. as a whole, as well as for each municipality in our data set. Our projections are for hate crimes overall and hate crimes per capita.

Creating the model presented challenges. For instance, the FBI hate crimes data set has sparsity issues, both in terms of hate crimes not being common everywhere and problems with underreporting. We emphasize and caution the reader that forecasts are only as accurate as the quality and comprehensiveness of historical data, as we are relying on reported hate crimes. That being said, as we are mostly concerned with privacy issues stemming from crimes that were reported, the data set is sufficient for our purposes.

We used the "Prophet" time series model to address these challenges as rigorously as possible. Developed by Facebook Research, we selected this model due to its auto-tuning nature and helper methods that follow the scikit-learn framework (i.e. instantiate, fit, predict methods)[3]. Also, it produces reliable forecasts that are robust to outliers and data sparsity issues. The AutoML nature of Prophet made this efficient to tune and straightforward to program, with only roughly 30 lines of code. Finally, this is an efficient model to train, only taking around 10-15 minutes for 276 cities.[4]

While there are several ways to test a time series model, Prophet provides built-in functionality for cross-validation testing. This has been useful, as it provides a way to assess various error statistics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and upper/lower confidence interval bounds across train/test splits. Further, by providing output across various error statistics, it provides an additional way to assess confidence in our model.[5]

As with any auto-tuning model, the lack of a thorough exploratory data analysis leads to the possibility of sub-optimal parameter tuning choices, over or under-fitting, assumption violations, and other errors that can only be avoided via manual engineering. Given the scale of the dataset that we were working with (276 cities), using an AutoML-enabled time series model was a necessity. Despite these limitations, our model produced forecasts with an MAE low enough for us to assume good model fit for cities with more than 20 hate crimes per year, on average, which amounted to 57 of the 276. We chose MAE as our primary diagnostic metric since we wanted the result to be in the same units as the forecast (hate crime volume) and we did not want to apply an increased penalty to higher magnitude errors as MSE or RMSE do. Forecasts for cities with sparse hate crime data are less reliable, with 95% confidence intervals often on both sides of zero.

The output of our model, for the U.S. and each city, includes two years of monthly predictions (y-hat), through the end of 2019. To assess the trends for each municipality, we took the difference between the last two years of historical data, 2016-17, and the two years of forecasted data, 2018-19, to see if hate crimes were likely to be increasing, decreasing, or remaining relatively flat. We then combined this with our Municipality Data Security and Doxing Risk Assessment to generate key insights and recommendations.

To explore the time series output further, see our Data Exploration Tool.

Endnotes

1. CDE :: Home. (n.d.). Retrieved April 13, 2019, from https://crime-data-explorer.fr.cloud.gov/
2. The most recent version of the Hate Crimes Explorer documents all reported hate crimes from 1991 to 2017.
3. Prophet. (n.d.). Retrieved March 11, 2019, from http://facebook.github.io/prophet/
4. Taylor, S. J., & Letham, B. (2017). *Forecasting at scale* (No. e3190v2). PeerJ Inc. https://doi.org/10.7287/peerj.preprints.3190v2
5. ibid