Hate Crime Analysis and Forecasting

Revati Trivedi (rst8739)
Yamini Narasimhan (yl9822)
Pranjal Jain (pj2069)

Background and Motivation

With the rising amount of Hate Crimes, it is crucial to analyze and understand the trends of Hate crimes. These statistics may be used to:

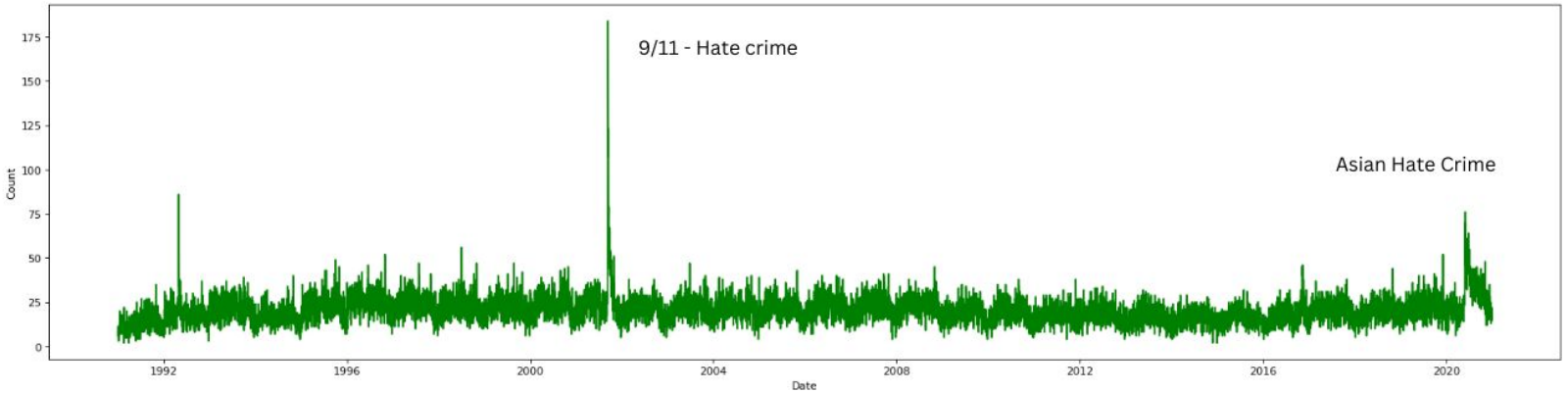
- help law enforcement address issues for their communities
- provide lawmakers with justification for certain legislation
- supply the media with credible information
- help researchers in determining trends in hate crimes
- Understand how previous instances of crime cause crime in the future

Dataset

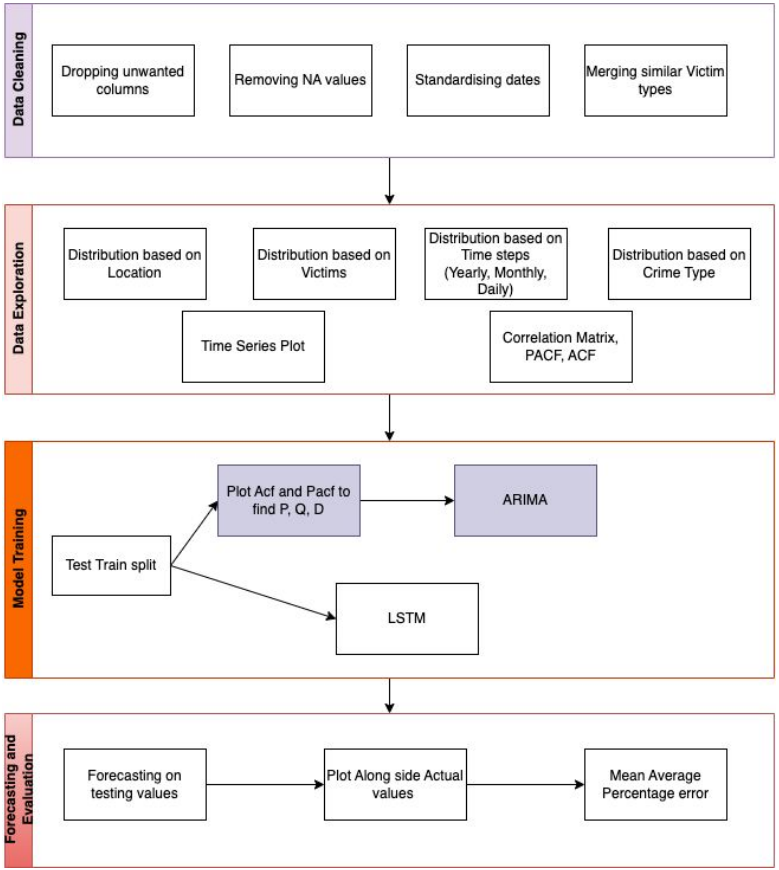
For our project, we will be using the FBI Crime Data Explorer's Hate Crime Statistics dataset. The dataset catalogs incidents of Hate crime from 1991 to 2020 with additional features such as the ethnicity of people involved, type of crime, location, and more. This data is collected by different agencies that all report the data to the FBI via the UCR program (Uniform Crime Reporting).

We have a total of 200,000 data points each covering attributes like date, count of victims, crime type, location, offender count.

Dataset

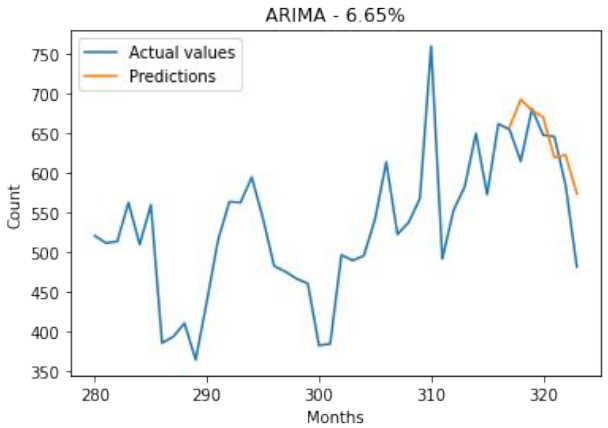

9/11 - Hate crime
Asian Hate Crime
Count
Date
0
25
50
75
100
125
150
175
1992
1996
2000
2004
2008
2012
2016
2020

Methodology

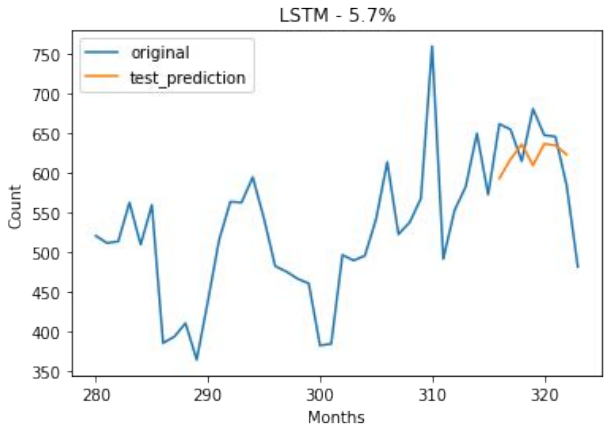


Models and Evaluation

1) Autoregressive integrated moving average



2) Long short-term memory



We have used mean average percentage error to verify the accuracy of our model along with the plotting the lines of predicted values in comparison to the actual values

Model Forecast on different horizons

Model	6 months	18 months	24 months
ARIMA	6.65	16.38	16.63
LSTM	5.7	7.74	10.8

Assumptions and Limitations

- Because the motivation of Hate crime is subjective, it is sometimes difficult to know with certainty whether a crime resulted from the offender's bias. This creeps into the dataset as its at the discretion of the investigating officer to classify a crime as Hate crime. This may result in many hate crimes not making into the dataset.
- The data is also not uniformly collected. There are locations where hate crime is not collected, making it difficult to make a general prediction.

Future work

- The current implementation can serve as an excellent base for creating a forecasting system that can actually predict trends of hate crime and even predict spikes in the dataset which are unrelated to past trends.
- We can use other sources such as economic trends, news or twitter sentiments and add that as a feedback in the LSTM we implement.

Questions?