Analysis of Global Terrorism

Shreehar Joshi, Debbie Yuster, Ph. D.
Computer Science and Data Science, School of Theoretical and Applied Sciences
Ramapo College of New Jersey, Mahwah, NJ 07430

Introduction

Terrorist attacks have been a significant concern for countries worldwide for many years, causing physical harm, emotional trauma and long-term socio-economic consequences. Despite efforts to prevent and counteract terrorism, the threat remains persistent, with new organizations forming and established ones changing their strategies. This project examines the terrorist incidents that occurred globally from 1970 to 2017, analyzing various aspects of the attacks, including methods, targets, motives, and frequency, and their impact on socio-economic factors such as GDP, migration, and population. It also evaluates the effectiveness of different deep learning models in predicting the number of casualties from existing incident data.

Methodologies

Three open-source datasets were used in this project. The first dataset is the Global Terrorism Database, which contains information on over 180,000 global terrorist attacks from 1970 to 2017. The second dataset is the World GDP dataset, which includes the GDP per Capita data of different countries globally from 1960 to 2021. Lastly, the third dataset is the World Population dataset, which provides the data on fertility rate and net migration of different countries from 1955 to 2020.

To process the data, NumPy and Pandas libraries were used while to visualize the results, Plotly, Seaborn, and BCR were primarily employed. For the modelling phase, the dataset was split into a train, validation and test sets in the ratio 70:15:15 to assess the efficiencies of different deep and machine learning models in predicting the number of casualties in any given attack. The models were built using the Tensorflow, Keras and Scikit learn frameworks from python.

Results

The key findings of the project are as follows:

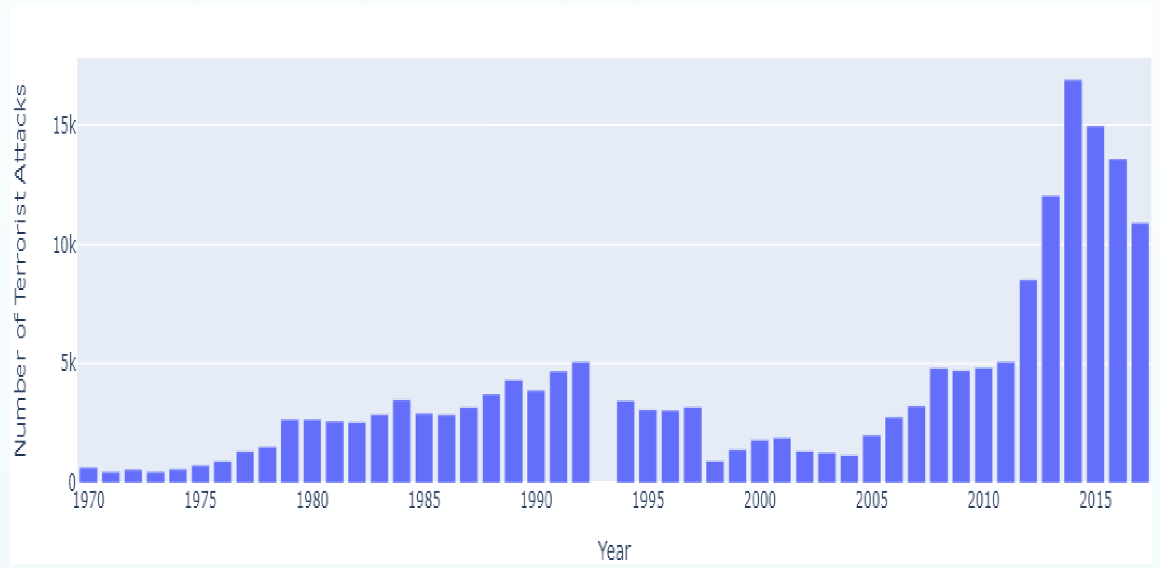


Fig 1: Frequency of terrorist attacks

The frequency of terrorist attacks was at its minimum at around 1972 and 2003 and has greatly increased over the last five years of the period that was analyzed.

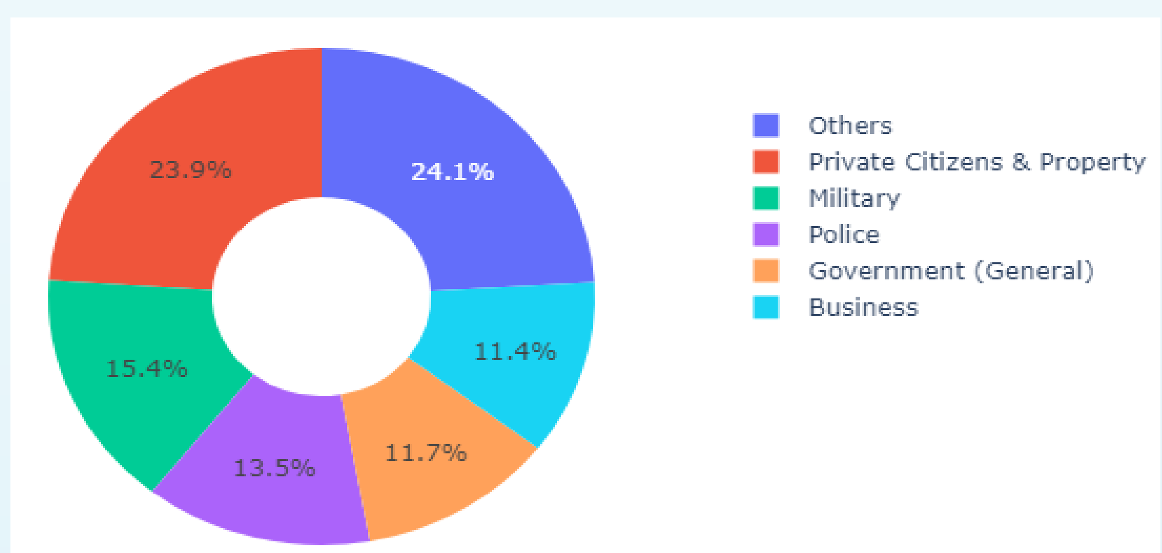


Fig 2: Target types of terrorists

Terrorist groups often targeted private citizens and property, military and police, and these accounted for more than 50% of the target types.

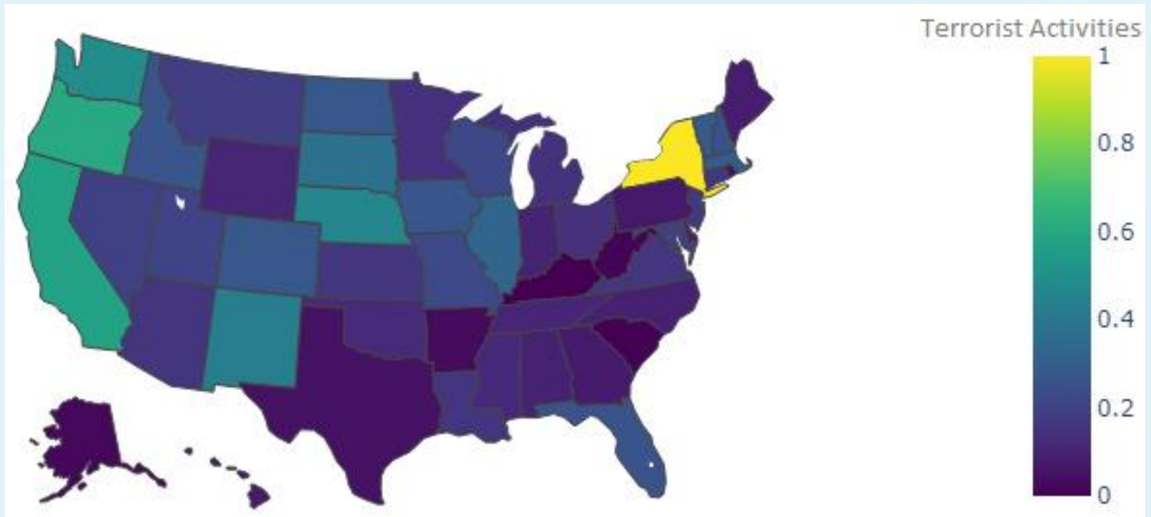


Fig 3: Terrorist Activities in the US

New York, Oregon, California, Washington and Nebraska were the top five US states which had the highest per capita terrorist incidents. Kentucky, South Carolina and West Virginia were the safest states in terms of terrorist attacks.

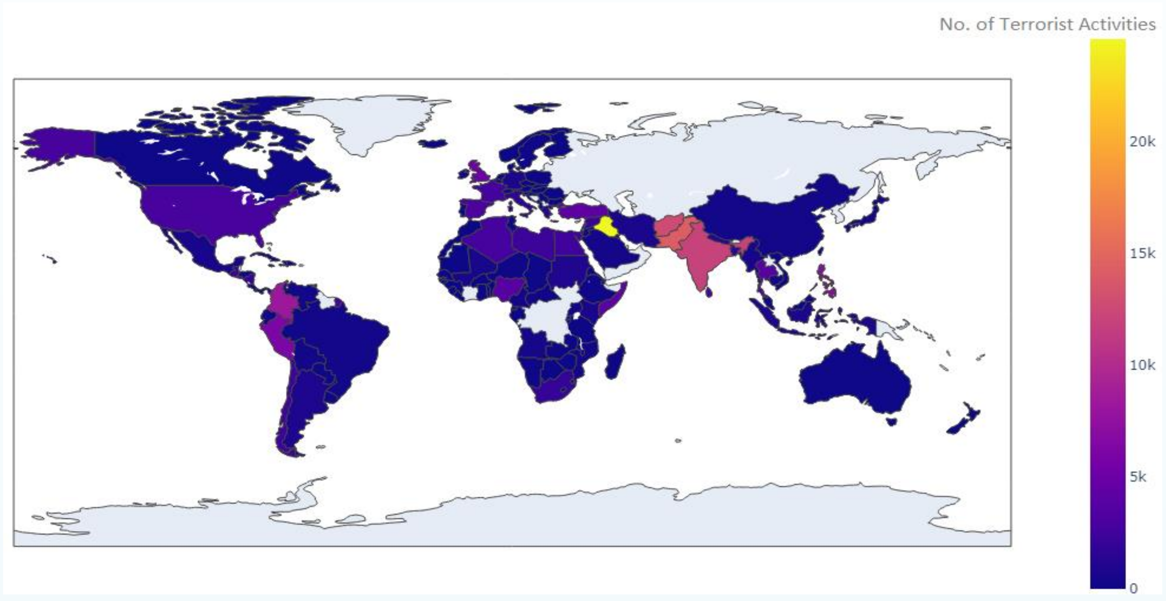
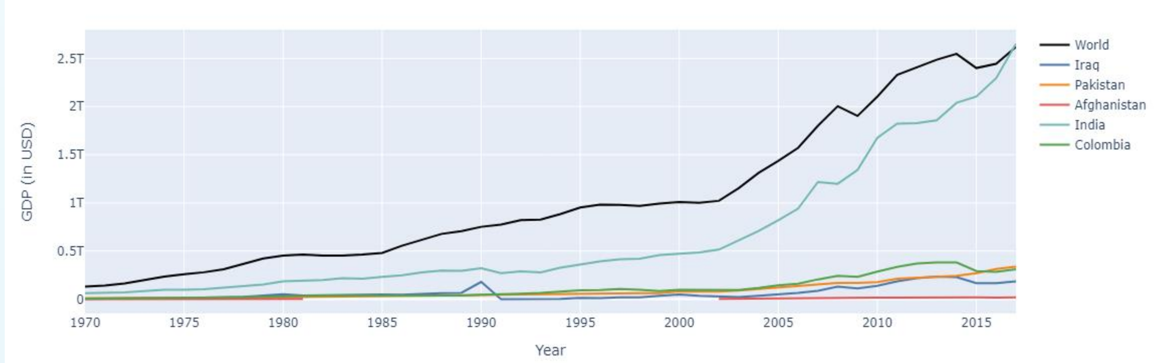


Fig 4: Global terrorist incidents

Iraq, Pakistan, Afghanistan, India, and Colombia were the most terrorist-prone countries.



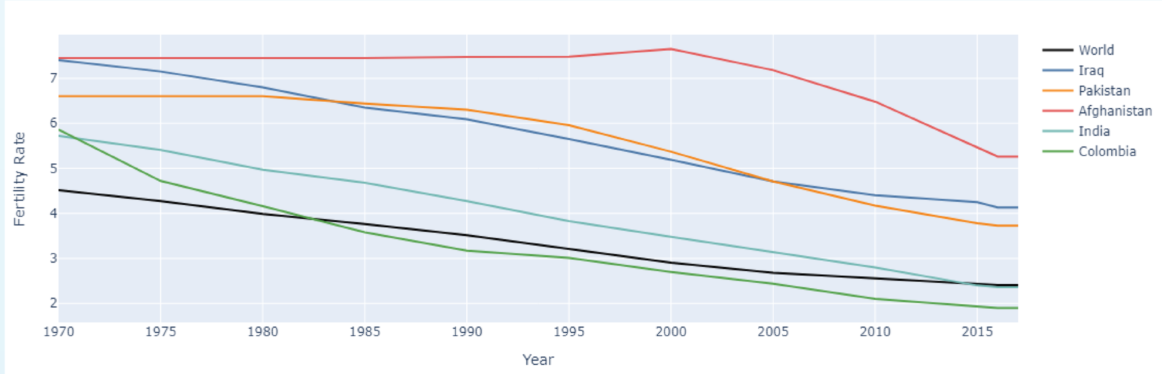


Fig 5.a: GDP of the terrorist-prone countries. Fig 5.b: Fertility rate of the terrorist-prone countries

In general, the terrorist-prone countries were the ones with their GDPs well below the global average and their fertility rates well above the global average.

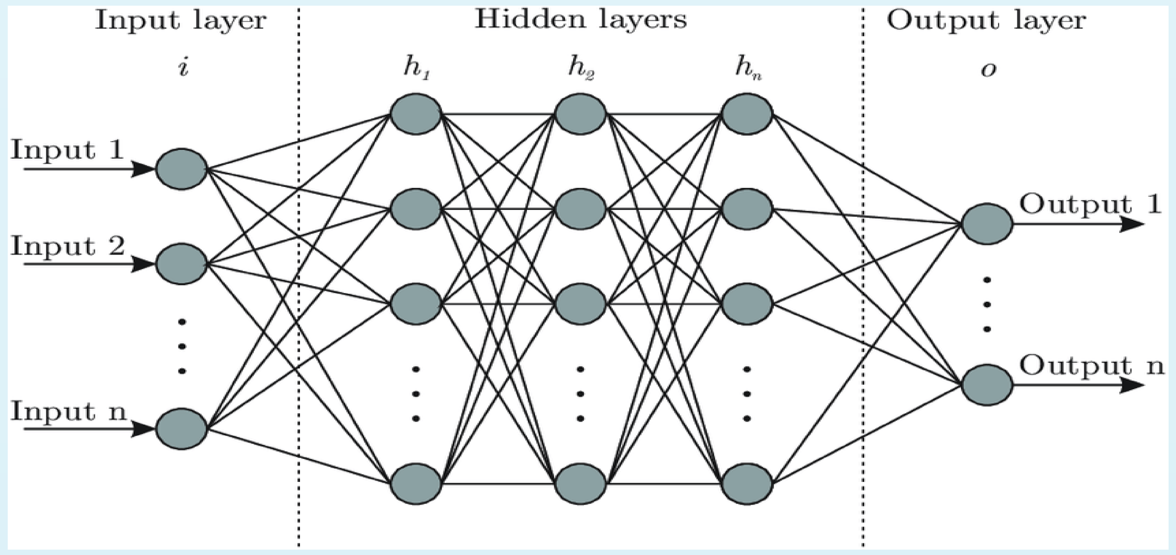


Fig 6: A simple Feed Forward Neural Network

Four different neural networks (Feed Forward NN, Bidirectional LSTM, Convolutional NN, and Gated Recurrent Units) and three different machine learning models (Random Forest, K Neighbors and Decision Trees) were trained to assess their efficiencies in predicting the number of casualties.

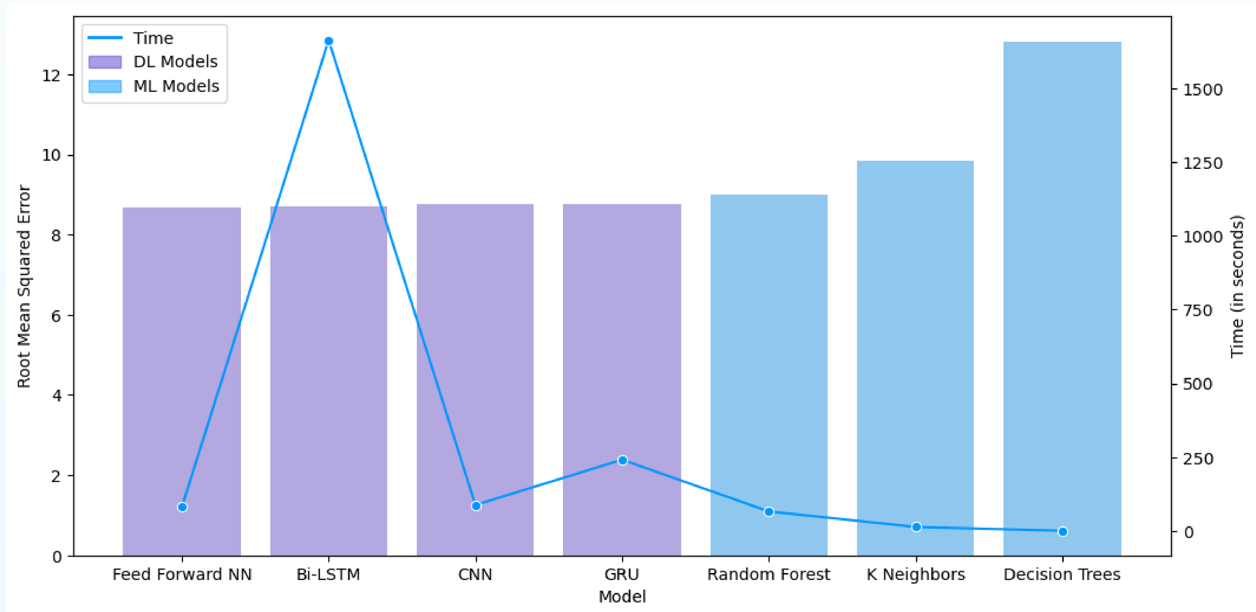


Fig 7: RMSE and time taken for different models

Feed Forward Neural Network turned out to be the most efficient model, achieving a RMSE score of 8.68 and KNN was the fastest model, completing the training and testing in 13.7 seconds. In general, neural networks had lower RMSE scores than machine learning models but they also took longer in prediction than their machine learning counterparts.

Future Works

The deep learning and machine learning models at their present states are far from efficient. In the future, the models will be tuned for their hyperparameters and trained on a larger subset of the main dataset to achieve the lowest RMSE score.

References

[1] Countries in the world by population (2023). Worldometer. Retrieved February 5, 2023,
[2] Lai, N. T. C. (2023, February 3). Word population (1955-2020). Kaggle. Retrieved February 5, 2023, from https://www.kaggle.com/datasets/nguyenthicamlai/population-2022
[3] Mishinev, T. (2022, September 9). World, region, country GDP/GDP per capita. Kaggle. Retrieved February 5, 2023, from https://www.kaggle.com/datasets/tmishinev/world-country-gdp-19602021
[4] National Consortium for the Study of Terrorism and Responses to Terrorism. Global terrorism database. Kaggle. Retrieved February 5, 2023, from https://www.kaggle.com/datasets/START-UMD/gtd
[5] World Bank. GDP (current US$). GDP National Accounts. Retrieved February 5, 2023, from https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

Scan the QR to view the animations