ANALYSIS OF GLOBAL TERRORISM

Kushagra Singh
PES1UG20CS657
Dept. of Computer Science PES University Bangalore, India
skushagra11@gmail.com

Pranav R
PES1UG20CS673
Dept. of Computer Science PES University Bangalore, India
pranavr25022003@gmail.com

Sujith Bonthala
PES1UG20CS694
Dept. of Computer Science PES University Bangalore, India
sujithbonthala7@gmail.com

***Abstract* — Terrorist attacks have developed into a significant source of risk that affects the security of the global society. This paper tends to analyze the Global Terrorism dataset to accurately put forth a number of facets of terrorism, from the kind of weapons used to the regions where terrorism is most common and many aspects. The dataset was thoroughly examined utilizing a range of tools and methods. The analysis provided as a part of this research could help governments and non-governmental organizations fight terrorism more effectively.**

***Keywords* — terrorist attack analysis, quantitative analysis, global terrorism database**

I. INTRODUCTION

Terrorism is a complex Political and Social phenomenon that has existed from time immemorial. Terrorist attacks are obviously destructive and violent and seemed to have worsened with the advancement of humankind. They are currently one of the biggest threats to the peace and security of the world community and a source of widespread worry. Terrorist attacks can be defined as involving illegal or criminal violence, threats, sabotage activities conducted by non-governmental organizations to achieve Political, Economic, Religious, Social goals through threats, coercion, intimidation . Attacks by terrorists frequently target civilians, putting their lives and property at peril while also disrupting social order and stability. Attacks by terrorists, which are unorthodox emergencies, frequently result in a large number of fatalities and have a significant social impact. [1]

Targets most at risk from terrorist attacks are civilians, and as a result, the risk index for these targets is significantly higher than it is for other targets. The findings of this study can assist individuals and the government in enabling a better understanding of terrorism, improving awareness to prevent terrorism, and enhancing emergency management and rescue, as well as providing a strong and reliable basis and reference for joint counterterrorism in various countries.

We give an in-depth examination of the Global Terrorism Dataset, which is a component of the Global Terrorism Database, as part of our research. For a clearer understanding of the analysis, numerous graphs have been employed on various components of it. Over the course of this research, we have analyzed a variety of features, including the frequency of attacks in a nation or region, the most often used weaponry, terrorist targets, and many others.

II. RELATED WORK

In this section, we summarize the existing work that has already been done by using the dataset for various purposes, especially the aspect that is relevant to our work. The summary of the approach used and the results obtained have been mentioned.

Xiaohui Pan [1] : In this study, through a quantitative analysis of the data in the Global Terrorism Dataset, ensemble machine learning has been used to construct five multiclass classification models for the prediction of terrorist organizations that perpetrated terrorist attacks.

The features and tendencies of 32 terrorist organizations with more than 500 terrorist acts were discussed in detail after the terrorist organizations were assessed according to the frequency of their assaults. Then, based on the feature selection approach, 36 feature attributes were chosen for the prediction of terrorist groups in terrorist acts, and five classifiers—decision tree, bagging, random forest, extra tree, and XGBoost—were built to accomplish so. The hold-out and 10-fold cross-validation procedures were used, respectively, to assess the performance and stability of the five models. For high-frequency activities in terrorist attacks, the models projected 32 terrorist organizations.

Finally, the experimental findings demonstrated that the five models had good stability and performance. The best prediction accuracies were attained by XGBoost and the random forest classifier, with 97.15% and 97.03%, respectively. They used the confusion matrix to better visualize and analyze the XGBoost model's prediction outcomes. The classification prediction accuracy of the random forest algorithms was consistently very good when taking into account the number of terrorist organization classifications based on the frequency of attacks. When there were few terrorist organizations (a few dozen, for example), XGBoost showed the highest level of prediction accuracy, and random forest performed almost as well.

The prediction model discussed here is capable of macroscopically predicting the terrorist groups responsible for major international terrorist attacks, uncovering the contributing causes of such attacks, and offering guidance on how antiterrorism organizations and related nations should prevent and control terrorism. We believe that these technologies can aid security departments in locating better algorithmic models and suitable datasets to increase the accuracy of predictions relating to terrorist attacks. This is contingent upon further advancements in the performance and accuracy of machine learning algorithms. Even with the ongoing development of machine learning, research on large-scale monitoring and prediction algorithms is still anticipated to be difficult given the local rarity of terrorist acts and their adaptability in planning and execution.

Zhongbei Li et.al [2] : The K-means cluster analysis method and the AHP (Analytic Hierarchy Process) were used to analyze the terrorist attacks from both qualitative and quantitative perspectives based on the GTD. A thorough analysis was done of the terrorist attack targets and major geographic areas throughout the previous three years.

The AHP was used to categorize the severity of the damage caused by terrorist attacks. Certain variables with overlapping data and complex relationships were attributed to a small number of unrelated, all-encompassing factors, and the grading standards for terrorist attacks were established. The indicator of the number of casualties had the biggest influence on the grade of terrorist attacks, followed by the indicators of property loss. The top 10 terrorist assaults that caused the most damage during the last 20 years were highlighted.

In order to categorise terrorists by geography, attack type, target type, and weapon type, the K-means cluster analysis method was employed. The top five suspects in the 2015 terrorist attacks, the identity of whose perpetrators was still unknown, were revealed.

Terrorist assaults are demonstrating a growing pattern in which the great majority of attacks and fatalities have occurred in a small number of locations. The distribution map of major terrorist attack locations indicated that the Middle East, Southeast Asia, Central Asia, and Africa would likely experience the greatest number of future worldwide terrorist incidents. The severity of terrorist attacks in Southeast Asia is likely to increase, and terrorist strikes in Africa are likely to expand in scale. It exhibits a propensity to develop into a fresh hotspot for terrorist assaults.

Civilians are the group most at risk from terrorist attacks, and their risk index is significantly greater than that of other targets, according to an examination of the targets of terrorist attacks during the last three years. Decision-makers ought to pay attention to this and enhance the system for responding to terrorist strikes.

III. DATASET

Global Terrorism Database (GTD) [3], Rand Database of Worldwide Terrorism Incidents (RDWTI), "International Terrorism: Attributes of Terrorist events" Dataset (iterate), World Incident Tracking System (WITS), and "Terrorism in Western Europe: events database" are the most prominent currently available International Terrorism Databases (tweed). The GTD, an open-source database of terrorist incidents, is the largest of them; it contains details on terrorist attacks globally from 1970 to 2020. The GTD already contains more than 200,000 episodes and, unlike many other event databases, incorporates comprehensive data relevant to Transnational and International Terrorist incidents that have happened during the chosen time period. The Terrorist attack data considered for this study Were obtained from the GTD (http://www.Start.Umd.Edu/gtd/, accessed on 30 April 2020). We have taken a subsection of the enormous dataset to do further analysis and visualization.

Exploratory Data Analysis and Visualizations done

The dataset has a total of 135 columns with 55 of them being float64, 22 being int64, and 58 of them belonging to the object category.

Upon data processing, it is observed that a lot of columns have many null values which led to the selection of 28 columns from 135 based on the above factor and their importance. Column names have been renamed to improve their readability during analysis.

On repeating the above step, it is seen that there are a few columns with NA entries that are replaced with appropriate values for some columns.

It is then observed that there are 204 countries and 12 regions in the dataset. A total of 3725 gangs have been observed using a certain set of weapons and targeting a specific group of people. But, there has been a noticeable reduction in the number of gangs from before the year 2000 (2337) to after the year 2000 (1686). Iraq is the most affected country due to terrorism after 2000

The above statement can be inferred from the below graph.

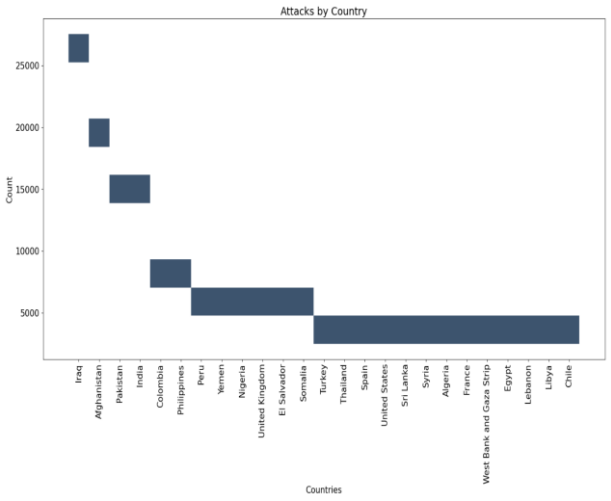

Fig. 1. Country vs attack count plot

Statistics tell that before the year 2000, South American and Europe regions contributed the most towards Global Terrorism, whereas, after 2000, the Middle East, North Africa, and South Asia experienced prominent terrorism.

Bombing/Explosions and Armed Assault are the most preferred attacking modes used by terrorists. This trend has been observed before and after the year 2000. Private citizens and property, the Military, Police, and the Government are the most targeted types. Explosives and Firearms are the weapons most used by terrorists.

Most of the terrorist acts haven't been categorized into gangs maybe because it was unknown. From the terrorism, acts have been assigned to a gang. Taliban has dominated other gangs in the number of crimes committed.

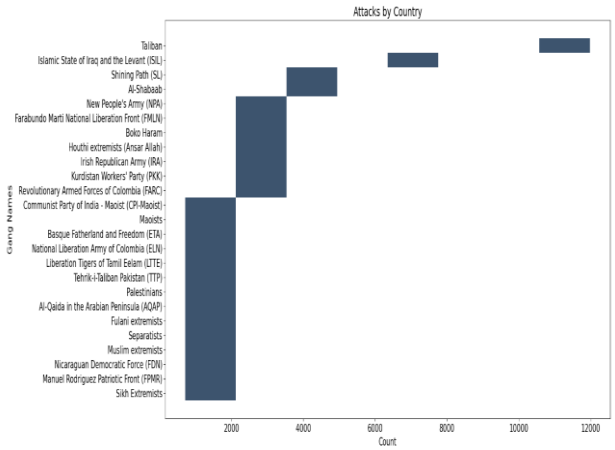

Fig. 2. Gang names vs attack count plot

The statistics tell that 2014 was the year with the most Terrorism. Since 2014, terrorism acts have been reduced. With almost the same number of attacks every month, it is observed that the number of attacks is high on the 1st and 15th of every month.

From the below 2 graphs, it is to be inferred that there has been a similar trend in the locations of terrorism acts before and after the tear 2000. The only difference between the below two World plots is that the acts have reduced in the South American region and increased in the North African region to a certain extent.

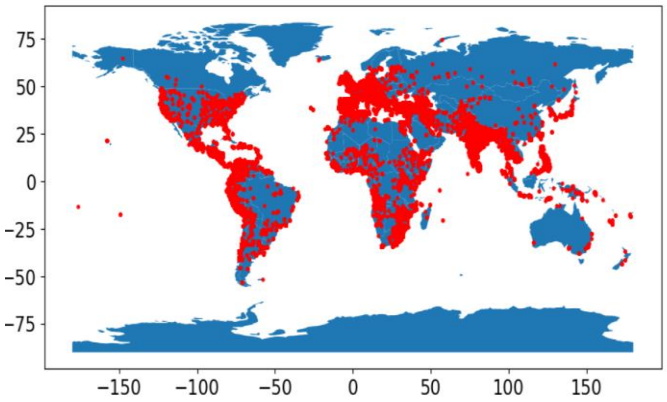

Fig. 3. Terrorism locations before 2000

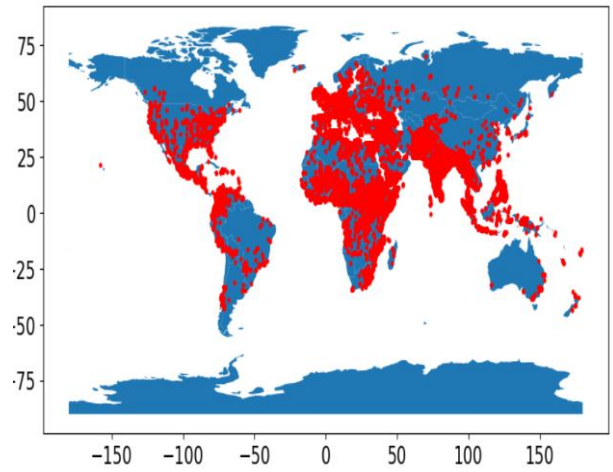

Fig. 4. Terrorism locations after 2000

Among the terrorist attacks, 14.59% of incidents are part of multiple terrorist acts. 56.66% of incidents resulted in destruction/damage to property. 3.55% of incidents were suicidal terrorist acts.

Out of all the terrorist acts, 927 have been categorized to have major damage to properties (likely > $1 million) and 6 have been categorized as catastrophic (likely > $1 billion).

IV) PROPOSED WORK

The report so far presents the analysis of terrorism and its impact on a global scale which could be further extended and improvised by finding the impact of different attributes by feature selection using the SelectKBest method and utilizing models like K-means clustering. As the research progresses we would be including a more extensive and quantitate analysis of the enormous database which would have a major impact in preventing such activities in the future.

The problem statement focuses on highlighting the numerous factors involved both before and after the terrorist attack occurred. We could make important inferences about the nature of global terrorism by comparing the numerous outcomes obtained as part of the investigation.

REFERENCES

[1] J Xiaohui Pan, "Quantitative Analysis and Prediction of Global Terrorist Attacks Based on Machine Learning", Scientific Programming, vol. 2021, Article ID 7890923, 15 pages, 2021. https://doi.org/10.1155/2021/7890923.

[2] Li, Zhongbei, Xiangchun Li, Chen Dong, Fanfan Guo, Fan Zhang, and Qi Zhang. 2021. "Quantitative Analysis of Global Terrorist Attacks Based on the Global Terrorism Database" *Sustainability* 13, no. 14: 7598. https://doi.org/10.3390/su13147598

[3] Li, G. LaFree and L. Dugan “Introducing the Global Terrorism Database , “ *Terrorism and the political violence*, vol. 19, no. 2, pp. 181204, 207. https://doi.org/10.1080/09546550701246817

[4] Hu, Xueli, Fujun Lai, Gufan Chen, Rongcheng Zou, and Qingxiang Feng. 2019. "Quantitative Research on Global Terrorist Attacks and Terrorist Attack Classification" *Sustainability* 11, no. 5: 1487. https://doi.org/10.3390/su11051487