**Literature Review:**

**Introduction:**

Counterterrorism efforts are crucial for establishing peace and stability in Afghanistan and bringing the perpetrators to justice. The country has been subjected to terrorist attacks for a long time, especially over the last two decades. Predictive modeling using machine learning (ML) offers a promising approach to anticipate and prevent terrorist attacks. By leveraging historical data and socio-economic factors, ML models can provide actionable insights, aiding in proactive security measures. It is important to monitor the related work on using machine learning in counterterrorism area to explore methodologies and key findings and gaps to establish the foundation for developing a predictive model tailored to Afghanistan.

**Related work:**

Defining terrorism is not a straightforward matter. There is no single internationally accepted definition of what constitutes terrorism, and the terrorism literature abounds with competing definitions and typologies. IEP accepts the terminology and definitions agreed to by TerrorismTracker. The GTI therefore defines terrorism as ‘the systematic threat or use of violence, by non-state actors, whether for or in opposition to established authority, with the intention of communicating a political, religious or ideological message to a group larger than the victim group, by generating fear and so altering (or attempting to alter) the behavior of the larger group.’ [1].

The number of countries experiencing at least one death from terrorism in the past year was 44, a slight increase compared with the 43 countries in 2020. Another 105 countries had no deaths or attacks from terrorism in 2021. This is the highest number of countries since 2007 [1]. According to the Global Terrorism Index (GTI) report of 2022 [1], Afghanistan was ranked first among countries with a very high impact of terrorism, scoring 9109 points. This ranking underscore the severe and pervasive impact of terrorist activities within the country. The GTI report evaluates countries based on various indicators, including the number of terrorist incidents, fatalities, injuries, and property damage. Afghanistan's high score reflects the significant and ongoing challenges it faces in combating terrorism and maintaining national security.

Terrorism can be claimed by known terrorist groups, affiliated groups, or individual terrorists. Terrorist attacks can have an enormous impact on wide sections of society [2]. Prediction of terrorist groups after an attack is one of the most important steps for counter terrorism. As soon as we are able to find the involved group name, we will be able to make strategies to catch the culprits [3].

Traditional methods for predicting the names of terrorist groups, such as email tracking, telephone signal analysis, and social network analytics, rely heavily on manual processes and have become less effective due to the evolving and dynamic nature of terrorist organizations and their activities. Given the importance of accurate predictions, there is a need for more advanced and reliable techniques that can handle the complexities associated with terrorist actions. Pattern recognition and machine learning, which have proven successful in other predictive tasks, offer promising solutions for predicting the names of terrorist groups [4], [5].

Pan and Zhang [6], presents a classification model framework for predicting terrorist attack targets. The model, utilizing data from the Global Terrorism Database (GTD) and employing machine learning algorithms such as SVM, decision tree, KNN, and neural networks, shows high performance in prediction tasks. Neural networks performed best, and future work will explore deep learning algorithms to predict other elements of terrorist attacks.

Enrique Lee Huamaní, Alva Mantari Alicia and Avid Roman-Gonzalez [7], used Global Terrorism Database and machine learning techniques to visualize and predict terrorist Attacks worldwide. The study demonstrates the effectiveness of machine learning techniques in visualizing and predicting terrorist attacks. The historical data from the GTD enabled the models to identify regions and types of attacks with high accuracy. The Middle East & North Africa and South Asia are the regions with the highest probabilities of attacks, primarily involving bombs and explosions. Decision Trees and Random Forest models showed assertiveness percentages ranging from 75.45% to 90.414%, validating the utility of these techniques for global event research​.

**Summary and Synthesis:**

* Global Terrorism Index 2022: Measuring the Impact of Terrorism [1]: This report by the Institute for Economics & Peace provides data on global terrorism and measuring the impact of terrorism worldwide, analyzing the impact of terrorism in various countries. The report likely uses data from terrorism tracker and other sources to measure terrorism-related deaths, attacks, and economic impact. Offers a resource to understand the current state of global terrorism and how it affects different regions.
* Global Terrorism Index 2022: Measuring the Impact of Terrorism [2]: Like the 2022 report, this provides data on global terrorism and measuring the impact of terrorism worldwide for an earlier year (2018).
* Prediction of Unknown Terrorist Group Names Responsible for Attacks inTurkey [3]:This paper focuses on developing methods to predict the names of terrorist groups responsible for attacks in Turkey. The authors likely employ machine learning techniques to analyze data on past attacks and identify patterns that can be used for prediction. Advances the field of terrorism analysis by developing methods to identify unknown terrorist group names.
* [4], [5], [6], [7]: These papers all explore the application of machine learning for various purposes related to terrorism analysis, including predicting terrorist group formation [5], identifying targets of terrorist attacks [6], and visualizing patterns of attacks worldwide [7]. These papers likely employ different machine learning algorithms depending on the specific task such as prediction, classification, visualization. Collectively, these papers demonstrate the potential of machine learning for various aspects of terrorism analysis, contributing to improved prediction and prevention efforts.

Papers [1] and [2] share a focus on data collection and analysis to understand global terrorism trends. Papers [3-7] all utilize machine learning for various tasks related to terrorism analysis.

While paper [3] present research on developing new methods using machine learning to predict unknown terrorist group names. Paper [6] tries to build a model for predicting targets of terrorist attacks. Paper [3] focuses on a specific country, while paper [7] tries to concentrate on the issue in a wide range.

Overall, these papers highlight the complementary nature of data analysis and machine learning in the field of terrorism research, but the specific applications of machine learning differ between papers. While reports like the Global Terrorism Index provide a foundational understanding of the current landscape, machine learning offers promising tools for prediction, identification, and potentially preventing future terrorist activities.

**Conclusion:**

The literature highlights the critical role of machine learning in enhancing counterterrorism efforts by providing predictive insights into terrorist activities. Traditional methods have proven less effective due to the dynamic nature of terrorist organizations, necessitating the adoption of more advanced techniques like pattern recognition and machine learning. The previous studies reviewed demonstrate the potential of these techniques in predicting terrorist group names, attack targets, and visualization.

Bye using the existing knowledge and methodologies with adding new perspective we can develop predictive models tailored to Afghanistan. That can help us provide actionable insights that aid in the prevention of terrorist attacks, enhance national security, and bring perpetrators to justice.

**References:**

[1] Institute for Economics & Peace. *Global Terrorism Index 2022: Measuring the Impact of Terrorism*, Sydney, March 2022.

[2] Institute for Economics & Peace, *Global terrorism index 2018. Measuring the impact of terrorism*. Sydney, Australia,2018.

[3] I. A. Fadel and C. Öz, *“Prediction of Unknown Terrorist Group Names Responsible for Attacks in Turkey”,* SAUCIS, vol. 5, no. 3, pp. 257–268, 2022, doi: 10.35377/saucis...879855.

[4] J. Feng, H. Xu, S. Mannor and S. Yan, “Robust Logistic Regression and Classification,” *Proc. - 27th Inter. Conf. on Neural Info. Processing Syst. (NIPS)*, Montréal, Canada, 08-13 December 2014.

[5] F. Gohar, W. Haider and U. Qamar, “Terrorist Group Prediction Using Data Classification,” *Proc. – Inter. Conf. on Artificial Intelligence and Pattern Recognition*, Kuala Lumpur, Malaysia, 17-19 November 2014.

[6] Xiaohui Pan and Tao Zhang, *“Machine learning-based target prediction for terrorist attacks”* 2023 J. Phys.: Conf. Ser. 2577 012007, **doi:** 10.1088/1742-6596/2577/1/012007

[7] Enrique Lee Huamaní, Alva Mantari Alicia and Avid Roman-Gonzalez, “*Machine Learning Techniques to Visualize and Predict Terrorist Attacks Worldwide using the Global Terrorism Database*” International Journal of Advanced Computer Science and Applications (IJACSA), 11(4), 2020. doi:[10.14569/IJACSA.2020.0110474](https://dx.doi.org/10.14569/IJACSA.2020.0110474)

**Data Research:**

**Introduction**

The Global Terrorism Database™ (GTD) serves as a crucial resource for understanding patterns and trends in terrorist activities globally from 1970 to 2020. This extensive open-source database, managed by the National Consortium for the Study of Terrorism and Responses to Terrorism (START), is unparalleled in its scope and depth, covering over 200,000 terrorist incidents. The significance of this research lies in its potential to inform policymaking, enhance counter-terrorism strategies, and contribute to academic and practical understandings of terrorism. Thorough exploration of this data is necessary to uncover insights into the evolution, methods, and impacts of terrorist activities, thus aiding in the development of more effective responses and preventive measures.

**Organization**

The data research findings are organized thematically to best address the multifaceted nature of the data. The themes include:

* Data Collection and Sources
* Characteristics and Scope of the GTD
* Key Insights and Trends from Data Analysis
* Implications for Counter-Terrorism Strategies

**Data Description**

The primary dataset utilized in this project is the Global Terrorism Database™ (GTD). This comprehensive database includes detailed information on terrorist incidents worldwide, covering the period from 1970 to 2020. The data source comprises over 200,000 entries, each documented with variables such as the date and location of the incident, the weapons used, the nature of the target, casualties, and, when identifiable, the responsible group or individual.

**Data Source and Format:**

Source: US National Consortium for the Study of Terrorism and Responses to Terrorism (START)

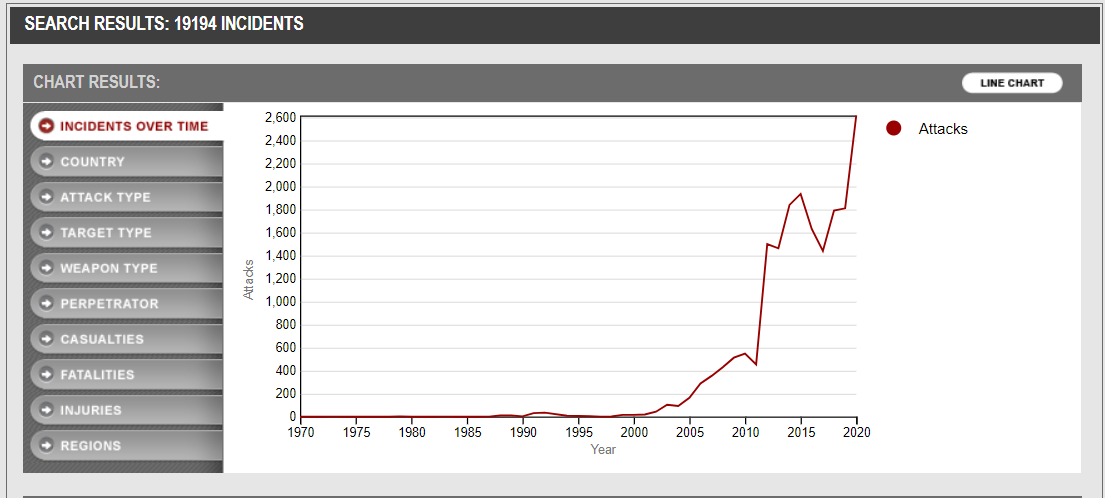
Format: The data is available in various formats, including CSV and SQL, to facilitate ease of analysis.

Size: The database includes information on over 200,000 incidents, derived from more than 4,000,000 news articles and 25,000 news sources from 1998 to 2017 alone. So, the size of the database is huge, but I prefer to use only the data which is related to Afghanistan for my project

I selected the GTD due to its comprehensive coverage and the reliability of its sources, which ensures a robust foundation for my project. This data is directly relevant to my project as it provides extensive insights into terrorist activities, crucial for understanding and developing counter-terrorism measures and models.

**Data Analysis and Insights**

Temporal Trends: There has been a significant fluctuation in the number of terrorist incidents over the decades, with notable peaks during periods of intense geopolitical conflicts.



Incidents over time in Afghanistan

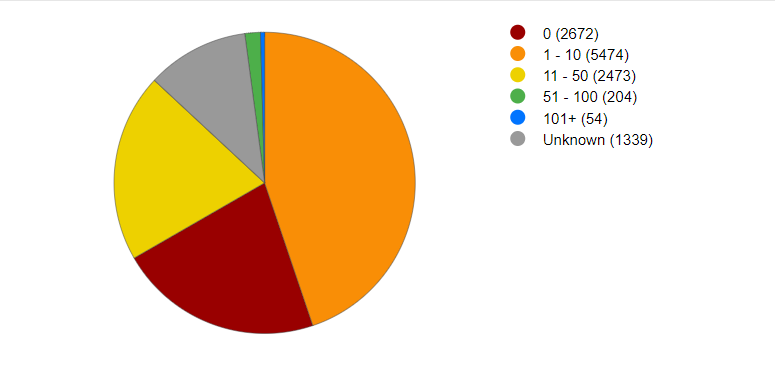
Geographical Distribution: Certain regions, particularly the Middle East and South Asia, have experienced a higher frequency of terrorist activities compared to other parts of the world. According to the database Afghanistan is among the countries which experienced/experiences very high frequency of terrorist activities.

Methods and Targets: Bombings are the most common method of attack, followed by armed assaults and kidnappings. Government buildings, military targets, and civilian populations are frequently targeted.

Casualties: The analysis indicates a varying number of casualties per incident, with some high-profile attacks resulting in significant loss of life and others causing minimal casualties. In the figure below the shows the casualties of attacks took place in Afghanistan.

* 0 casualty (26.72%)
* 1-10 casualties (54.74%)
* 11-50 casualties (24.73%)
* 51-100 casualties (2.04%)
* 101+ casualties (0.54%)
* Unknown casualties (13.39%)

The largest portion of the pie chart (54.74%) consists of attacks which had 1-10 casualties.

casualties of terrorist attacks

**Conclusion**

The key findings from data analysis of the GTD underscore the dynamic and complex nature of global terrorism. The research highlights significant temporal and geographical variations in terrorist activities, as well as the diverse methods and targets employed by terrorist groups. These insights are critical for informing counter-terrorism strategies and policies, emphasizing the need for adaptive and region-specific responses. The thorough exploration of the GTD not only enhances our understanding of past and present terrorist activities but also aids in anticipating and mitigating future threats.

**References**

[1] National Consortium for the Study of Terrorism and Responses to Terrorism (START). (2020). Global Terrorism Database (GTD). https://www.start.umd.edu/gtd

[2] LaFree, G., & Dugan, L. (2007). Introducing the Global Terrorism Database. \*Terrorism and Political Violence, 19\*(2), 181-204.

[3] START. (2021). Global Terrorism Database Codebook. https://www.start.umd.edu/gtd/downloads/Codebook.pdf

**Technology Review:**

**Introduction**

In this technology review, I am trying to explore the key technologies and tools planned to use in our project analyzing data from the Global Terrorism Database (GTD) and building a predictive model for Afghanistan. The importance of this review lies in identifying and understanding the most effective methods and tools for processing, analyzing, visualizing, and implementing complex datasets for building the model. The selected technologies—Support Vector Machines (SVM), Decision Trees, k-Nearest Neighbors (KNN), Deep Neural Networks (DNN), Power BI, TensorFlow, Keras, and Jupyter Notebook and other technologies that will be used in this project are highly relevant to our project goals of uncovering patterns and insights in GTD to build our model.

**Support Vector Machines (SVM)**

The purpose of using SVM is to do classification and regression analysis. It is effective in high-dimensional spaces, versatile kernel functions, robust against overfitting. Which is widely used in text classification, image recognition, and bioinformatics.

**Decision Trees**

I will use decision trees for decision making and classification in this project. Which is highly effective, handles both numerical and categorical data, and can capture non-linear relationships.it is mostly used in finance, healthcare, and marketing for predictive modeling.

**k-Nearest Neighbors (KNN)**

The same as SVM it is also used for classification and regression. It has a simple algorithm, instance-based learning and requires no training phase. It is used in recommendation systems, image recognition, and anomaly detection.

**Deep Neural Networks (DNN)**

It will be used in complex pattern recognition in this project which has multiple layers of neurons, capability to model complex non-linear relationships and can have high accuracy with large datasets.

**Power BI**

I will use Power Bi for data visualization and business intelligence. The interactive

**Python libraries, TensorFlow and Keras**

Various Extensive libraries Pandas, NumPy, Scikit-learn etc. Will be used in this project.

Machine learning and deep learning frameworks Tensorflow and Keras will be used particularly for the deep learning part of the project.

SVM, Decision Trees, KNN, and DNN: These algorithms provide diverse approaches to classification and pattern recognition, essential for building the model.

Power BI: Facilitates the visualization of data trends and patterns, making it easier to communicate findings.

Python, TensorFlow, Keras, Jupyter Notebook: Provide a comprehensive ecosystem for data analysis, machine learning model development, and experimentation.

**Conclusion**

This technology review highlights the importance and suitability of selected tools and algorithms for the project. The combination of machine learning algorithms, data visualization tools, and programming environments ensures comprehensive data analysis, insightful visualizations, and model building. These technologies will significantly benefit my project by providing robust methodologies for uncovering patterns, trends, and actionable insights in GTD.

**References**

[1] Abadi, M., et al. (2016). TensorFlow: A System for Large-Scale Machine Learning. OSDI, 265-283.

[2] "User Guide." scikit-learn. <https://scikit-learn.org/stable/user_guide.html>.

[3] Géron, Aurélien. “*Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*”. Sebastopol, CA: O'Reilly Media, Incorporated, 2019.