Global Journal of Computer Sciences: Theory and Research
ISSN 2301-2587
Volume 14, Issue 2, (2024) 30-48

www.gjcs.eu

Data science insights and the classification of terrorist attacks in Nigeria using machine learning techniques

Ahmad Mustapha Bello[a], American University of Nigeria, 98 Lamido Zubairu Way, Yola Township bypass, PMB 2250, Yola, Adamawa State, Nigeria, ahmad.mustapha@aun.edu.ng
Aamo Iorliam[b], American University of Nigeria, 98 Lamido Zubairu Way, Yola Township bypass, PMB 2250, Yola, Adamawa State, Nigeria, aamoiorliam@gmail.com
Ozcan Asilkan[c1], Higher Colleges of Technology, United Arab Emirates. oasilkan@hct.ac.ae

Suggested Citation:
Bello A. M., Iorliam, A. & Asilkan 0. (2024). Data science insights and the classification of terrorist attacks in Nigeria using machine learning techniques. *Global Journal of Computer Sciences: Theory and Research*. *14*(2), 30-48. https://doi.org/10.18844/gjcs.v14i2.9669

Received from; March 2, 2024, revised from; June 22, 2024 and accepted from October 2.
Selection and peer review under the responsibility of Assist. Prof. Dr. Ezgi Pelin YILDIZ, Kars Kafkas University, Department of Computer Technology, Turkey.

Abstract

Terrorism has become a critical concern, with extremist groups orchestrating widespread violence resulting in casualties, displacement, and societal instability. Despite growing interest in counterterrorism, limited research has applied data science to classify and interpret patterns of terrorist activity. This study addresses this gap by employing data driven approaches to analyze terrorism in Nigeria. Utilizing data from the Global Terrorism Database and the Armed Conflict Location and Event Data Project, the study conducts exploratory data analysis to uncover patterns and trends in terrorist incidents. The analysis, performed using Power Business Intelligence, reveals key insights into the distribution, characteristics, and relationships within the datasets. Machine learning classifiers including Decision Tree, Random Forest, and Logistic Regression are trained and evaluated using metrics such as accuracy, precision, and recall. Among the models tested, Random Forest demonstrated the highest accuracy. The findings reveal a consistent pattern of successful attacks and underscore the potential of data science techniques in understanding and predicting terrorist activities. This research highlights the value of analytical methodologies in supporting informed decision making and developing strategic interventions for counterterrorism efforts.

Keywords: Armed conflict location; decision tree; event data project; machine learning classifiers; random forest.

* ADDRESS FOR CORRESPONDENCE: Özcan Asilkan, Business Analytics Department Higher Colleges of Technology, United Arab Emirates.
E-mail address: oasilkan@hct.ac.ae

Bello A. M., Iorliam, A. & Asilkan 0. (2024). Data science insights and the classification of terrorist attacks in Nigeria using machine learning techniques. *Global Journal of Computer Sciences: Theory and Research*. *14*(2), 30-48. https://doi.org/10.18844/gjcs.v14i2.9669

1. INTRODUCTION

Terrorism is a persistent global issue and has been driven by various motivations, including religious, political, and social factors. Terrorist groups often employ violence to achieve their objectives, resulting in significant societal impacts, including human casualties and economic costs [1]. The 20th century witnessed the global proliferation of terrorism, primarily fueled by rising nationalism and religious extremism, leading to the emergence of groups like Al-Qaeda, ISIS, and Boko Haram [2]. These organizations executed high-profile attacks, including the notorious 9/11 incidents, the Paris attacks, and the Manchester Arena bombing, amongst several others. Various factors contribute to terrorism, encompassing political instability, economic disparities, and social injustices. Yet, the root causes of terrorism remain intricate and multifaceted [1,2].

Terrorism has witnessed significant global growth over the years, negatively impacting economies and populations worldwide. Notably, Afghanistan remained the country most affected by terrorism for the fourth consecutive year worldwide. In 2022, the Islamic State (IS) and its affiliates ranked as the deadliest terrorist organizations worldwide, followed by Jamaat Nusrat al-Islam wal Muslimeen (JNIM), Al-Shabaab, and the Balochistan Liberation Army (BLA). The proliferation of terrorism has been particularly conspicuous in Africa. The Sahel region of Sub-Saharan Africa accounted for more terrorist-related deaths than South Asia, the Middle East, and North Africa combined in 2022, establishing itself as the epicenter of terrorism [3]. The Sahel's share of global terrorism deaths surged from 1% in 2007 to 43% in 2022. Burkina Faso and Mali contributed significantly to the Sahel terrorism deaths, with unidentified jihadists attributed to the majority of attacks in 2022. Both countries witnessed a significant increase in terrorism-related fatalities, with Burkina Faso and Mali recording a 73% number of deaths due to terrorism-related deaths [3, 4].

In Nigeria, the Global Terrorism Index reported a decline in terrorism-related fatalities in 2022 when compared to the 2015 peak of terrorist attacks [3; 5]. Furthermore, Nigeria dropped to 8th position globally in the 2023 Global Terrorism Index with a score of 8.065. The country faces an array of security challenges, including the enduring Boko Haram insurgency in the North East region, prolonged discontent and militancy in the Niger Delta region, escalating conflicts between pastoralists and farming communities in the Middle Belt region, and the Biafra agitation in the South East, Nigeria [3]. The violence has caused over two million people to flee their houses, primarily due to the Boko Haram insurgency, resulting in a severe humanitarian crisis [6].

Terrorism remains a critical global concern, with Nigeria facing significant challenges in combating this threat [7]. Osadola and Emah [8] pointed out the fact that Nigeria shares boundaries with Cameroun, Chad, Benin, Niger, and Guinea. As such, solving terrorism issues in Nigeria should also be the responsibility of neighboring countries as well. Again, Onana Ibogo [9] emphasized that the Boko Haram insurgency is not only in Nigeria but also neighboring countries such as Niger, Chad, and the Cameroon Republic. They stated that the governments, non-governmental organizations, and other international countries should come together in solving this terrible menace. Furthermore, Jackson [10] also pointed out that the Boko Haram terrorist group has grown tremendously in Nigeria due to Nigeria's delay in brutally curbing this group on time. The author stated that the Multinational Joint Task Force (MNJTF), which was established in 1994, is tasked with the responsibility of curbing terrorists and their activities in affected countries such as Chad, Niger, Nigeria, and Cameroon, and the MNJTF is strongly supported by the African Union (AU).

Over the years, researchers have proposed machine learning approaches to studying terrorist attacks. For example, Ogundunmade and Adepoju [11] utilized the Bayesian Neural Network to predict terrorists' attack nature. They showed that the hyperbolic tangent activation function of the Bayesian Neural Network achieved the best terrorist attack prediction accuracy for the train and test dataset they considered for their experiment. Idakwo et al., [12] utilized the Capsule Network to develop an improved technique for the detection and classification of weapons used by terrorists. They used the gun detection datasets from González et al., [13] and achieved 99.43%, 98.14%, 98.77%, and 98.45% for the average accuracy, precision, recall, and F1-score, respectively. Bordeanu [14] utilized cybersecurity datasets to have insights about crime patterns using techniques such as Natural Language Processing (NLP) and deep learning techniques in order

to keep under control cybercrimes. The author detected download activities that are malicious and Internet of Things botnet attacks using NLP and deep learning techniques and achieved reasonably high accuracies.

1.1. Purpose of study

Therefore, this paper first gets clear insights into terrorist activities in Nigeria using data science techniques. Furthermore, it utilizes machine learning classifiers, namely, Decision Tree, Random Forest, and Logistic Regression on the Global Terrorism Database (GTD) [15], and the Armed Conflict Location and Event Data Project (ACLED) [16] in order to classify if or not a terrorist attack in Nigeria was successful.

1.2. Related works

The persistent and multifaceted threats of terrorism have plagued the world for centuries, fuelled by diverse motivations. Terrorist organizations utilize violence to further their goals, causing substantial harm to human lives and economic stability [1]. Nigeria faces significant challenges in combating terrorism. Nigeria's terrorism landscape began evolving in the year 2000 with the emergence of notorious groups like Boko Haram. Various factors contribute to terrorism, encompassing political instability, economic disparities, and social injustices [17]. These extremist organizations have inflicted widespread devastation, causing significant loss of life, displacement of citizens, and destruction of property across the country. The relentless violence perpetrated by these groups has instilled a pervasive atmosphere of fear and insecurity among Nigerians [17].

The ongoing threat posed by terrorism remains a formidable challenge for the nation, requiring concerted efforts from both domestic and international stakeholders to address and mitigate its impact on society [18]. In the 2023 Global Terrorism Index, Nigeria is ranked number 8th in the world, dropping from 5th in 2021, with a score of 8.065 in the impact terrorism has on the country [3]. Obasi, [6] the senior advisor of the International Crisis Group, said Nigeria suffers from an increasing number of security challenges, such as the resilient Boko Haram militant insurgency in the North East, long-running discontent along with militancy in the Niger Delta region, increasing disputes between shepherds and farming communities in the Middle Belt, and Biafra agitation in the Igbo South East [6]. Over two million people have been forced to flee their homes due to violence caused by the Boko Haram insurgency. This violence has also caused a serious humanitarian crisis and given rise to vigilante self-defense groups made up of civilians, which has new policy implications and potential security issues [6].

Nigeria has faced an unprecedented surge in insecurity and terrorism since the inception of the current democratic dispensation [2]. The insecurity landscape has taken on a regional pattern, with militia groups in the South, insurgency in the North, kidnapping prevalent in the East and South, and ritual killings occurring in the East and West. Additionally, political and non-political assassinations have been carried out across the nation. This regionalization of insecurity has prompted the establishment of regional security formations aimed at addressing the escalating insecurity [19, 20].

The conflicts in the Niger Delta region, initiated in the 1990s due to the actions of various militant groups, have had adverse effects on economic development in Nigeria. These militant factions, such as the Movement for the Survival of the Ogoni People (MOSOP), Ijaw Youth Congress (IYC), Movement for the Emancipation of the Niger Delta (MEND), the Niger Delta Vigilante Force (NDVF), and the Niger Delta People's Volunteer Force (NDPVF), among others, have orchestrated destructive attacks on oil and gas facilities, Nigerian naval officers, and oil company personnel [17]. These attacks have resulted in fatalities and severe injuries. The groups have also engaged in criminal activities such as hostage-taking, kidnapping, bombings, rape, and assassinations.

Another significant security challenge confronting Nigeria is the actions of the Fulani herdsmen, who have conducted numerous attacks nationwide, causing loss of life and displacing communities [21]. Poverty, unemployment, and social grievances have been cited as contributing factors to the ongoing threat of terrorism, highlighting the need for comprehensive strategies to address the root causes of extremism [22]. The Terrorism Provision Act of 2011 has played a crucial role in providing a legal framework for prosecuting and preventing terrorism in Nigeria. Addressing the underlying drivers of extremism and enhancing implementation mechanisms are essential for effectively combating this threat in the country [18].

1.2.1. Data science and terrorism

With emphasis on data science and terrorism, its emergence in counterterrorism represents a paradigm shift in how security and intelligence agencies address terrorism's challenges. Data science, which involves data collection, analysis, and interpretation, is transforming counterterrorism [25, 26]. It plays a pivotal role in understanding and analyzing the root causes of terrorism. It integrates various data sources, identifies patterns, and employs predictive analytics [26]. This approach enables researchers to gain comprehensive insights into the factors contributing to terrorism. Data science aids in predicting potential terrorism hotspots based on historical data and high-risk indicators [27]. It also involves analyzing social networks to identify individuals susceptible to radicalization. Moreover, data science creates early warning systems that detect shifts in political, economic, or social conditions, which may increase the risk of terrorism.

1.2.2. Digital devices and threat detection

Digital devices and their interconnection to the internet have provided abundant data sources, offering insights into potential threats and the intricate dynamics of terrorist networks. The tremendous growth of terrorism has made it increasingly complex, necessitating the use of data science to uncover hidden patterns and connections [28]. Data science enables real-time analysis, allowing for swift detection and response to emerging threats. Moreover, machine learning and predictive analytics play a crucial role in anticipating potential terrorist activities and enhancing proactive measures [29].

1.2.3. Challenges of data science in counterterrorism

While data science offers significant benefits in counterterrorism, it also poses certain challenges. On the positive side, it provides early warning and threat detection by identifying suspicious activities through pattern analysis, offering early warnings. Additionally, data science aids in mapping terrorist networks, thus assisting law enforcement agencies in dealing with terrorists. It optimizes resource allocation, improving the efficiency of counterterrorism efforts. Furthermore, online data analysis can help identify individuals at risk of radicalization, enabling the development of effective counter-radicalization strategies [27].

On the other side, balancing national security with privacy concerns is a significant challenge, particularly regarding data collection and surveillance. Ensuring data accuracy is critical to avoid false positives or negatives in threat detection. However, ethical considerations must be addressed, especially when dealing with sensitive data and vulnerable populations [30, 31, 32]. Managing and interpreting vast volumes of data can be overwhelming, and it is crucial to have the necessary infrastructure and tools for effective data handling. Lastly, terrorist groups continuously adapt to counterterrorism efforts, necessitating ongoing advancements in data science techniques to stay ahead of their changing tactics [33, 34].

The proliferation of terrorist attacks, characterized by their high lethality and destructive power, leads to substantial casualties, property losses, and psychological pressure on societies. These attacks disrupt the normal order of life and work, impeding economic development and causing social unrest. To address these challenges, the analysis, insights, and prediction of terrorist attacks have become crucial components of global security governance [22]. This approach not only supports targeted counterterrorism efforts but also provides valuable information for pre-emptive measures, enabling authorities to identify new or hidden threats promptly, thereby reducing human and property losses, preventing crises, and enhancing overall societal security and stability. It is important to note that terrorist attacks, though seemingly random on the surface, often exhibit organized and premeditated characteristics [35].

1.2.4. Analyzing terrorist patterns

Analyzing the patterns and rules that guide the activities of terrorist organizations is key to making more accurate predictions and improving the efficiency of counterterrorism efforts. The Global Terrorism Database (GTD) has been a valuable resource for researchers in this regard, offering open-source data, reliable, and comprehensive. Over the past decades, academics have developed various models and algorithms for the early warning and prediction of terrorist attacks. These models incorporate machine learning techniques to include Hidden Markov Models, K-Means Clustering, and Decision Trees. Some have employed hybrid

classification frameworks, social network analysis, wavelet transform, pattern recognition approaches, and risk assessment systems [36, 37]. This analysis unveils critical insights that can inform targeted counterterrorism strategies.

1.2.5. The role of data visualization

Data visualization plays a pivotal role in the analysis of global terrorism data. The visual representation of data facilitates the comprehension of complex datasets and the identification of meaningful insights [38]. Visualization techniques are essential in exploratory data analysis (EDA), where they provide a means to uncover trends, geographic clusters, and temporal patterns in terrorist activities [39]. Geospatial visualization, in particular, allows researchers to map the geographic distribution of terrorist attacks, hotspots, and their evolution over time. Through the use of Geographic Information Systems (GIS), analysts can identify regions prone to terrorist activities and assess the effectiveness of counterterrorism measures in specific areas. Such visualizations provide actionable information for law enforcement and policy development [26]. Additionally, temporal analysis through visual representation enables researchers to understand the patterns of terrorist incidents over time. Overall, data visualization contributes to the synthesis and interpretation of terrorism data, enhancing decision-making processes and preventive measures [27].

1.2.6. Machine learning models in terrorism prediction

Machine learning models serve as invaluable tools in predicting terrorist activities globally, and in Nigeria. The analysis of extensive historical data, discerning patterns, and generating predictive insights are important machine learning techniques in curbing terrorist attacks [40]. These models conduct thorough data analysis, scrutinizing various facets of terrorist incidents such as attack types, affected regions, weaponry employed, and group dynamics. By uncovering trends and correlations within this data, machine learning models offer invaluable insights that aid in forecasting and anticipating future attacks, thereby enhancing counterterrorism efforts.

1.2.7. Recent advancements in counterterrorism systems

In the realm of counterterrorism research, recent advancements have seen the emergence of innovative systems, which harness artificial intelligence techniques to automatically identify terrorist attacks [41, 42]. Leveraging machine learning algorithms, particularly the Random Forest model, innovative systems like Terror Mine aims to discern the perpetrating groups behind such incidents by analyzing data sourced from the Global Terrorism Database (GTD). Key features such as the temporal and spatial characteristics of attacks, target specifics, and weaponry employed are scrutinized to classify and identify responsible entities accurately [42]. Notably, experimental evaluations reveal that the Random Forest model demonstrates the highest weighted F1-score among the techniques employed [42]. Furthermore, the predictive capabilities of Terror Mine extend beyond retrospective analysis, with a focus on forecasting future terrorist activities. The system utilizes advanced forecasting models like Prophet, renowned for its adeptness in capturing complex temporal patterns [42]. Prophet's additive regression framework incorporates various factors including growth, seasonality, holidays, and custom white noise error, contributing to its ability to generate accurate predictions [42]. In comparative assessments, Prophet exhibits slightly superior performance metrics, including accuracy and weighted F1-score, in contrast to the widely used Autoregressive Integrated Moving Average (ARIMA) model. This underscores Prophet's potential as a robust tool for anticipatory analysis and proactive counterterrorism measures.

1.2.8. Artificial intelligence and machine learning for terrorist attack prediction

The "Artificial Intelligence Approach for Terror Attack Prediction" paper employs artificial intelligence and machine learning methodologies for the prediction of terrorist attacks, with a primary focus on leveraging the Global Terrorism Database (GTD) and graph databases [43]. Various machine learning models were employed to categorize these incidents, yielding a commendable classification accuracy of approximately 90% [43]. The study utilized terrorism prediction approaches such as neural networks, logistic regression, support vector machines (SVM), and ensemble methods like XGBoost and Random Forest [43]. The study concludes by

emphasizing the necessity of real-world testing for these models and suggests exploring advanced deep learning techniques like recurrent neural networks and graph adversarial networks to further enhance predictive accuracy.

The study conducted by Odeniyi et al., [44] highlighted the effectiveness of advanced machine learning techniques, particularly the heterogeneous neural network model, in improving the prediction accuracy of terrorist activities. This model outperformed other conventional machine learning approaches such as logistic regression, support vector machines, k-nearest neighbors, boosting, and random forest. The research emphasized the importance of understanding the key factors that influence terrorist attacks, such as the number of perpetrators, the type of attack, weaponry used, victim demographics, and the location of incidents. Recognizing these variables allows authorities to tailor counterterrorism strategies more effectively to mitigate risks and improve security.

The findings are particularly significant when considering the high success rate of terrorist attacks in Nigeria, which stands at 91.6%, a concerning statistic that underlines the urgent need for proactive measures from law enforcement and government agencies. Utilizing classification models derived from machine learning can help authorities enhance their capacity to prevent attacks and address security challenges more effectively. Recent advancements, such as the use of Bayesian neural networks to forecast variables associated with terrorist attacks in Nigeria, further contribute to the field of predictive counterterrorism.

Ogundunmade and Adedayo [11] investigated the use of Bayesian neural networks (BNN) for forecasting key variables related to terrorist incidents in Nigeria. Their study focused on identifying patterns and predicting the nature of attacks. Various activation functions, including sigmoid, hyperbolic tangent (Tanh), and rectified linear unit (ReLU), were tested to optimize the model's performance. The results revealed that the hyperbolic tangent function demonstrated superior performance in predicting critical variables, emphasizing its potential for improving predictive accuracy in counterterrorism efforts. The research also explored the BNN model's performance under varying training set proportions, confirming its robustness and reliability across different data conditions.

Furthermore, the study by Iorlíam et al., [45] utilized the Apriori algorithm on the Global Terrorism Database (GTD) dataset with a focus on Nigeria. Motivated by these findings, this paper proposes the novel application of decision trees, random forests, and logistic regression for attack classification on both the GTD and ACLED datasets. The overarching goal of this research is to provide actionable insights that help government and security agencies strengthen resilience against terrorism, safeguarding both the population and national infrastructure from the growing threat of violent extremism.

In summary, the integration of machine learning techniques, especially advanced models like neural networks and Bayesian networks, plays a pivotal role in enhancing the predictive capabilities of counterterrorism efforts. By analyzing terrorist attack patterns and identifying key influencing factors, these models provide valuable insights that can inform targeted strategies, improve security measures, and mitigate the impact of terrorist threats.

2. METHOD AND MATERIALS

2.1. Dataset description and pre-processing

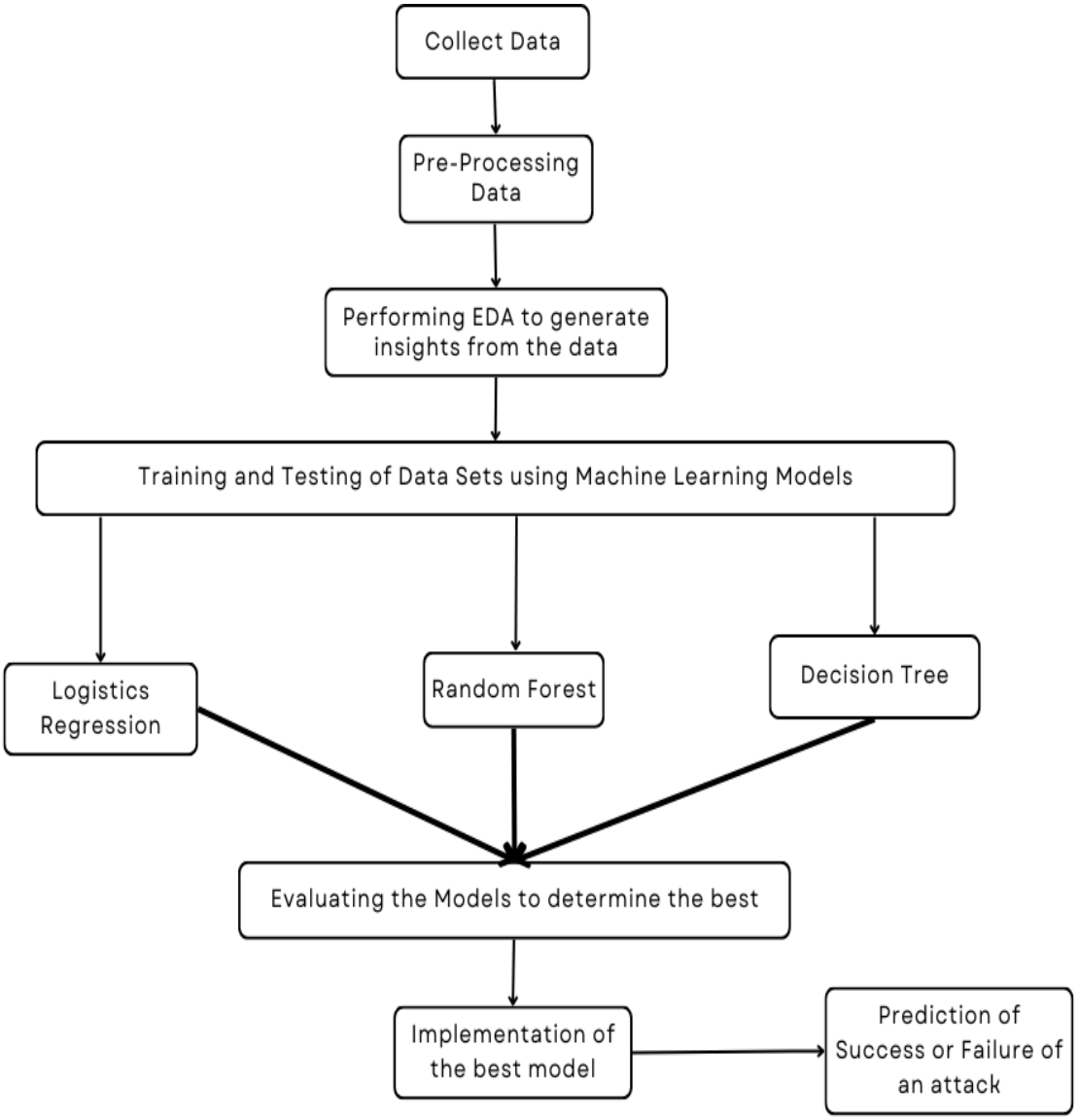
The data sources for this paper are the Global Terrorism Database (GTD) and the Armed Conflict Location and Event Data Project (ACLED). For the pre-processing, we handled missing values, removed duplicates, and filtered the data of importance for this research. Furthermore, we performed exploratory data analysis specifically targeting nearly three decades of attacks occurring within Nigeria, utilizing a combination of Excel, Microsoft Power BI, and Google Colab to generate insights into the patterns and dynamics of violence in the region.

2.2. Procedure

The dataset is divided into two sets: a training and test set. The training set is used to train machine learning models, while the test set is used to evaluate the model's performance on unseen data. Three machine learning classifiers are applied to the dataset: decision tree, random forest, and logistic regression for attack classification purposes in Nigeria. Figure 1 illustrates the proposed methodology.

Figure 1

Proposed methodology for insights and uncovering terrorist activities in Nigeria



This paper further adopts and uses the evaluation metrics commonly used in similar research, including accuracy, precision, recall, and f1-score [46].

3. RESULTS

3.1. Experiments and results analysis

An analysis was conducted utilizing two renowned datasets: the Global Terrorism Database (GTD) and the Armed Conflict Location and Event Data Project (ACLED). The GTD dataset spans from 1970 to 2020, with a focus on incidents occurring between 1997 and 2020. Furthermore, the ACLED covers data from 1997 to 2024.

Specifically, GTD tracks terrorism-related incidents, whereas ACLED concentrates on armed conflict-related events. Figure 2 shows the top 10 affected states in Nigeria from ACLED (1997-2024 January).

Figure 2

Top 10 affected states in Nigeria from ACLED (1997-2024 January)

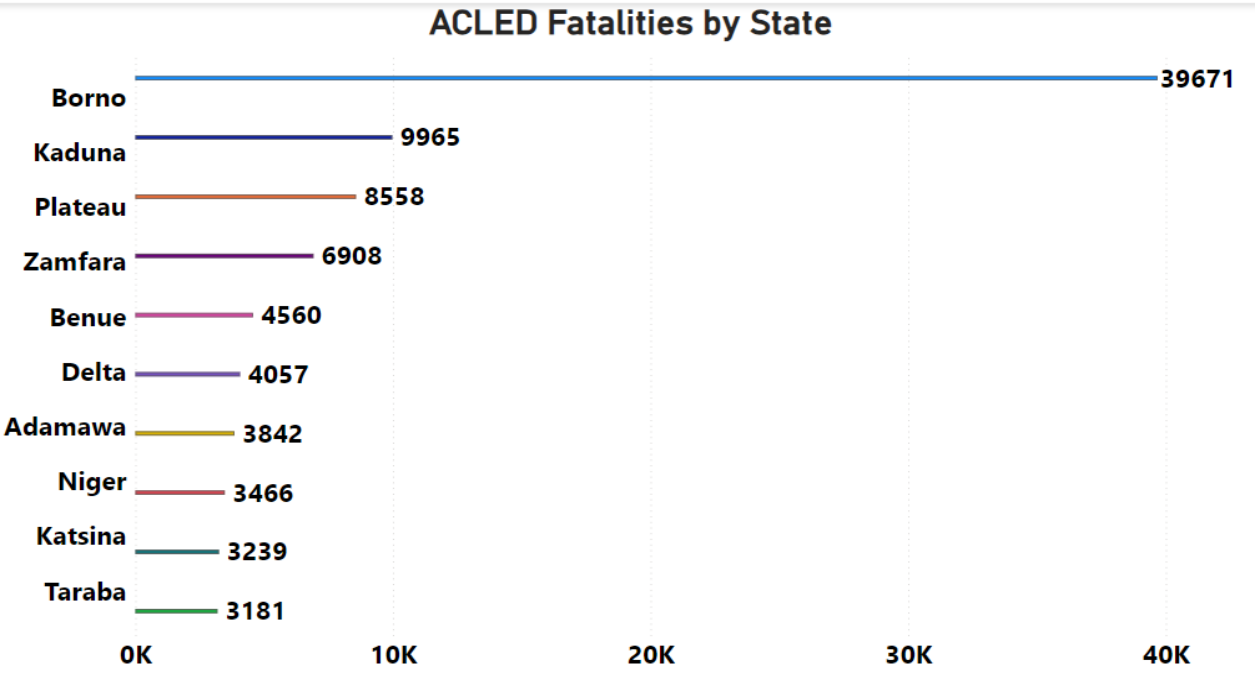
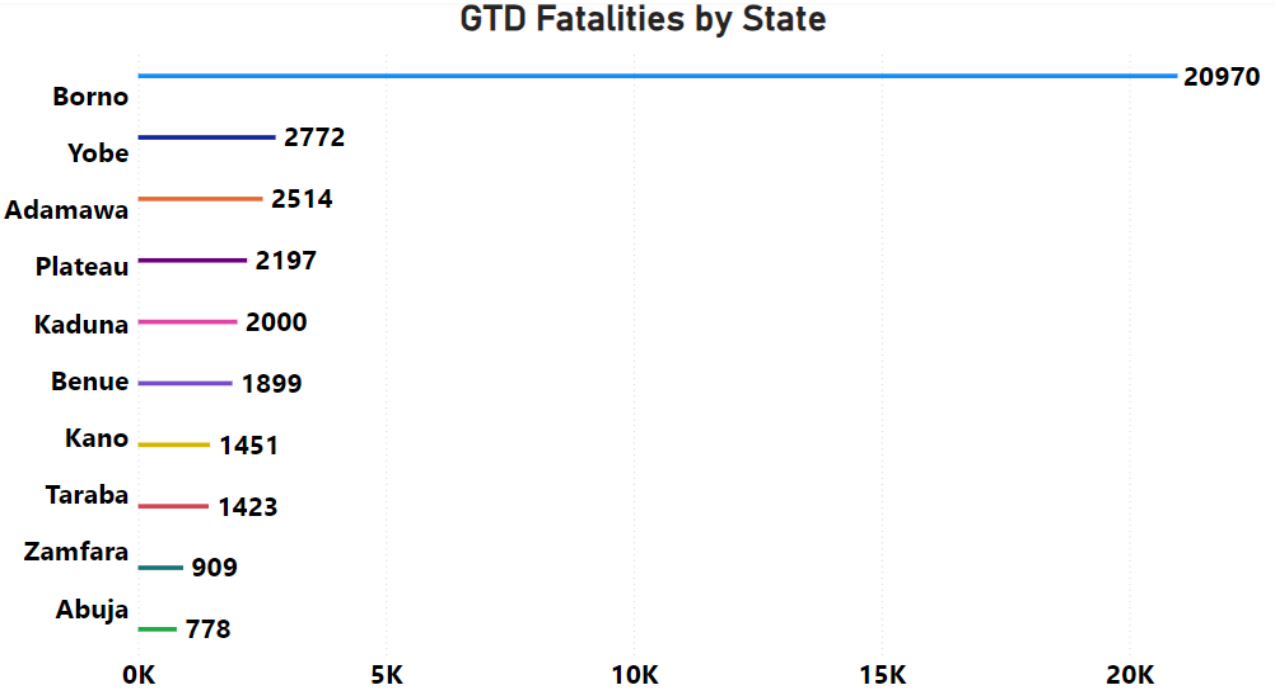


Figure 3

Top 10 affected states in Nigeria from GTD (1997-2020)



A comparison of Figures 2 and 3 reveals discrepancies in the ranking of states by fatalities between the two databases. Nevertheless, some states exhibit consistent trends across both datasets. Notably, Borno State consistently records high fatality numbers in both databases, indicating a significant prevalence of conflict and terrorism-related incidents in the region.

Figure 4

Fatalities by year in Nigeria from ACLED (997-2024 January)

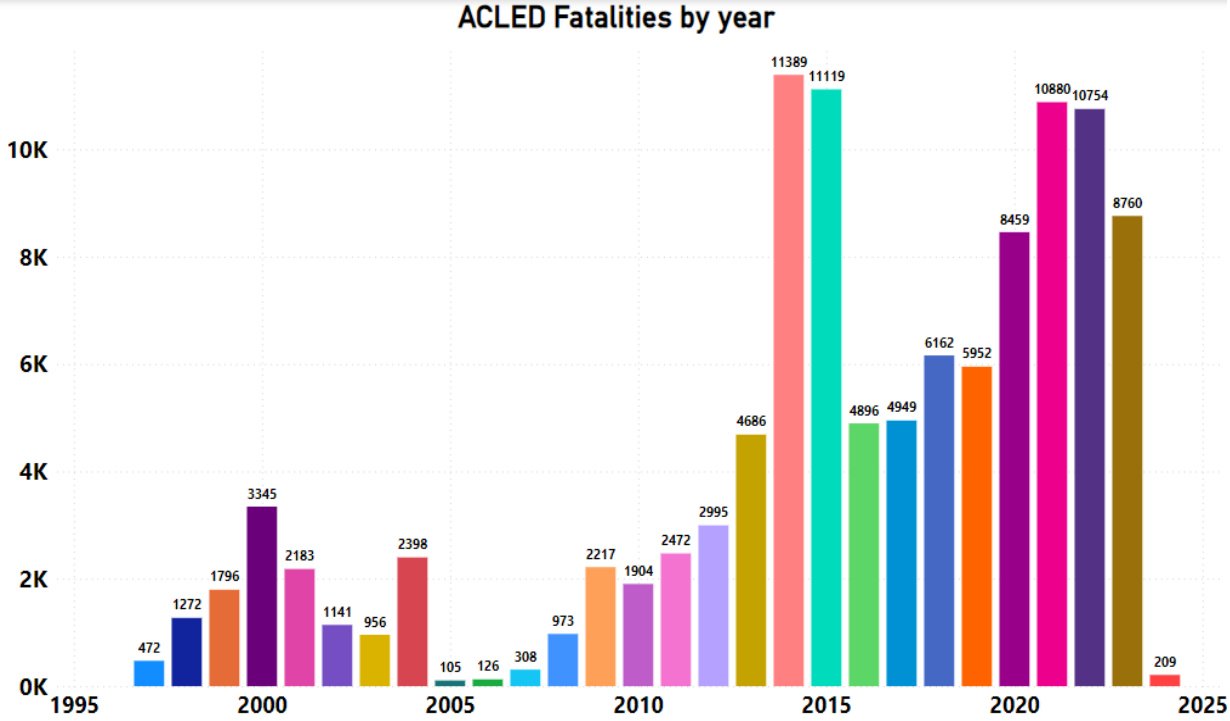
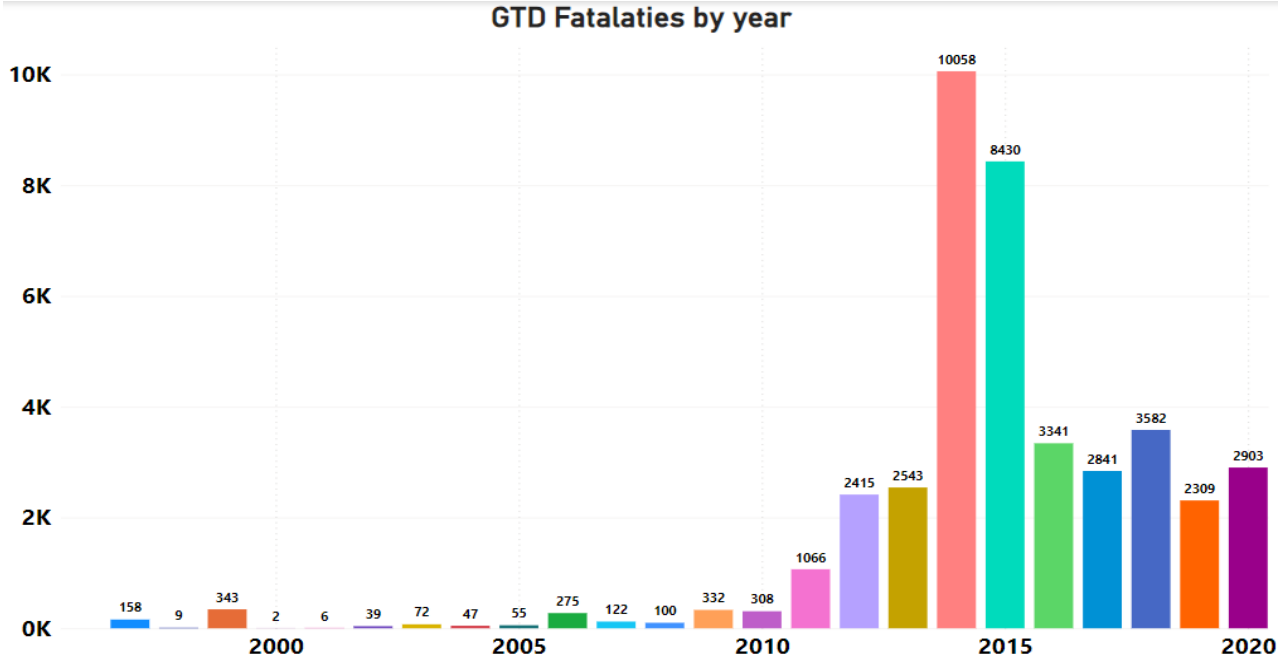


Figure 5

Fatalities by year in Nigeria from GTD (1997-2020)



The ACLED chart illustrates fluctuations in fatalities over the years, characterized by varying peaks and troughs at different points in time. Conversely, the GTD chart depicts a general upward trend in fatalities observed throughout the years under examination. Interestingly, both datasets (see Figures 4 and 5) align notably in 2014 and 2015, indicating these years as periods with the highest recorded fatalities.

Figure 6

Fatalities by attack type in Nigeria from ACLED (1997-2024 January)

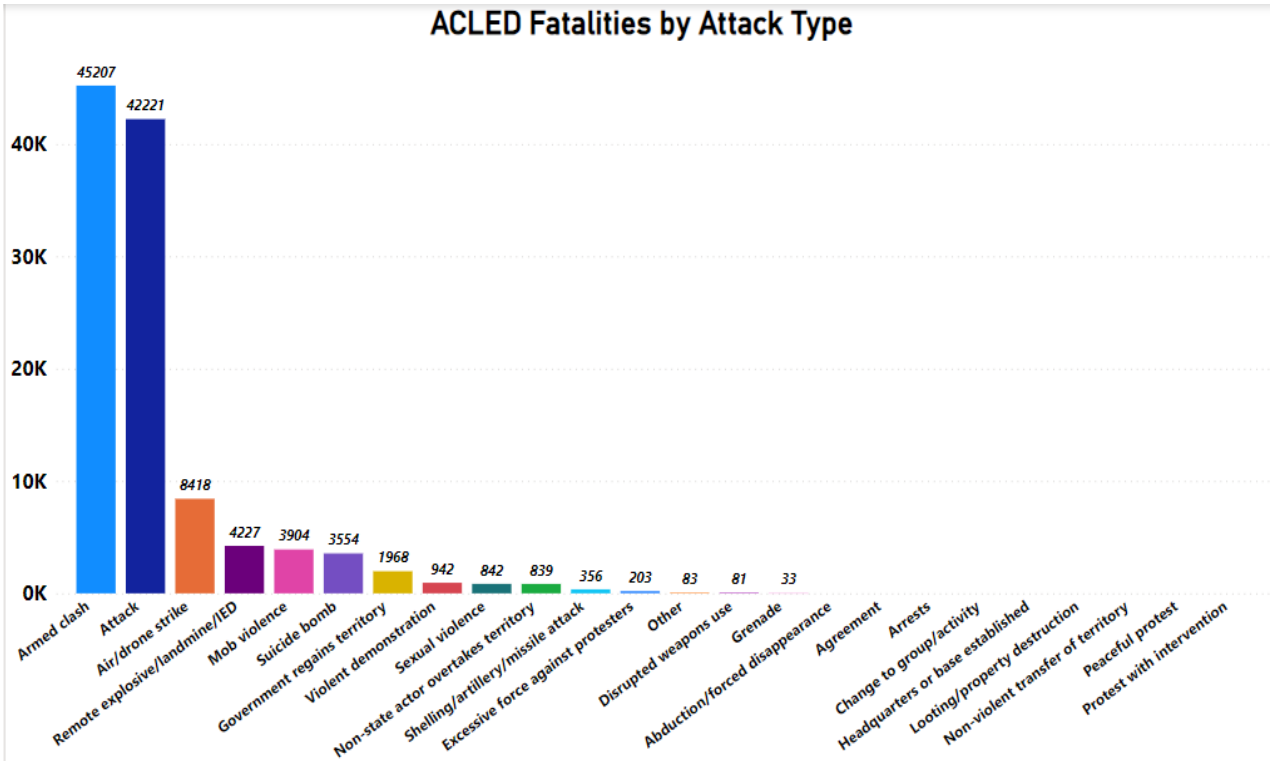
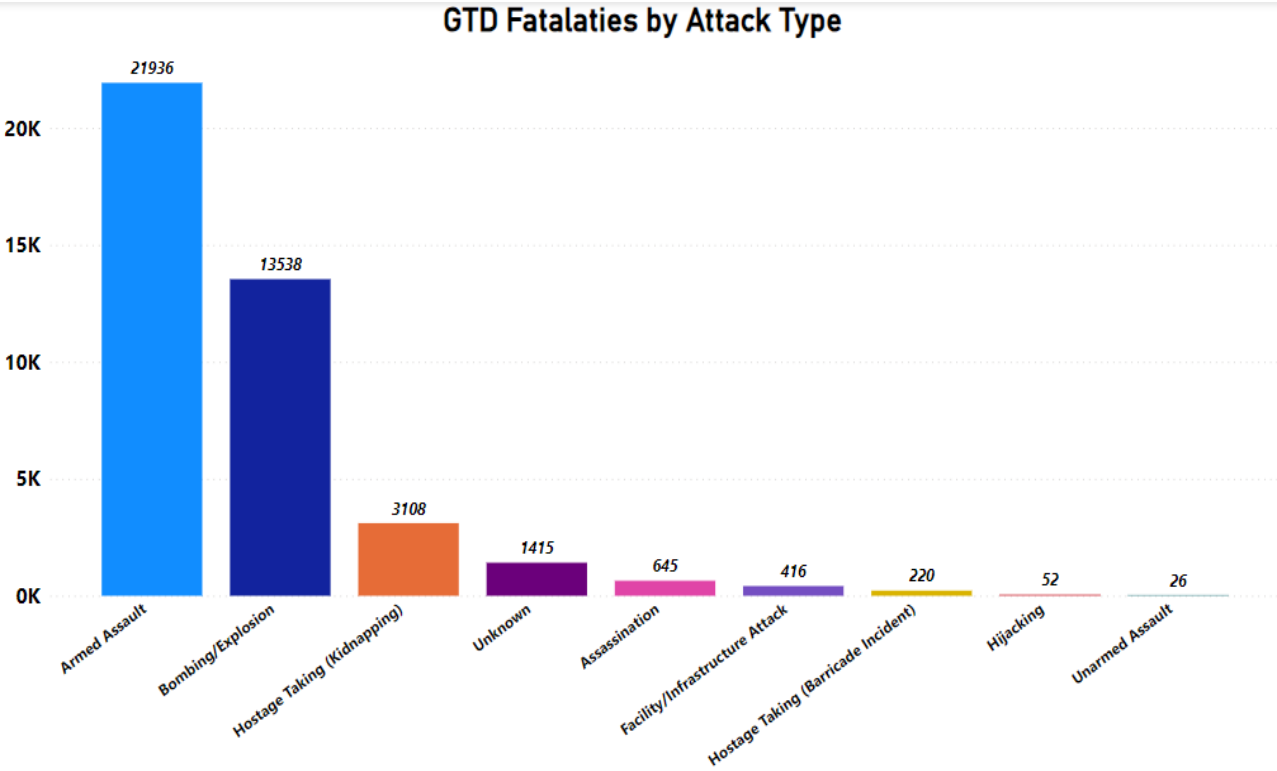


Figure 7

Fatalities by attack type in Nigeria from GTD (1997-2020)



A comparative analysis of Figures 6 and 7 reveals distinct categorizations of fatalities by attack types. The GTD chart provides a detailed breakdown of specific attack types, notably armed assault and bombing/explosion. In contrast, ACLED broader categorization encompasses additional attack types, including air/drone strikes, mob violence, and sexual violence. This difference in categorization scope may contribute to discrepancies in total fatalities between the two datasets, as ACLED includes violence types with a potentially lower fatality rate. The variation in attack type categorization between GTD and ACLED underscores the importance of standardized classification systems in conflict data analysis.

Figure 8

Top 10 Terrorist Groups in Nigeria from ACLED (1997-2024 January)

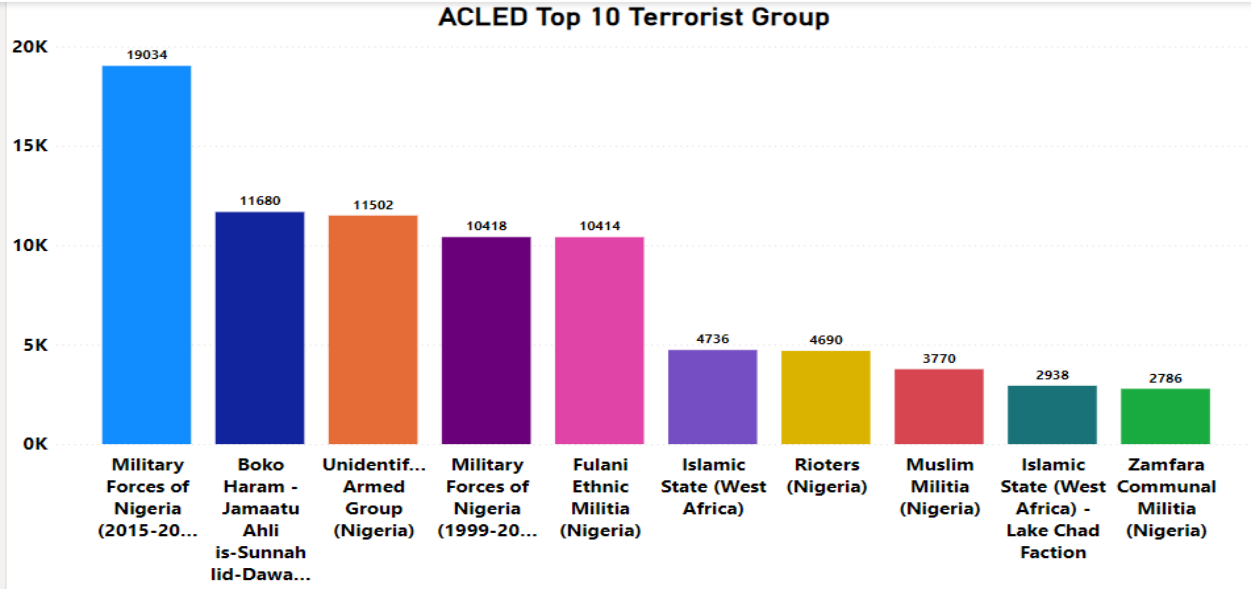
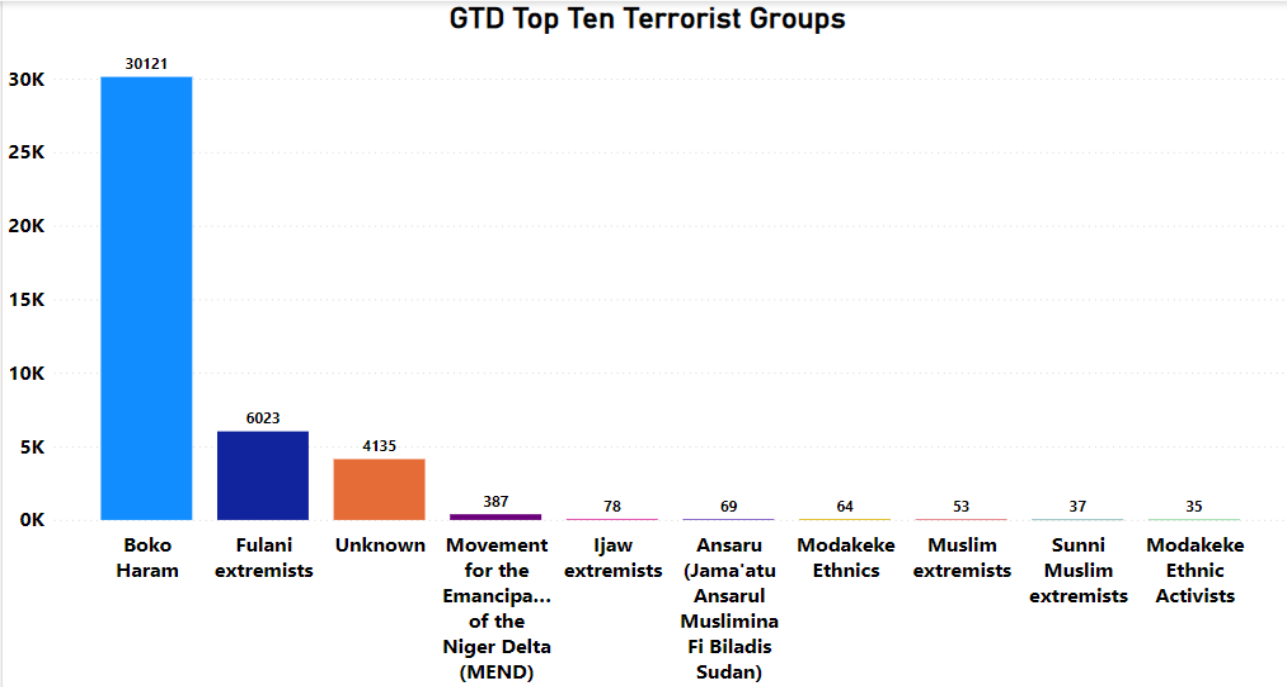


Figure 9

Top 10 terrorist groups in Nigeria from GTD (1997-2020)



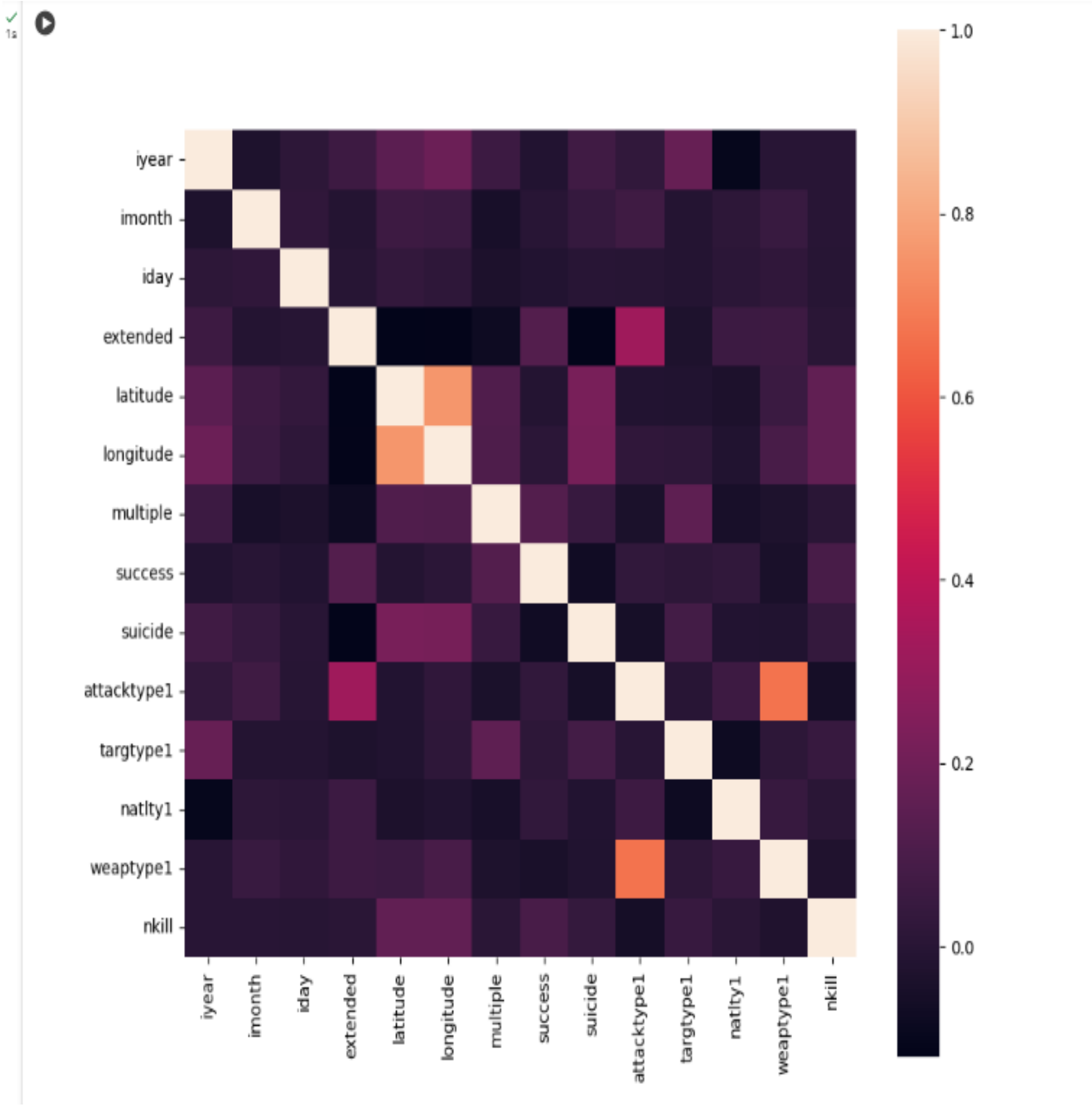
Boko Haram emerges as the primary terrorist group responsible for significant havoc and fatalities in Nigeria, according to both the ACLED (Figure 8) and GTD (Figure 9) datasets. The military forces of Nigeria ranked highly in Figure 8, do not qualify as a terrorist group.

3.2. Machine learning model

Given the GTD dataset's focus on terrorism-related incidents and relevant variables for attack classification, it serves as the primary feature set for machine learning techniques employed in this study. A correlation matrix, illustrated in Figure 10, was generated based on the GTD dataset features.

Figure 10

Correlation matrix visualizing



There is a noticeable strong correlation between the variables attack type and weapon type. However, aside from this correlation, the remaining variables exhibit relatively low correlations with each other, which is advantageous for the model we developed.

We therefore split the data into a training set comprising 70% of the data and a test set containing 30%. The variable we aim to predict is "success," which indicates whether a terrorist attack will be successful or not. To ensure consistency in the split and avoid introducing sampling bias, the “random state” variable in Python was set to '42', thereby keeping the random number generator constant. This approach ensures that we consistently obtain the same split each time the data is partitioned.

3.2.1. Decision tree results

A decision tree model was implemented, with the code snippet displayed in Figure 11.

Figure 11

Code snippet for decision tree classifier

```
[19]  1 from sklearn import tree
      2 import pandas as pd
      3 from sklearn.tree import DecisionTreeClassifier
      4 from sklearn.model_selection import train_test_split
      5 from sklearn import metrics
      6 y = df['success']
      7 X = df[features]
      8 dtree = tree.DecisionTreeClassifier()
      9 dtree = dtree.fit(X_train,y_train)
```

```
[20]  1 dtree_pred = dtree.predict(X_test)
      2 from sklearn.metrics import classification_report,confusion_matrix
      3 print(classification_report(y_test,dtree_pred))
```

Figure 12

Decision tree performance evaluation

```
                   precision    recall  f1-score   support

                0       0.48      0.50      0.49       123
                1       0.96      0.96      0.96      1542

         accuracy                           0.92      1665
        macro avg       0.72      0.73      0.72      1665
     weighted avg       0.92      0.92      0.92      1665
```

```
[22]  1 print(confusion_matrix(y_test,dtree_pred))
```

```
[[  61   62]
 [  66 1476]]
```

The decision tree model yielded an accuracy of 92%, as illustrated in Figure 12.

3.2.2. Random forest results

A Random Forest model was implemented in Python, with the corresponding code snippet displayed in Figure 13.

Figure 13

Code snippet for random forest classifier

```
[23]  1 from sklearn.ensemble import RandomForestClassifier
      2 rf = RandomForestClassifier(n_estimators=400)
      3 rf = rf.fit(X_train, y_train)
      4 rf_pred = rf.predict(X_test)
```

```
[24]  1 print(classification_report(y_test,rf_pred))
```

Figure 14

Random forest performance evaluation

```
[24]                precision    recall  f1-score   support

           0       0.86      0.44      0.58       123
           1       0.96      0.99      0.98      1542

    accuracy                           0.95      1665
   macro avg       0.91      0.72      0.78      1665
weighted avg       0.95      0.95      0.95      1665
```

```
[25]  1 print(confusion_matrix(y_test,rf_pred))
```

```
[[  54   69]
 [   9 1533]]
```

The Random Forest model achieved an accuracy of 95%, as illustrated in Figure 14.

3.2.3. Logistic regression results

The logistic regression model was implemented in Python, with the corresponding code snippet displayed in Figure 15.

Figure 15

Code snippet for logistic regression classifier

```
[26]  1 from sklearn.linear_model import LogisticRegression
      2 logreg = LogisticRegression(random_state=42)
      3 logreg.fit(X_train, y_train)
      4 y_pred = logreg.predict(X_test)
```

```
[27]  1 print(classification_report(y_test,y_pred))
```

Figure 16

Logistic regression performance evaluation

```
[27]
0s
                   precision    recall  f1-score   support

                0       0.00      0.00      0.00       123
                1       0.93      1.00      0.96      1542

         accuracy                           0.92      1665
        macro avg       0.46      0.50      0.48      1665
     weighted avg       0.86      0.92      0.89      1665
```

```
[28]  1 print(confusion_matrix(y_test,y_pred))
0s
[[   0  123]
 [   3 1539]]
```

The logistic regression achieved an accuracy of 92%. Overall, comparing the Confusion Matrix of Decision Tree, Random Forest, and Logistic Regression, the Random Forest model achieved the highest accuracy of 95%, and as such, it is more suitable for attack classification in this paper.

4. DISCUSSION

The results of this study underscore the complexity and utility of multiple datasets in analyzing terrorist activities and armed conflicts in Nigeria. When comparing the Global Terrorism Database (GTD) and the Armed Conflict Location and Event Data Project (ACLED), it is evident that while both datasets provide valuable insights, discrepancies in their findings, such as the varying rankings of affected states and the categorization of fatalities highlight the need for standardized classification systems in conflict data analysis. Notably, Borno State consistently emerges as a hotspot for both terrorism and armed conflict, demonstrating the region's significant role in the ongoing violence. Furthermore, the comparison of fatalities over time from both datasets reveals a convergence around the years 2014 and 2015, which stands out as a period marked by particularly high fatality numbers in Nigeria, underlining the intensity of violence during that period.

The differences in attack type categorization between the GTD and ACLED datasets also warrant attention. While the GTD provides a more granular classification, focusing on specific attack types such as armed assault and bombings, the ACLED dataset includes a broader range of violence types, including mob violence and air strikes, which could lead to discrepancies in the total fatalities reported. This difference in scope may explain some of the variations in fatality counts between the two datasets, underscoring the importance of having consistent and well-defined categories when analyzing conflict data. These findings stress the need for improved standardization and alignment between datasets to facilitate more accurate and meaningful comparisons of conflict-related statistics.

From a machine learning perspective, the study employed several models to predict the success of terrorist attacks using the GTD dataset. The decision tree model, random forest model, and logistic regression model were all tested for their accuracy in classifying whether an attack would be successful. Of these, the random forest model outperformed the others, achieving an accuracy of 95%. This suggests that random forests are particularly well-suited for this task, as they effectively handle the complexities of the data without overfitting. The decision tree and logistic regression models also performed well, with accuracies of 92%, but the random

forest model demonstrated superior performance, highlighting its utility in the context of attack classification and its potential for future research in terrorist attack prediction.

5. CONCLUSION AND FUTURE WORK

This research makes a significant contribution to the understanding of terrorism in Nigeria by demonstrating the effectiveness of data science, particularly machine learning techniques, in analyzing and predicting terrorist attacks. By utilizing advanced machine learning models, the study showcases how these tools can effectively classify attacks, identify key dynamics of terrorist activities, and support the development of data-driven counterterrorism strategies. The ability to analyze large datasets and uncover patterns in terrorist activities offers valuable insights that can inform more targeted and efficient responses to emerging threats.

The analysis underscores the importance of machine learning for anticipating and mitigating potential threats. Among the models tested, the Random Forest model emerged as the most effective, achieving an impressive 95% classification accuracy. This high level of performance positions the model as a promising tool for predicting the success of terrorist attacks. Such predictive capabilities can enhance the preparedness and response strategies of security and intelligence agencies, enabling them to take proactive measures in real-time to thwart potential attacks. The study highlights how integrating these models into security frameworks can significantly improve decision-making and resource allocation in counterterrorism efforts.

Furthermore, the research emphasizes the need for further exploration of machine learning techniques to refine the classification and prediction of terrorist activities. Expanding the analysis to include terrorist attack patterns in other countries facing similar security challenges, such as Cameroon, Chad, Benin, Niger, and Guinea, could provide a more comprehensive understanding of cross-border terrorist threats. A regional approach to studying terrorism would allow for the identification of shared patterns and risk factors, facilitating the development of more effective regional and global counterterrorism strategies. The next steps in this research, including the investigation of additional machine learning models and comparative regional analyses, are crucial for advancing predictive capabilities and strengthening security frameworks in response to the evolving tactics of terrorist organizations.

Conflict of Interest: The authors declare no conflict of interest.

Ethical Approval: The study adheres to the ethical guidelines for conducting research.

Funding: This research received no external funding.

REFERENCES

[1] A. C. Okoli and P. Iortyer, "Terrorism and humanitarian crisis in Nigeria: Insights from Boko Haram insurgency," *Global Journal of Human Social Science*, vol. 14, no. 1, pp. 39–49, 2014.

[2] E. Ben-Edet, *Terrorism: A Case Study of the Global Security Threat of Boko Haram and The ISIS Alliance in Nigeria*, Doctorate dissertation, Texas Southern University, 2022.

[3] The Institute for Economics & Peace, *Global Terrorism Index 2023: Measuring the Impact of Terrorism*, Sydney, 2023. [Online]. Available: https://www.dragonflyintelligence.com/intelligence/terrorismtracker/

[4] A. Bolpagni, "The Central Sahel's Powder Keg: The Inexorable Ascendancy of Jihadist Groups in the Shadow of a Renewed International Competition," *INSIGHT*, 2023.

[5] S. Wojciechowski, "The Spirit of Terrorism–its Contemporary Evolution and Escalation," *Przeglqd Strategiczny*, vol. 13, no. 16, pp. 9–22, 2023.

[6] N. Obasi, *Nigeria*, International Crisis Group, 2023. [Online]. Available: https://www.crisisgroup.org/africa/west-africa/nigeria

[7] A. P. Schmid, *Defining terrorism*, International Centre for Counter-Terrorism, 2022.

Bello A. M., Iorliam, A. & Asilkan 0. (2024). Data science insights and the classification of terrorist attacks in Nigeria using machine learning techniques. *Global Journal of Computer Sciences: Theory and Research*. *14*(2), 30-48. https://doi.org/10.18844/gjcs.v14i2.9669

[8] O. S. Osadola and G. S. Emah, "Terrorism in Nigeria and Her Neighbours," *Konfrontasi: Jurnal Kultural, Ekonomi Dan Perubahan Sosial*, vol. 9, no. 3, pp. 439–448, 2022.

[9] F. M. Onana Ibogo, *A Framework for the Sustainable Management of Water Resources in Lake Chad*, Doctoral dissertation, Wien, 2023.

[10] S. Jackson, *The African Union Solution to the Challenge of Terrorism: An Assessment of Regional Actors' Role in Countering Terrorism in Africa*, 2024.

[11] T. P. Ogundunmade and A. A. Adepoju, "Predicting the Nature of Terrorist Attacks in Nigeria Using Bayesian Neural Network Model," in *Sustainable Statistical and Data Science Methods and Practices: Reports from LISA 2020 Global Network, Ghana, 2022*, Cham: Springer Nature Switzerland, 2024, pp. 271–286.

[12] M. A. Idakwo, R. E. Yoro, P. Achimugu, and O. Achimugu, "An Improved Weapons Detection and Classification System," *Journal of Network and Innovative Computing*, vol. 12, pp. 10–10, 2024.

[13] J. L. S. González, C. Zaccaro, J. A. Álvarez-García, L. M. S. Morillo, and F. S. Caparrini, "Real-time gun detection in CCTV: An open problem," *Neural Networks*, vol. 132, pp. 297–308, 2020.

[14] O. C. Bordeanu, *From Data to Insights: Unraveling Spatio-Temporal Patterns of Cybercrime using NLP and Deep Learning*, Doctoral dissertation, UCL (University College London), 2024.

[15] Global Terrorism Database, *Codebook: Methodology, Inclusion Criteria, and Variables*, Maryland, 2021. [Online]. Available: https://www.start.umd.edu/gtd

[16] ACLED, *Armed Conflict Location & Events Data*, 2024. [Online]. Available: https://acleddata.com/

[17] O. Ajala, "Formation of insurgent groups: MEND and Boko Haram in Nigeria," *Small Wars & Insurgencies*, vol. 29, no. 1, pp. 112–130, 2018.

[18] O. S. Oladimeji, *Terrorism in Nigeria: Causes, Consequence and Panacea*, 2019.

[19] C. Obi, "Challenges of insecurity and terrorism in Nigeria: Implication for national development," *OIDA International Journal of Sustainable Development*, vol. 8, no. 2, pp. 11–18, 2015.

[20] N. Zubairu, "Rising insecurity in Nigeria: Causes and solution," *Journal of Studies in Social Sciences*, vol. 19, 2020.

[21] G. Odafe, "Boko Haram, Fulani Herdsmen and Possible Humanitarian Crises: Evaluation and Recommendations," *Journal of African Studies and Sustainable Development*, vol. 4, no. 5, 2021.

[22] M. Yamamoto, *Terrorism against democracy*, Center for International & Security Studies at Maryland, 2015.

[23] R. Douglas, *Law, liberty, and the pursuit of terrorism*, University of Michigan Press, p. 336, 2014.

[24] X. Pan, "Quantitative analysis and prediction of global terrorist attacks based on machine learning," *Scientific Programming*, vol. 2021, no. 1, p. 7890923, 2021.

[25] D. Lewinsky, D. Te'eni, I. Yahav-Shenberger, D. G. Schwartz, G. Silverman, and Y. Mann, "Detecting terrorist influencers using reciprocal human-machine learning: The case of militant Jihadist Da'wa on the Darknet," *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1–11, 2024. [Online]. Available: https://www.nature.com/articles/s41599-024-03920-7

[26] E. Bakker, "Forecasting terrorism: The need for a more systematic approach," *Journal of Strategic Security*, vol. 5, no. 4, pp. 69–84, 2012.

[27] G. M. Tolan and O. S. Soliman, "An experimental study of classification algorithms for terrorism prediction," *International Journal of Knowledge Engineering-IACSIT*, vol. 1, no. 2, pp. 107–112, 2015.

[28] N. E. M. Khalifa, M. H. N. Taha, S. H. N. Taha, and A. E. Hassanien, "Statistical insights and association mining for terrorist attacks in Egypt," in *International Conference on Advanced Machine Learning Technologies and Applications*, Cham: Springer International Publishing, 2019, pp. 291–300.

[29] S. J. Krieg, C. W. Smith, R. Chatterjee, and N. V. Chawla, "Predicting terrorist attacks in the United States using localized news data," *PloS one*, vol. 17, no. 6, p. e0270681, 2022.

Bello A. M., Iorliam, A. & Asilkan 0. (2024). Data science insights and the classification of terrorist attacks in Nigeria using machine learning techniques. *Global Journal of Computer Sciences: Theory and Research*. *14*(2), 30-48. https://doi.org/10.18844/gjcs.v14i2.9669

[30] T. Saheb, "Ethically contentious aspects of artificial intelligence surveillance: a social science perspective," *AI and Ethics*, vol. 3, no. 2, pp. 369–379, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s43681-022-00196-y

[31] W. R. W. Rosli, "Waging warfare against states: the deployment of artificial intelligence in cyber espionage," *AI and Ethics*, pp. 1–7, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s43681-024-00628-x

[32] A. B. Krueger, *What makes a terrorist: Economics and the roots of terrorism*, Princeton University Press, 2008.

[33] F. Gohar, W. H. Butt, and U. Qamar, "Terrorist group prediction using data classification," *Work. MultiRelational Data Min. MRDM2003*, vol. 10, pp. 199–208, 2014.

[34] V. Kumar, M. Mazzara, A. Messina, and J. Lee, "A conjoint application of data mining techniques for analysis of global terrorist attacks: Prevention and prediction for combating terrorism," in *Proceedings of 6th International Conference in Software Engineering for Defence Applications: SEDA 2018 6*, Springer International Publishing, 2020, pp. 146–158.

[35] R. Douglas, *Law, liberty, and the pursuit of terrorism*, University of Michigan Press, p. 336, 2014.

[36] A. Karakikes, P. Alexiadis, and K. Kotis, "Bias in X (Twitter) and Telegram Based Intelligence Analysis: Exploring Challenges and Potential Mitigating Roles of AI," *SN Computer Science*, vol. 5, no. 5, p. 574, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s42979-024-02935-w

[37] X. Pan, "Quantitative analysis and prediction of global terrorist attacks based on machine learning," *Scientific Programming*, vol. 2021, no. 1, p. 7890923, 2021.

[38] H. Chen, A. L. Houston, R. R. Sewell, and B. R. Schatz, "Internet browsing and searching: User evaluations of category map and concept space techniques," *Journal of the American Society for Information Science*, vol. 49, no. 7, pp. 582–603, 1998.

[39] N. R. Vajjhala, K. D. Strang, and Z. Sun, "Statistical modeling and visualizing open big data using a terrorism case study," in *2015 3rd International Conference on Future Internet of Things and Cloud*, 2015, pp. 489–496, IEEE.

[40] A. Majekodunmi, "Terrorism and counter-terrorism in contemporary Nigeria: Understanding the emerging trends," *Journal of Policy and Development Studies*, vol. 289, no. 2379, pp. 1–18, 2015.

[41] A. Blanchard and M. Taddeo, "The ethics of artificial intelligence for intelligence analysis: a review of the key challenges with recommendations," *Digital Society*, vol. 2, no. 1, p. 12, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s44206-023-00036-4

[42] A. Falzon and J. Azzopardi, "Terror Mine: Automatically Identifying the Group behind a Terrorist Attack," in *KDIR*, 2022, pp. 221–228.

[43] S. Shinde, S. Khoje, A. Raj, L. Wadhwa, and A. S. Shaikha, "Artificial intelligence approach for terror attacks prediction through machine learning," *Multidisciplinary Science Journal*, vol. 6, no. 1, pp. 2024011–2024011, 2024.

[44] O. A. Odeniyi, M. E. Adeosun, and T. P. Ogundunmade, "Prediction of terrorist activities in Nigeria using machine learning models," *Innovations*, vol. 71, pp. 87–96, 2022.

[45] A. Iorliam, R. U. Dugeri, B. O. Akumba, S. Otor, and Y. I. Shehu, "An Investigation and Insight Into Terrorism In Nigeria," *arXiv preprint arXiv:2109.11023*, 2021.

[46] A. Iorliam, S. Bum, I. S. Aondoakaa, I. B. Iorlıam, and Y. Shehu, "Machine learning techniques for the classification of IoT-enabled smart irrigation data for agricultural purposes," *Gazi University Journal of Science Part A: Engineering and Innovation*, vol. 9, no. 4, pp. 378–391, 2022.