DESIGN OF AN ENSEMBLE MACHINE LEARNING BASED RECOMMENDER SYSTEM FRAMEWORK FOR TERRORISM PREVENTION

[1]Jimoh, I. T. and [2]Daramola, C. Y.

ABSTRACT

Terrorists cause a lot of unrest, fear and destruction of live and property worth trillions of Naira in Nigeria and the entire world. Several authors had deployed hard computing techniques like kinetic approach to prevent and combat the heinous act but the menace kept increasing. Hence, there is need for deployment of soft computing techniques such as machine learning to combat the problem of terrorism. This study created a machine learning method to forecast terrorist activity and warn the public and security organizations so they can take preventative action. The paper proposes bagging techniques, consisting of the traditional ensemble module (logistic regressing, random forest and support vector machine) and deep learning module (bidirectional long short-term memory and bidirectional encoder representation from transformer) to explore both the global terrorist dataset (GTD) and dataset obtained from social media platform for predicting the likelihood of the terrorist attack, the likely time and possible location of future attack.

Keywords: Terrorism, Assemble, Prediction, Twister, GTD

[1]*Department of Software Engineering, The Federal University of Technology, Akure, Nigeria*

[2]*Department of Computer Science, Federal University, Oye-Ekiti, Nigeria*

Correspondence
itjimoh@futa.edu.ng

History
Received: 21/02/2025
Accepted: 16/05/2025
Published: 30/05/2025

1 | Introduction

The term "terrorism" was first used during the French revolution to simply means the “reign of terror”, in the 1790s (UNODC, 2025). Falk (2004) described terrorism as a political and military tool, employed by certain parties dated back to the time immemorial, making it just as old as armed conflict and governance. Terrorism in Nigeria dated back to 1986, when Dele Giwa was killed with letter bomb. Several lives and properties worth trillions of dollars have been lost to terrorism activities, prediction and elimination (Afolabi, 2013). The permeable land borders and inadequate digital surveillance of Nigeria's digital environment have made terrorism a concern to the country (Iorliam *et al.*, 2021). Because of the ever-changing nature of terrorism, the impact of the efforts made by the established authority is not readily apparent. Terrorism ranges from suicide bombing, kidnapping, hostage taking, vandalism, religion intolerance to maiming and killing of innocent citizens. Terrorists are the actors that cause a lot of unrest, fear, and destruction in Nigeria and the entire world (Adewale *et al*., 2024 and Jimoh *et al*., 2024).

Technology advancements have improved the social media platforms and modern communication systems, awareness of the need for self-enrichment is growing. User-generated data from social media can assist analysts in profiling consumers,

https://www.futa.edu.ng
JOSTIR
JOURNAL OF SCIENCE, TECHNOLOGY AND INNOVATION RESEARCH
https://www.jostir.futa.ng

https://doi.org/10.51459/jostir.2025.1.1.010

but there can be major hurdles due to the noise produced and the expansion of data sizes (Gilbert *et al*., 2019). Treistman (2023) asserted that studies on the origins of terrorism frequently overlook contextual elements that have an impact on individuals, such as social exclusion, in favor of macro trends at the national level. Minaee *et al*. (2021) summarized over forty (40) popular frequently used text classification datasets with deep learning models to offers a thorough analysis of approximations. It also discusses the technical contributions, similarities, and strengths of each model. Cheong (2018) examined the connection between terrorism and feelings of guilt and retaliation.

1.1 | Literature Review

Soliman *et al*. (2019) presented an artificial neural network-based system for terrorism prediction (ANN). The system identifies the terrorist group or groups that are responsible for terror attacks in various nations. This system also serves as alarm tool to identify the networks of these terrorist groups. The work explores terrorism phenomenon, which is accountable for the terrorist incidents that occurred in Egypt between 1996 and 2017, by creating hybrid computational intelligence framework. With the help of the metaheuristics optimization algorithm, the implemented algorithms combine a feature selection approach that uses both filter and wrapper approaches in a hybrid system using k-nearest neighbour (KNN) and random forest (RF). The algorithm determines the smallest number of features that are selected that achieve the highest classification accuracy using the GTD dataset, and accuracy is used as a performance metric. As the last stage of the prediction process, the results show promising prediction accuracy based on Artificial Neural Networks, enabling the prediction system to be used as an alert in the future as a researched tool for the existence of terrorist groups' networks and to reduce the frequency of terrorist attacks. Instead of employing the wrapper technique and a variety of metaheuristic optimization algorithms, the work did not take into account a hybrid embedded feature selection approach.

Huamaní *et al*. (2020) utilized the Global Terrorism Database, machine learning (ML) techniques to visualize and forecast terrorist events around the globe. The investigation sought to pinpoint the locations of the attacks, potential attack targets, and remedies to these occurrences. The scientists used artificial intelligence (AI) approaches to visualize and forecast potential attacks using two categorization models: random forests and decision trees and a database that has a comprehensive record of terrorist acts that have occurred worldwide between 1970 and 2018. Comparison of the accuracy of the two models were done to evaluate the model. With 500 leaves overall, Decision Tree strikes a balance between overuse and underuse. With regard to the assertiveness model, predictive results were produced with an accuracy percentage of 75.45%. The Decision Tree's prediction of the sorts of terrorist attacks has an accuracy rate of 79.24%. This percentage is highly good because, with Random Forest, the aggressiveness percentage is 89.544%, however a larger modification can display a percentage almost to 100%, which is an unfavorable result for the model to learn with fresh data. Likewise, the work achieved a 90.414% assertiveness rate when combined with the types of terrorist attacks conducted using Random Forest. However, the work did not explore Big Data techniques with social media dataset to determine possible threats.

Babitha *et al.* (2020) used ML techniques for the analysis of terrorism activities. The work was motivated by the need to have an early alert system that will trigger appropriate preventive measures to eradicate the terrorism and making the country safe

https://doi.org/10.51459/jostir.2025.1.1.010

and secure in order to improve the foreign investments. The project's goal was to forecast which terrorist organizations are most likely to attack a country by analyzing several terrorist elements through data mining from previous data. The research analyzed the GTD data, in order to find hidden patterns and insights. The following ML algorithms were used, including decision tree, naïve bayes, support vector machine, ensemble approaches, and random forest classification. Python and jupyter notebook were used to create visual representations of patterns and predictions. The outcomes demonstrate that the random forest classification algorithm has the highest accuracy among the models. The study estimated the number of injuries and fatalities and revealed the top 10 terrorist-affected regions in India. The investigation also determined which states are most likely to be targeted and exposed the main extremist groups that carry out the majority of attacks in India. In contrast to external threats, the experimental results indicate that extremist internal organizations are primarily responsible for the rise of terrorism in India.

Rigterink (2021) proofed that the rise in terrorism acts following the death of a terrorist leader can be attributed to retaliation. Höflinger (2021) posited that the lines separating the operations of terror groups and private citizens have become hazier due to advancements in communications technology. This work will design a ML system for terrorism prediction using data from Twitter (now X) incorporation and the Global Terrorism Database (GTD). Maniraj *et al.* (2019) explored ML algorithms to predict terror group. The authors were motivated by lack of existing work that examines the development or decline of terrorist organizations based on time, places of activity, attack types, targets of motivation, weapon availability and proficiency. The research work identified patterns and hidden structures in the terrorist activity and forecast the timing and nature of future attacks. The study employed logistic regression (LR), support vector machines (SVM), and kernel density normalization to aggregate and group real-time data from prior terrorist attacks around the world. To gain more insight and determine the best algorithm, evaluation criteria like accuracy, precision, recall, F1 scores, and ROC curves were employed. SVM produced best accuracy, followed by KNN and logistic regression. Other categorization algorithms could be investigated to enhance this suggested system.

Uddin *et al*. (2021) predicted the future activities of the terrorist using deep neural network (DNN)**.** The authors were motivated by the dynamism of terrorism in response to civilization. The work developed DNN model and compare the result with ML models. Five DNN based models are developed to comprehend terrorist behavior, including whether or not an attack will be effective or whether or not the attack will be a suicide? Or what kind of weapon will be employed in the assault? Or what kind of assault will be launched? or which area will be the target of an attack?). Three traditional ML algorithms are used to create the models: naïve bayes, logistic regression, and support vector machines (SVM). Also employed were a five-layer DNN and a single-layer neural network (NN). The efficacy of the NN and DNN-based model is juxtaposed with conventional ML algorithms. This study compares average precision, recall, F1-score, and average train and test accuracy. These findings show that DNN is the best model for this kind of dataset since it is an example of big data, and deep networks perform better when there is a greater amount of data. The highest performance of about 84% is attained. However, DNN's average accuracy was 95%.

Abdalsalam *et. al.* (2021) examined how textual characteristics affect the GTD dataset's ability to predict terrorist events. The detection of terrorist trends and behaviors is crucial for global

https://doi.org/10.51459/jostir.2025.1.1.010

counterterrorism strategies due to their regularity. The project's objectives were to improve the forecast accuracy of terrorist attack types and develop a framework for terrorism attack prediction using the global terrorism database. Seven (7) features were chosen to be combined with textual features after text features were extracted and represented using several text representation techniques, including Word Embedding (W2Vec), Bag of Words (BoW), and Term Frequency-Inverse Document Frequency (TF-IDF). Accuracy, precision, recall, and f1-score were the four performance indicators used to assess the performance. The outcomes demonstrated that distinct text features combined with important features (numeric, categorical) can significantly outperform conventional algorithms in identifying the kind of terrorist attack. Some observations regarding features extraction techniques and classifiers were also made in the three sets of tests carried out on GTD corpora. By expanding the suggested framework to include group names for the attack, further terrorist acts can be predicted, further advancing this research. Additionally, feature reduction approaches and auto encoder feature selection utilizing a deep learning model were not employed to improve prediction accuracy and reduce the complexity of the proposed framework.

Olabanjo *et al.* (2021) predicted the risk zone using ML model to achieve worldwide counterterrorism. The project's goal was to create a computer model for terrorism-related risk zone prediction. The researchers developed a stack ensemble ML model that combines SVM and KNN to predict the continents where a specific type of terrorism may occur. The dataset was taken from the GTD, an online database. Chi-squared (CS) and Information Gain (IG), as well as a hybrid of the two (HB), where the two feature selection methods were used on the dataset before a stack ensemble model was deployed. When it came to forecasting the location of potential terrorist occurrences, the HB approach yielded the greatest results in the shortest amount of time. Radial basis function (RBF), an SVM model, was chosen because it outperformed the sigmoid function in this experiment and had a localized, finite response throughout the whole x-axis. The continents where a specific occurrence is likely to occur were predicted using the stack ensemble model, and feature selection procedures produced 97.8% accuracy. The work did not focus on predicting the country but the Continent.

Wang *et al*. (2021) presented learning algorithms for predicting suicide from messages on social media. Using social media post data from the CLPsych 2021 shared challenge, the study created the C-Attention model of deep learning architecture and evaluated three other ML models to automatically identify people who might attempt suicide within 30 days and 6 months. In addition to the latent feature (Doc2Vec), these models were constructed utilising hand-crafted features such as emotions, POS, the three-step theory of suicide, and suicide dictionaries. According to the results, KNN and SVM (EB) both received the best F1 scores of 0.741 and F2 scores of 0.833 on subtask 1 (prediction of a suicide attempt days before), while C-Attention acquired the highest F1 scores of 0.737 and F2 scores of 0.843 on subtask 2 (prediction of suicide six months prior).

Odeniyi *et al.* (2022) developed ML models to forecast terrorists' activities and prevent its threat to today's civilization. The Nigeria Terrorism Database (NTD) provided the daily terrorism incidences across Nigeria, which served as the source of data for this study. The data includes the various attack types, their success and suicide rates, and the various weapon types that were employed in each attack. Highlights in the database included the objectives or victims of the terrorist

https://doi.org/10.51459/jostir.2025.1.1.010

attack, details on the perpetrators, casualties, and the aftermath of the act. The study employed a Heterogeneous Neural Network (HETNN) model and evaluated its performance against five different ML models: RF, Boosting, KNN, SVM, and LR. The results show that HETNN performs better in prediction than the other models. Social media data and the location of terror acts were not taken into account in this study. Bridgelall (2022) applied the Natural Language Processing to classify the intent of the terrorists. The author was motivated by lack of counterterrorism research to classify the general aims of perpetrators. The goal of the work is to identify the categories of perpetrator aim. The work used the GTD dataset, eleven ML algorithms namely: LR, SVM, SGD, DT, RF, ADB, MLP, NB, kNN, GB, XGB were used for comparison. The performance evaluation procedure used accuracy, precision, recall, F1-score, and ROC-AUC score. The applied approach did not consider the reporting biases is the motives and summary narratives of the GTD. The work established the goal of a terrorist is to bring about change by undermining their target, imposing their will on circumstances, or threatening their opponents and stirring up feelings of hatred, dissent, or retaliation. The study did not look at whether terrorist activity patterns fit into one of the six objectives.

Saidi & Trabelsi (2022) presented hybrid deep learning-based model for future terrorist activities prediction. The authors observed the need to bridge DL models' research gap to increase accuracy and classify terrorist activities based on bi-label and multilabel data sets. The work's goals were to propose a hybrid model that evaluated the performance of two deep learning models—the Long Short-Term Memory (LSTM) and the Convolutional Neural Network (CNN)—for both bi-classification models and multi-classification problems in order to predict terrorist attacks. To predict impending terrorist attacks, the authors proposed a CNN-LSTM hybrid model. When the CNN model learns the local features of the data sets, the LSTM increases its accuracy by extracting the context-dependent features. The model's core elements include a filtered input data set, a fully connected layer, a 1D convolution layer, an LSTM layer, and a classification SoftMax function. The CNN-LSTM model for terrorist activity forecasting consists of one input layer, one pooling layer, convolutional layer, LSTM layer, fully connected dense layers, and an output layer based on a soft-max function. The study used Uddin et al. (2020) as a benchmark to evaluate the modified DNN and the suggested CNN-LSTM. The split ratio of the training and testing is kept at 80:20 and results are generated. The accepted standard for suicide prediction accuracy was 98%. The CNN-LSTM model had the greatest accuracy of 99%, while the DNN model produced 98.6% accuracy. As a result, both models worked effectively. The criterion for predicting assault success was 93%. In weapon type prediction, CNN-LSTM only managed to attain 89.7% accuracy, falling short of the model's 94% benchmark, whereas DNN obtained 99.5% accuracy. The local geographic characteristics of terrorist actions were not taken into account in the work in order to have a greater understanding of forecasting future terrorist activities. The suggested hybrid approach works well on the same dataset, but it performs poorly on more varied datasets. Future research should look into ResNet topologies and transfer learning in the suggested manner to address the problem and further enhance the network's performance.

George *et al.* (2023) designed a model for intelligent pattern recognition to evaluate the activities of terrorists in Nigeria. There is no known regional pattern of armament, attack kinds, or victim types of terrorist operations, and many of the models for assessing terrorist activities are incapable of learning from past patterns to assist

https://doi.org/10.51459/jostir.2025.1.1.010

preventive actions against future occurrences. The goal of the research work was to develop an intelligent model that can identify various terrorist trends in each of Nigeria's six geopolitical zones. Using ANN and a data set of 5,503 occurrences of terrorist activity in Nigeria, a pattern recognition model was constructed with 70%, 15%, and 15% data splits for training, validation, and testing, respectively. The scaled conjugate gradient backpropagation technique was used to create and train a 10-10-6 ANN architecture. Neural networks were used in the suggested pattern recognition system. A neural network called Kohonen Self-Organizing Map (SOP), formerly called Self-Organizing Feature Map (SOFM), was utilized to identify patterns in 400 photos taken from the AT&T database, including human faces. The average percentage scores for accuracy, precision, recall, and F1-score were 99.89%, 99.96%, 100%, and 99.98%, respectively. The models contained a five-layer DNN in addition to three traditional ML algorithms: SVM, naive bayes, and others. Results revealed that DNN beat ANN, logistic regression, SVM, and Naïve Bayes, with scores above 95% in accuracy, precision, recall, and F1-Score, while ANN, SVM, and Naïve Bayes only managed a maximum accuracy of 83%. A two-layer feed forward neural network was constructed with a sigmoid activation function at the hidden layer and a SoftMax activation function at the output layer. This work can be expanded for future research by identifying the patterns of terrorists using additional popular ML technologies, which will give researchers a platform for comparison analysis.

The limitation of Maniraj *et. al.* (2019), Soliman (2019), Sarker *et al.* (2020), Huamaní *et al.* (2020), Babitha *et. al.* (2020), Uddin *et al*. (2021), Abdalsalam *et. al.* (2021), Olabanjo *et al.* (2021), Wang *et al*. (2021), Odeniyi *et al.* (2022), Saidi and Trabelsi (2022), and George *et al.* (2023) are the key motivation for this research work. These includes the failure to appropriate both structure and unstructured datasets with ensemble ML techniques that encompasses the traditional ML models and deep learning model to design a recommender system that will help the user to predict the likelihood of terror attack.

2.0 | Materials and Methods

This work will design bootstrap and aggregation (bagging) ensemble ML model i.e. Logistic regressing, random forest and support vector machine and BiLSTM with BERT embeddings. Figure 1 depict the architecture of the system which is made-up of data collection, ensemble module, deep learning module, network layer and classification/prediction database, application layer. The proposed system will get datasets from global terrorism database that consist of structured data of worldwide terrorist attacks accessible through the uniform resource locator: http://www.start.umd.edu/gtd/download. The second dataset will come from twitter incorporation through the use of snsrape module. This work explores keywords such as "Boko haram, Bandits, Herdsmen, IPOB, Unknown gunmen, Suicide, Bombing" etc. To predict the terrorism activities in Nigeria.

From Figure 1, the ensemble module further expanded in Figure 2, which shows the ensemble of logistic regression, support vector machine, and random forest to explore the GTD structured dataset. The ensemble model learns and predicts from the following factors of terrorism activities.

1. Suicide: 0 = "no," the incident was not suicide attack; 1 = "yes," means otherwise. The dimension of the is 5550 x 24. 80% instances were used for training while 20% instances were used for testing.
2. Success: 1= "yes" incident was successful while 0 = "no" otherwise.

https://doi.org/10.51459/jostir.2025.1.1.010

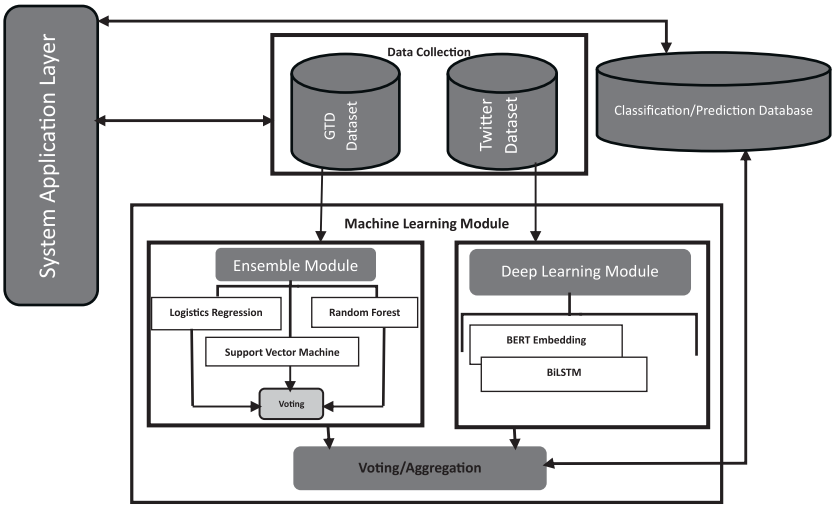

Figure 1 | Architecture of the Recommender System

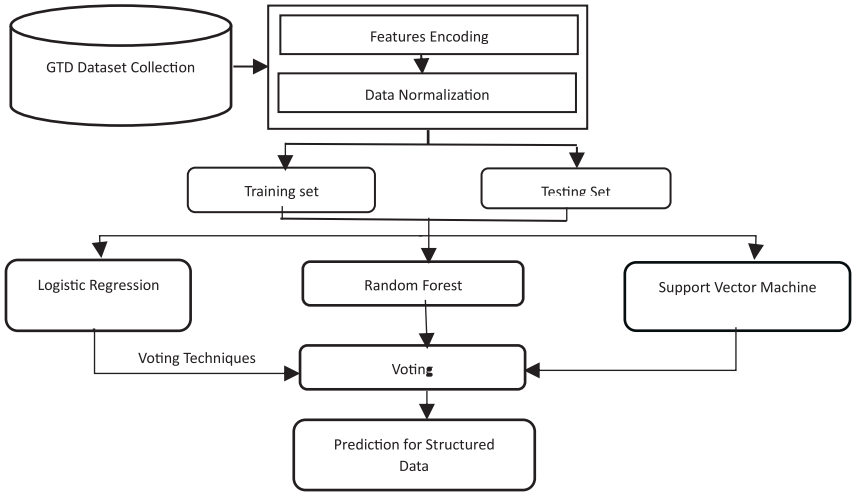

Figure 2 | Architecture of Ensemble Module for Structured Dataset

3. Weapon type: denotes the type of weapon employed for the attack.
4. Geo-political zone: this indicates six different regions in Nigeria.
5. Attack type: attack method.

Features Encoding

DicVectorizer of sklearn library is been proposed to convert non-numeric data to numeric data as label can be hash and compare to numerical labels.

Data Normalization

In this work, the sklearn library's MinMaxScalar function converts data in the range of -1 to 1 using the standardization formula expressed in Equation 1, where X_i are all the samples for a given feature, $\overline{X}$ is the average of all samples by the feature, s and is the standard deviation

$$Z_i = \frac{X_i - \overline{X}}{S} \quad (1)$$

Logistic Regression

Logistic regression divides the probability of an outcome into two classes, which are either class 0 or class 1. It is mathematically represented in Equation (2).

$$y_i = \frac{e^{(a+bX)}}{1+e^{(a+bX)}} \quad (2)$$

https://doi.org/10.51459/jostir.2025.1.1.010

where X = is input data of the sample features from GTD dataset, y_i = predicted output, a and b parameters of the model.

Random Forest (RF)

RF will use the class and probability to calculate the Gini Index of each terror attack from the dataset and predict the likelihood of an attack as shown in Equation (3 and 4).

$$Gini\ Index = 1 - \sum_{i=1}^{c} p_i^2 \quad (3)$$

where, p_i, c denotes probabilty of attack in the observing dataset and number of attacks respectively in Equation 3 and p_+ and p_- denote the probability of the positive attack class and negative attack class respectively in Equation 4.

$$\sum_{i=1}^{c} p_i^2 = 1 - [(p_+^2 + p_-^2)] \quad (4)$$

Support Vector Machine (SVM)

The SVM classifier model predict and classify the class into the new instance x by simply calculating the decision function. Equations (5) – (7) explain each component of SVM.

$$w^T x + b = 0 \quad (5)$$

$$y_i(w^T x + b) = \begin{cases} \geq 0 \\ < 0 \end{cases} \quad (6)$$

where, w - weight vector, T - transpose, b - bias and x - symbolizes the training examples closest to the hyperplane in Equation (5 and 6).

$$y_i = \min \frac{||w||}{2} + c * \sum_{i=1}^{n} \varepsilon_i \quad (7)$$

where, c denotes number of misclassified attacks, ε_i magnitude of the misclassified errors in Equation (7).

Unstructured Data

From Figure 1, the deep learning module further expanded in Figure 3, where Bi-LSTM and BERT explores the unstructured data gathered from the twitter.

Data Pre-Processing

The original tweet is broken up into discrete parts known as tokens, and each token is represented by a unique ID. Tweets are normalized using Feature scaling to ensure all values are in range [0,1] using Equation 8.

$$X_{new} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (8)$$

where, X is the value of the original tweet, X_{new} is the normalized tweet, X_{min} is the minimum value in the tweet, X_{max} is the maximum value in the tweet.

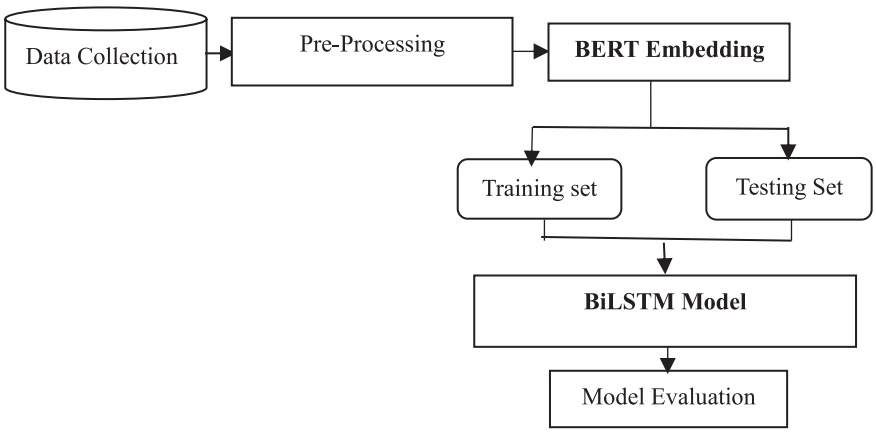

Figure 3 | Architecture of the Deep Learning Model

https://doi.org/10.51459/jostir.2025.1.1.010

BERT

A pre-trained transformer-based model that uses masked language modelling and next sentence prediction tasks by learning contextualized embeddings from words representation. Given a tweet input X that consist a set of elements x words range from x_i to x_n.

$$X = \{x_1, x_2, \ldots, x_n\} \quad (9)$$

where n means the total number of elements (that is, words) in the set *X*, BERT maps *X* into a series of word vectors $\overline{X}$ as shown in Equation 10.

$$\overline{X} = \{e_1, e_2, \ldots, e_n\} \quad (10)$$

The contextual representation of every token, sentence, sentence pair, or paragraph is produced by the BERT. Using its related tokeniser, BERT preprocesses text by frequently dividing words into subwords and other special tokens.

LSTM and BiLSTM

By retaining long-term knowledge, LSTM creates long-term dependency. Each memory cell in the LSTM layer contains four components known as gates: an input gate (i_t), an output gate (o_t), a forget gate (f_t), and a Ct cell state gate, which controls the flow of information into and out of the cell gate. The memory cell itself retains values for arbitrary periods of time. By squashing the output values in the closed range [0,1], the ft uses the sigmoid function (σ) to determine what should be forgotten from the cell state as shown in Equation (11), where 0 means the text is irrelevant to the context and 1 means the text is relevant and should be kept using Equation (12)

$$\sigma(x_t) = SoftMax(x_t) = \frac{e^{x_t}}{\sum_{t=1}^{T} e^{x_t}} \quad (11)$$

$$f_t = \sigma\left(W_f . [h_{t-1}, x_t] + b_f\right) \quad (12)$$

Using a sigmoid function that determines what needs to be changed and observes point-wise product operation with a candidate gate, input gate it determines what new data should be added back to the cell updates of the cell state with word embedding of the tweet x_t in Equation (13).

$$i_t = \sigma(W_i . [h_{t-1}, x_t] + b_i) \quad (13)$$

where, W_i is encoded vector weight, x_t means the tweet data current input, h_{t-1} denotes previous hidden state and h_{t-1}, $x_t \rightarrow \sigma$. σ decides which values should be updated. Recall that σ operation gives values between 0 and 1.

Candidate cell state $\hat{c}_t$ of x_t is mathematically represented as shown in Equation (14).

$$\hat{c}_t = \tanh.(W_{\hat{c}} . [h_{t-1}, x_t] + b_{\hat{c}}) \quad (14)$$

where, *tanh* is the Tanh function that values in range -1 to 1.

Equation (15) regulate the values traversing the network by squishes the values between -1 and 1

$$f(x_t) = tanh(x_t)\frac{e^{x_t} - e^{-x_t}}{e^{x_t} + e^{-x_t}} \quad (15)$$

Equation (16) regulate the values traversing the network by squishes the values between -1 and 1.

$$c_t = f_t * c_{t-1} + i_t * \hat{c}_t \quad (16)$$

where, c_t store and transmit x_t information from one time step to another.

Equation (17), o_t represent the output gate.

$$o_t = \sigma(W_o . [h_{t-1}, x_t] + b_o) \quad (17)$$

Hidden layer h_t of LSTM is obtained in Equation (18);

$$h_t = o_t * c_t \quad (18)$$

where, h_t is the hidden state, o_t is the output gate and c_t is the cell state.

Bi-LSTM collected semantic features in both directions and combined their representations as an output as shown in Equations (19) - (21).

$$\vec{h}_t = LSTM\left(x_t, \vec{h}_{t-1}\right) \quad (19)$$

https://doi.org/10.51459/jostir.2025.1.1.010

$$\overleftarrow{h}_t = LSTM\left(x_t, \overleftarrow{h}_{t-1}\right) \quad (20)$$

$$h_t = \left[\vec{h}_t \oplus \overleftarrow{h}_t\right] \quad (21)$$

Bootstrap and Aggregation (Bagging)

The output from ensemble module and deep learning module are combined using bagging approach, this approach trains independent datasets of different sources using data points and later aggregate the training results by employs majority voting techniques. The output from the bagging serves as inference for the web-based recommender system to assist the security agencies in discharging their duties.

3 | Conclusion

This work presented an ensemble ML based system to enhance the prediction of terrorism activities in Nigeria. In the future, the Python programming language will be used in the future to implement this framework while the web based application will be design for the user to query, analyze and get response from the database of the bagging.

Reference

Abdalsalam, M., Li, C., Dahou, A. and Noor, S. (2021). A Study of the Effects of Textual Features on Prediction of Terrorism Attacks in GTD Dataset. *Engineering Letters*, 29(2).

Adewale, O. S., Jimoh, I. T. Makinde, I. A., Adeleye, S. A. (2024). A Framework of Deep Learning Based Terrorist Classification Model. Journal of Computing, Science &Technology 1(2) 8-15.

Afolabi, O. (2013). Terrorism in Nigeria: Culmination of Economic Disenfranchisement, Social Marginalization and Political Instability.

Bridgelall, R. (2022). An application of natural language processing to classify what terrorists say they want. *Social Sciences*, *11*(1), 23.

Cheong, D. D. (2018). Strategic communication and violent extremism: The importance of state action. *Journal of Asian Security and International Affairs*. *5*(2), 129-148.

Falk, R. (2004). The changing role of global civil society. In *Global Civil Society* (pp. 69-84).

George, U. D., Udoh, S. S., and Obot, O. U. (2023). An intelligent pattern recognition model for assessment of terrorists' activities in Nigeria. *International Journal of Science and Research Archive*, 9(2), 231-244.

Gilbert, J., Hamid, S., Hashem, I. A. T., Ghani, N. A., & Boluwatife, F. F. (2023). The rise of user profiling in social media: review, challenges and future direction. *Social Network Analysis and Mining*, *13*(1), 137.

Huamaní, E. L., Alva, M. A. and Roman-Gonzalez, A. (2020). Machine learning techniques to visualize and predict terrorist attacks worldwide using the global terrorism database. International Journal of Advanced Computer Science and Applications, 11(4).

Höflinger, T. (2021). Modern terrorism: motives of individual terrorists or the strategies of terrorist groups? *Global Change, Peace & Security*, *33*(1), 77-83.

https://doi.org/10.51459/jostir.2025.1.1.010

Iorliam, A., Dugeri, R. U., Akumba, B. O., Otor, S. and Shehu, Y. I. (2021). An Investigation and Insight into Terrorism In Nigeria.

Jimoh, I. T., Adewale, O. S., Kuboye, B. M., Gabrie A. J., Makinde, I. A., (2024). Development of Deep Learning Based System for Terrorism Prevention. Journal of Computing, Science &Technology 1(2) 46-62.

Maniraj, S. P., Chaudhary, D., Deep, V. H. and Singh, V. P. (2019). Data aggregation and terror group prediction using machine learning algorithms. *International Journal of Recent Technology and Engineering*, 8(4), 1467-1469.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M. and Gao, J. (2021). Deep learning--based text classification: a comprehensive review. *ACM computing surveys (CSUR)*, *54*(3), 1-40.

Odeniyi, O. A., Adeosun, M. E. and Ogundunmade, T. P. (2022). Prediction of terrorist activities in Nigeria using machine learning models. Innovations, 71, 87-96.

Olabanjo, O. A., Aribisala, B. S., Mazzara, M. and Wusu, A. S. (2021). An ensemble machine learning model for the prediction of danger zones: Towards a global counter-terrorism. *Soft Computing Letters*, 3, 100020.

Rigterink, A. S. (2021). The wane of command: Evidence on drone strikes and control within terrorist organizations. *American Political Science Review*, *115*(1), 31-50.

Sarker, A., Chakraborty, P., Sha, S. S., Khatun, M., Hasan, M. R. and Banerjee, K. (2020). Improvised technique for analyzing data and detecting terrorist attack using machine learning approach based on twitter data. *Journal of Computer and Communications*, 8(7), 50-62.

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y. and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1631-1642).

Wang, J., Qin, J. H. Xiang, X. Y. Tan, Y. and Pan, N. (2019). CAPTCHA recognition based on deep convolutional neural network. *Mathematical Biosciences and Eng*ineering 16 (5), 5851-5861.

Treistman, J. (2024). Social exclusion and political violence: Multilevel analysis of the justification of terrorism. *Studies in Conflict & Terrorism*, *47*(7), 701-724.

https://doi.org/10.51459/jostir.2025.1.1.010