IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 2 August 2024, accepted 27 September 2024, date of publication 4 November 2024, date of current version 20 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3485913

RESEARCH ARTICLE

Towards AI-Driven Prediction of Terrorism Risk Based on the Analysis of Localized Web News

GEORGIOS KOUTIDIS, KOSTAS LOUMPONIAS, THEODORA TSIKRIKA, STEFANOS VROCHIDIS, (Member, IEEE), AND IOANNIS KOMPATSIARIS, (Senior Member, IEEE)

Information Technologies Institute (ITI), Centre for Research and Technology Hellas (CERTH), GR-57001 Thessaloniki, Greece

Corresponding author: Georgios Koutidis (gkout@iti.gr)

This work has received funding from the European Union Research and Innovation Framework Programme under Horizon Europe project PERIVALLON (Grant Agreement No 101073952) and Horizon 2020 project APPRAISE (Grant Agreement No 101021981). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

ABSTRACT This work presents a framework for predicting the likelihood of terrorist incidents within a specific time period in a target country based on machine learning and deep learning models. These models are trained on news data for the target country sourced from the Global Database of Events, Language, and Tone (GDELT), while the terrorist incidents recorded in the Global Terrorist Database (GTD) serve as the ground-truth. To cope with the massive volume of such localised news and the features extracted from them, a novel feature selection process based on Random Forest models is proposed, where the predictive power of each feature is calculated and only features estimated to have sufficient predictive power are subsequently considered. Experiments are performed by considering the United Kingdom as a case study and six predictive models are evaluated across four evaluation metrics. The experimental evaluation results indicate that the proposed framework enhances the predictive power of both machine learning and deep learning models that use the proposed feature selection in forecasting the likelihood of terrorist incidents, surpassing models trained 1) on all baseline features and 2) on features selected by two well established feature selection approaches. Overall, the proposed framework aims at assisting counter-terrorism efforts by providing a projection of the likelihood of terrorist incidents and thus the overall terrorism risk in a target country, using the most valuable features provided by GDELT.

INDEX TERMS Terrorist incidents, predictive models, deep learning, machine learning, feature selection, global terrorism database (GTD), global database of events, language, and tone (GDELT).

I. INTRODUCTION

Terrorism has caused substantial economic, political, and humanitarian crises worldwide, along with thousands of casualties, during the past 50 years. Significant effort has been placed on the study of the many facets of terrorism and the multiple dimensions of counter-terrorism, with particular focus on geographical areas with many terrorist activities, such as the Middle East and Asia [1], as well as other parts of the world, including Europe, where the impact of terrorism has also been considerable [2]. To enhance counter-terrorism efforts, the ability to predict (forecast) in an effective and efficient manner the likelihood of the occurrence of future terrorist incidents, as well as their characteristics, is of the essence.

To this end, many strategies and approaches have been proposed to counter terrorism and its effects based on the prediction of terrorist incidents through the analysis of data primarily in the Global Terrorism Database (GTD) provided by the National Consortium for the Study of Terrorism And Responses of Terrorism (START) [3]. GTD contains more than 200,000 instances of terrorist incidents around the world since 1970, making it the most comprehensive real-world terrorism-related dataset. Each terrorist incident is described by 145 variables, which include information about the date/time of the incident, the number of casualties, the

The associate editor coordinating the review of this manuscript and approving it for publication was Ahmed Farouk.

type of weapons used, the nature of the target, as well as the group responsible for the terrorist attack. This database has been used effectively for predicting several aspects of future terrorist incidents given historical information, such as predicting the group or organisation responsible [4], [5], whether an attack will be a success or not, whether it will be a suicide attack or not, the type of attack, the type of weapon employed, and the region where the attack will take place [6]. As, though, the available variables mainly describe the aftermath of an attack and do not reflect the overall situation in a region and what is taking place in the time period before an attack occurs, there are, to the best of our knowledge, no successful studies on the use of GTD for predicting the likelihood of occurrence of future terrorist incidents.

For this purpose, the utilisation of news data has been proposed in order to uncover more sophisticated and explanatory variables, since socio-political and economic instability may be a precursor to potential terrorist strikes [7]. To this end, the Global Database of Events, Language, and Tone (GDELT) [8], a publicly available collection of news sources since 1979, in over 100 languages from around the world, has been used [7], [9]. This allows for a more comprehensive analysis that has the potential to enable the prediction of future terrorist incidents. It should be noted though that the extensive volume of global news data in GDELT, along with their inherent heterogeneity, given that they encompass multiple aspects beyond the provision of terrorism-related insights, necessitate robust selection of both the data and variables that would be useful towards the predictions of terrorism risk.

In this context, this work proposes a framework for predicting the likelihood of future terrorist incidents that may take place within a specific time period in a specific target country based on machine learning (ML) and deep learning (DL) models trained on localised news data sourced from GDELT. To select the most relevant features from the news data, a feature selection process is proposed in which the predictive power of individual features is evaluated using a Random Forest (RF) model [10], as RFs excel in predictive capabilities [7], [11], [12]. If a feature alone is found to enhance the forecasting results, it is retained and grouped with other selected features to improve the overall predictive performance. This use of RF allows for the identification of the most important features related to terrorist incidents and the selection of these features for further analysis. Following the process of feature selection, ML and DL methods are employed to forecast the likelihood of a terrorist incident. This approach aims to focus on the most critical information and enhance the accuracy of predictions. The terrorist incidents provided by GTD are used as ground-truth by the RF in the feature selection stage, as well as to train and evaluate the ML- and DL-based models that provide forecasts for potential terrorist incidents.

Overall, this work primarily contributes a framework for predicting the likelihood of future terrorist incidents and assessing terrorism risk by integrating data from diverse sources, namely the data from GDELT and the terrorist incident records from GTD. The main novelty lies in the incorporation of feature selection, which focuses on the individual predictive power of each feature. This is accomplished by identifying, selecting, and employing specific features from the GDELT in a novel manner using RF models. The proposed Artificial Intelligence (AI)-driven framework then employs a wide range of ML and DL models for providing a prediction and is evaluated using data related to a European country. The feature set resulting from the novel feature selection methodology proposed in this work demonstrates superior performance compared to other feature selection techniques and also against the baseline feature set that encompasses the entirety of information from the GDELT. This outcome underscores the significance of the proposed feature selection and the originality of the overall approach presented in this work. In general, the estimated risk level based on the forecast of future terrorist incidents using AI methods has the potential to assist counter-terrorism efforts by providing a projection of the likelihood of such an event, thus assisting in providing alerts regarding areas and populations that may be more vulnerable, and also highlighting the need for further investigation and analysis of factors, such as targets and motives.

The rest of this paper is organised as follows: Section II provides an overview of related work on the use of the GTD and GDELT datasets for predicting different aspects of terrorist incidents. Section III details the proposed framework and its mathematical formulation. The experimental evaluation is presented in Section IV, while Section V discusses the main findings, as well as the limitations of this work. Finally, Section VI presents our conclusions and outlines directions for future research.

II. RELATED WORK

A. GTD DATABASE

Several works have used GTD to predict various aspects of terrorist incidents. In [5] a Terrorist Group Prediction Model (TGPM) was proposed for identifying terrorist groups likely to carry out attacks in the near future, using the GTD India data from 1998 to 2008. The TGPM identifies groups likely to carry out attacks based on the past behaviour of several parameters, including attack type, location, target type, weapon type, etc. When all parameters were given equal weight, the model achieved an overall performance of around 66%; however, when the weight of the location parameter was increased, the performance reached around 80%.

In [4], a data mining approach was presented for predicting the likelihood of terrorist attacks by specific groups based on an ensemble predictive model relying on a combination of four ML models: Naïve Bayes (NB), K-Nearest Neighbour, Iterative Dichotomiser 3, and decision stump, for identifying patterns in the data indicative of predicting the responsible terrorist group. When applied to the GTD data, the proposed majority vote classifier outperformed the baseline classifiers, reaching almost 94% accuracy.

A hybrid classification approach was presented in [1] for predicting terrorist groups likely to carry out attacks in the Middle East and North Africa regions between 2009 and 2013. The developed ML model combines multiple classification algorithms to improve the accuracy of predictions. A comparison was carried out between various classification models including Hybrid Hoeffding Tree, Functional Tree, Hybrid Naive Bayes, and Decision Tree (DT), as well as Support Vector Machine (SVM), RF, Naive Bayes (NB), and stacking classifiers. The results indicated that the hybrid classifiers performed better than the standard methods and ensemble techniques across all evaluation metrics.

In [13], the focus was on the use of ML techniques, specifically SVM, NB, and logistic regression (LR), to predict the type of potential terrorist attacks (political, separatist, religional, and gangdom) using data from GTD. Two feature selection methods, minimal-redundancy maximal-relevancy and maximal relevance, were used to improve the accuracy of their predictions. The results showed that the LR model with a subset of seven optimal features achieved a classification precision of around 78%, indicating that feature selection can reduce classification errors.

The use of a deep neural network (DNN) model was proposed in [6] to predict various aspects related to terrorist incidents, such as whether an attack will be a success or not, whether it will be a suicide attack or not, the type of attack, the type of weapon employed, and the region where the attack took place. The proposed DNN model was compared to a single-layer neural network (NN) and three traditional ML algorithms (LR, SVM, and NB) and the results showed that the DNN model performs significantly better, with an accuracy of over 95%, while the other models had a maximum accuracy of 83%.

In [14], advanced gradient decision trees (XGBoost) [15] were used for predicting whether a terrorist attack will lead to casualties. The algorithm incorporates RF and principal component analysis (PCA) for feature selection and uses a genetic algorithm to optimise the hyperparameters of XGBoost. The proposed model achieved 94% accuracy on GTD data related to China and proved its higher generalisation ability by achieving 87% accuracy on data from other countries.

In [16], the application of temporal meta-graphs combined with bi-directional LSTM (Long Short-Term Memory) models was introduced to predict future terrorist targets. Using GTD data on terrorist attacks in Afghanistan and Iraq in 2001–2018 and focusing on three facets (weapons used, tactics employed, and targets selected), these meta-graphs draw parallels between temporally adjacent attacks, highlighting operational correlations. Based on these, two-day interval time series were created to track the prominence of each attribute over time. This study underscored the benefits of temporal meta-graphs that offer deeper insights compared to conventional time-series based on shallow feature frequency.

In [12] , the GTD was utilised alongside data from the International Country Risk Guide (ICRG) [17] and WorldBank [18] to predict annual human and property loss, "successful" terrorist incidents, and the likelihood of more than one severe terrorist incident taking place. A recursive feature elimination method with RF kernels (RF-RFE) was proposed to identify crucial features and several ML (SVM, LR, RF) and DL (LSTM, full-connected DNN) models leveraged these crucial (selected) features to provide predictions. The RF model provided the best overall performance according to mean absolute error; it should be noted though that the minimal set features used for understanding the risk of terrorist attacks did not at the end include any features provided by GTD, but only from the other two sources.

In [19], the attention-based spatial-temporal multigraph convolutional network (AST-MGCN) was proposed to predict the daily civilian casualties per region (e.g. West Europe, North America, etc.) caused by terrorist groups. To that end, the proposed model took into account all worldwide terrorist incidents listed in GTD and used the wavelet transform as a preliminary step to extract the temporal dynamics information of these terrorist incidents, such as their trend, period, and proximity. Then, the proposed AST-MGCN used a spatial multi-graph convolution to capture rich social-spatial features in multi-views and a temporal convolution to capture time-dependent relationships. The final prediction resulted from fusing the three trainable temporal-dimensional components (trend, period, proximity).

B. GDELT DATABASE AND OTHER NEWS SOURCES

Several works have also proposed the use of the GDELT for predicting different aspects of terrorist incidents. In particular, two case studies using the GDELT data were presented on the Sri Lankan civil war and the 2006 Fijian coup [20]. These case studies demonstrated how change point analysis can be used to identify important trends in the data, along with a conceptual model for a multi-level analysis dashboard system that could be used by peacekeeping organisations.

A graph-based approach was presented in [21] for detecting and forecasting domestic political crises. By building a network model of the political system in a country, with nodes representing political actors and edges representing relationships between them, key political actors could be identified and their interactions could be analysed to detect early warning signs of a crisis. A least absolute shrinkage and selection operator regression model was also developed to forecast at a monthly level the likelihood of a crisis occurring, using GDELT data of Latin American countries.

In [22], the development of a real-time data collection system and various methodologies was described for predicting and projecting the risk of terrorist incidents. The data collection system gathers terrorist incident information from reliable sources (such as CNN and NY times) based on a crawler [23], while the risk model uses this information to calculate the terrorism risk level for Mali, Yemen, Nigeria, Kenya, and Libya respectively. A manual set of rules and a risk projection model were developed to predict and project the risk of future terrorist incidents within the next 10 days.

In [24], a framework was presented for predicting indicators of country instability using posts/articles of protests encoded in GDELT. The framework used Hidden Markov Models to analyse the temporal burst patterns in GDELT event streams and formulate the prediction of social unrest events as a sequence classification problem based on Bayes decision. The framework was tested on data from five countries in Southeast Asia and was found to be more effective than the LR model and baseline methods.

The use of the Autoregressive Integrated Moving Average (ARIMA) model and the Markov Switching-regime model was employed in [25] to forecast the number of mass incidents (large-scale disturbances or protests) in mainland China in the short term. The study found that the two-regime Markov switching model can achieve a more accurate prediction of mass incidents than the ARIMA model alone.

In [26], ML techniques (such as RF, Boosting, and DNN) were used to understand and predict instances of social unrest in the United States. The authors used news articles from GDELT to study the factors that contribute to such events at the state and country level. They found that the volume of news articles with negative sentiment increased after events that led to major civil unrest. Experimental evaluation indicated that RF achieved the best results on the country level.

In [7], GDELT data were leveraged to forecast whether there will be a terrorist attack within the next few days in five different states of the USA. In order to provide forecasts for a particular day, authors considered features (including from GDELT) regarding the previous days. Furthermore, the optimal number of prior days (time window) was defined according to a two-sample Kolmogorov-Smirnov test. Finally, the RF algorithm and the Multilayer perceptron (MLP) model were used to provide forecasts and the proposed approach was evaluated by only one metric, the ROC-AUC score [27].

C. DISCUSSION

Overall, GTD has been primarily used for predicting various aspects related to terrorist attacks, rather than their occurrence. For instance, [1], [4], and [5] predict the terrorist group responsible for carrying out the attack, while [13] and [6] focus on the prediction of the type of attack, and [6] on other characteristics, such as the success of an attack, its type, and the region. Similarly, in more recent studies based only on GTD, [14] and [19] predict whether a terrorist incident will cause causalities, while [16] forecast future targets.

Although these approaches are undoubtedly beneficial, they inherently lack the capacity to deliver predictive information to relevant stakeholders related to the potential occurrence of terrorist incidents. Thus, they fail to offer preventative insights that could potentially forestall such incidents from happening in the first place. To that end, an alternative way is the tracking of socio-political instability from online news sources (such as the GDELT) in order to identify trends [20], predict political crises [21], social unrest [26], country instability [24], and mass incidents [25], and overall project terrorism risk level [7], [22], [23].

In this context, the framework proposed in this work employs a novel feature selection method to the GDELT in order to identify a set of features that aim to predict the likelihood of occurrence of terrorist incidents within a specific time period in a target country, while it employs GTD as ground truth. Compared to previous work and in particular to the approach by Krieg et al. [7] which has similar objectives to our work, the proposed framework selectively utilises specific features of GDELT based on a novel methodology, as opposed to employing all available features. Moreover, the proposed approach considers additional ML and DL models for providing the prediction and is evaluated more robustly across four evaluation metrics, thus increasing the reliability of the findings. Finally, as substantiated in the experimental evaluation (see Section IV), the proposed approach outperforms the baseline methodology and other feature selection methods, thus highlighting the efficacy of the proposed feature selection and usage in predicting the likelihood of potential terrorist incidents.

III. PROPOSED METHODOLOGY: TERRORIST INCIDENTS PREDICTION FRAMEWORK

This section presents in detail the proposed framework for predicting terrorist incidents likely to occur within a specific number of days following a given date, for a target country, based on ML and DL models trained on GDELT data associated with the target country. The proposed framework consists of three main stages: (i) data collection from GDELT, (ii) feature selection, and (iii) prediction of the likelihood of terrorist incidents using ML and DL methods. More specifically, in the first stage, features (measurements) provided by GDELT that refer to the country of interest are acquired. Next, the RF model is applied for each feature (obtained from the previous stage) to compute its predictive power and the features with the higher scores are selected in order to improve the forecasting performance. Finally, ML and DL models are employed for predicting the likelihood of occurrence of terrorist incidents within the next days; these models are trained by considering the GTD data as ground truth. The proposed framework is illustrated in Figure 1.

A. DATA COLLECTION

In this work, the latest version 2.0 of GDELT is used, where all online news translated to English are available from February 2015 onwards. GDELT contains two major sources of information: the Event Database and the Knowledge Graph that attempt to create a global representation of *what is happening* and *how the world feels about it*, respectively.

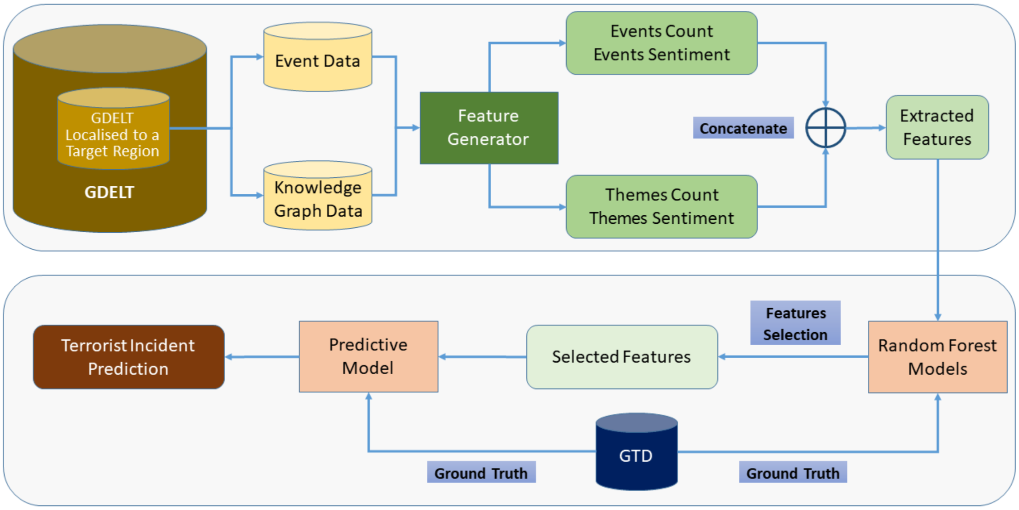


FIGURE 1. Framework of the proposed approach for predicting the likelihood of terrorist incidents within a specific time period in a target country.

The *GDELT Event Database* is a publicly available dataset that records events around the world that have been extracted from news articles and other sources. The dataset is structured with five primary fields: location, date, event type, and detailed information about the two countries involved in an event of any type, referred to as "actor 1" and "actor 2", following the Conflict and Mediation Event Observations (CAMEO) taxonomy [28]. In particular, the "Actor1CountryCode" field contains the country code for the primary actor involved in an event, "actor1KnownGroupCode" contains a code if a known organisation took part in that event (e.g., United Nations, World Bank, al-Qaeda, etc), while the "EventCode" field contains codes describing the type of event. There are a total of 196 unique event codes (e.g., code 1721 corresponds to "Impose restrictions on political freedoms") [28] from which the *Event Count* and *Event Sentiment* features are generated. The Events Count feature counts occurrences of each base code each day, while the Events Sentiment feature provides an average sentiment score of all articles associated with each event code aggregated daily.

The *GDELT Knowledge Graph* (GKG) is a database that represents up-to-date information about global events, organisations, people, and countries, as well as the relationships between these entities. It includes a wide range of data extracted from news articles, social media, and other sources using ML algorithms. The GKG database is updated every fifteen minutes and each entry includes references to people and locations, textual themes, and other metadata extracted from each published article, video, or other news item. In this work, the *textual themes*, which mention the country of interest (target country), are considered. The extracted textual themes are described in the GKG Category List [29]. The GKG database provides information on the published date, 284 textual themes (e.g., the "*Terror*" theme), sentiment score, and location of each record. From this information, two feature categories, *Themes Count* and *Themes Sentiment* are generated [7], [8] . The Themes Count feature corresponds to the number of records mentioning a certain theme for each date, while the Themes Sentiment feature provides the average sentiment score of each theme per day based on the records of that date. This allows for analysis of the frequency and sentiment of specific themes over time and in specific locations.

The features derived from GDELT, including themes count, themes sentiment, events count, and events sentiment, provide information on a daily basis. From now on, the extracted features on a daily basis will be denoted as $\{x_{i,j}\}_{i=1:n,\, j=1:T}$, where n stands for the total number of extracted features and T for the total number of days. However, due to the sparsity of terrorist incidents (i.e., their infrequent occurrence), the dynamic predictive power of the derived features when considered on a per-day basis could be weak and therefore likely to result in poor performance. To overcome this limitation, the selected features are aggregated over w-day(s) spans. Specifically, the average value of the derived features is calculated for each w-days period. Then, w-days periods for which at least one terrorist incident is recorded in the GTD are labelled as 1, while all other periods are labelled as 0. The aggregated features (without loss of generality) are denoted as $\{x_{i,j}\}_{i=1:n,\, j=1:T_w}$, where T_w stands for the total number of w-days periods.

As GDELT is an expansive source of information that encompasses numerous elements that may not be inherently associated with terrorist incidents, the framework introduced in this work proposes a feature selection process, which is tailored to capture and utilise only the most relevant aspects from the vast array of GDELT data, as described next.

B. FEATURE SELECTION

Identifying the most characteristic variables is crucial for enhancing prediction accuracy. The proposed approach involves the design of a procedure for selecting the features employed by models trained towards predicting the likelihood of the occurrence of terrorist incidents within the next w-days in the target country. To that end, each feature $\{x_{i,j}\}_{j=1:T_w}$ is considered as a stand-alone predictive variable and the predictive power of each feature is calculated and evaluated.

To calculate the predictive power of each feature, the RF model is used as follows:

$$\hat{y}^i_{j+1} = RF(x_{i,j-dt:j}) \tag{1}$$

where $\hat{y}^i_{j+1}$ represents the prediction output at time $j+1$ using only the feature i, while dt denotes an observation window that corresponds to the number of previous w-days spans considered as historical information. The performance of each feature is evaluated using the average of precision and recall. Next, only those features that provide, for at least one dt value, an average greater than a threshold θ are considered. In this way, an initial examination of the predictive power of each feature is performed, while the main goal is to select the most relevant ones regarding the prediction of the likelihood of terrorist incidents.

The RF model has been selected as it is a well established method that has been shown to outperform other ML and DL models in various predictive tasks (e.g. [11]), including when used for predicting the likelihood of occurrence of terrorist incidents (e.g., [7], [12]). Here, RF is one of the predictive models under consideration (see Section III-C), but it is also employed in a different manner at the feature selection stage for estimating the predictive power of each feature provided in GDELT by following the novel methodology described above. The proposed feature selection approach is also compared against two other well established feature selection techniques: the Permutation Feature Importance (PFI) [10] and the Recursive Feature Elimination with cross-validation (RFE) [12], [30]. PFI assesses feature importance by measuring the increase in prediction error when a feature's value is randomly permuted, while RFE recursively fits a RF model, ranking features based on their importance and eliminating the least important ones using cross-validation, until the desired number of features is reached.

The crucial role of these derived features is their application in detecting socio-political unrest - a significant predictor of potential future terrorist incidents. By monitoring signs of such unrest, the framework effectively leverages these derived features to anticipate and forewarn of the likelihood of possible future terrorist threats using the predictive models described next.

C. PREDICTIVE MODELS

The features selected using the approach described above are denoted as $\{x'_{i,j}\}_{i=1:m,\, j=1:T_w}$, where m stands for the total number of the selected features and T_w is defined as the total number of w-days (same as above). Then, several ML and DL methods are employed to provide a forecast of possible future terrorist incidents (i.e., a binary classification problem) using the selected features from GDELT. In particular, the following ML models are employed:

1) LR

This model is instrumental in estimating the likelihood of binary outcomes, utilising the logistic function to model the dependent variable when responses are dichotomous.

2) SVM

Known for its robustness and versatility, SVM performs classification tasks by establishing a hyperplane or a set of hyperplanes in a high-dimensional space, which maximises the margin between different classes [13].

3) RF

This ensemble method enhances prediction accuracy through numerous DTs at training, and outputs the class that is the mode of the classes predicted by the individual trees [11], [12].

4) HYBRID ENSEMBLE MODEL (HYBRID)

Integrating 20 weaker learners, this ensemble model comprises five DTs, five LR models, five SVM, and five NB classifiers, all working collectively on a binary classification problem [1].

Additionally, the following DL models are considered:

5) GATED RECURRENT UNIT (GRU)

GRUs are adept at capturing dependencies in time-series data, benefiting from gating mechanisms that manage the information flow within the unit, thus enhancing the analysis of time-dependent data [31].

6) MLP

This basic yet powerful neural network architecture features interconnected layers and nonlinear activation functions, enabling it to identify complex patterns and relationships within the data [32].

The ML and DL techniques employed in this work form a comprehensive framework, each contributing uniquely to our overarching goal of effectively forecasting terrorist incidents using selected features from the GDELT dataset. All models are evaluated based on their ability to predict whether a terrorist incident is likely to occur within the next w-days. The scope of this paper is to enhance the performance of the aforementioned ML and DL models in predicting terrorist incidents through the proposed feature selection process.

In all methods, except from the GRU, the selected input features are handled as follows:

$$\hat{y}_{j+1} = method(average(x'_{i:m,j-dt:j})) \tag{2}$$

where $average(x'_{i:m,j-dt:j})$ represents the average value of the selected features in the observation window dt, thus the input data are aggregated to form a single vector to be processed from the ML/DL models. In the GRU model, the predictions are computed as follows:

$$\hat{y}_{j+1}^{GRU} = GRU(x'_{i:m,j-dt}, \ldots, x'_{i:m,j}). \tag{3}$$

In this way, the GRU model exploits the temporal dependencies of the features within the observation window dt.

IV. EXPERIMENTAL EVALUATION

In this section, the proposed RF-based framework (PRF) is evaluated on terrorist incidents regarding Europe and in particular the United Kingdom (UK). The UK has been selected as the target country for our evaluation experiments as it is the country most affected by terrorism in Western Europe. During the 2016-2020 period, for instance, the UK experienced the highest number of terrorist incidents in Europe (503 in total), while Germany and France, two major European countries, had significantly fewer incidents (147 and 125, respectively), indicating the variance in the occurrence of terrorist incidents in Europe [33].

In our experiments, three feature selection methods, i.e., PRF, PFI, and RFE, and six predictive models, i.e., RF, Hybrid, SVM, LR, MLP, and GRU, are evaluated. The ML and DL models that employ a feature selection method are indicated by the respective superscript "*PRF*", "*PFI*" and "*RFE*"; otherwise, no superscript is used. The primary objective of these experiments is to determine the performance of the proposed framework in comparison to the baseline methodology described in [7], which uses all features provided by GDELT and also in comparison to the two feature selection methods (PFI and RFE) as implemented in the *sklearn* python library [34]. To that end, three sets of experiments are conducted. First, the *event* features derived from the GDELT Event Database (whole set) and the selected features (derived using the PRF, PFI, and RFE methods) are used to predict terrorist incidents within the next w-days. Then, in the same way, the *theme* features derived from GKG (whole set) and the selected theme features are used to provide predictions for the next w-days. Finally, *all* features from both databases (i.e., GDELT Event and GKG) and the selected features from both sources are utilised. Table 1 illustrates the total number of features (provided by GDELT) for the UK case, which are used by the baseline methodology [7], as well as those selected by the PRF, PFI, and RFE approach, respectively.

TABLE 1. Number of event and theme features provided by GDELT (and employed by the baseline method in [7]) and selected by the PRF framework and the PFI and RFE methods.

	All	PRF	PFI	RFE
Number of Event features	392	64	8	32
Number of Theme features	568	70	32	132
Total number of features	960	134	40	164

A. EXPERIMENTAL SETUP

The evaluation is performed based on online news in GDELT from January 1, 2016 to July 31, 2021 that mention the target country (namely the UK). The average number of terrorist incidents listed in GTD that took place in the UK during this time period (January 1, 2016 to July 31, 2021) is 0.19 incidents/day. As mentioned in Section III-A, the forecasting of terrorist incidents by considering an 1-day span is likely to provide poor performance, due to the sparsity of terrorist incidents [16]. Therefore, the selected features are aggregated over a $w = 2$-days span; by doing so, the average number of incidents increases to 0.31 incidents/2-days. For higher values of w, the average number of terrorist incidents per w-days will increase even more (e.g. 0.48 incidents/3-days, 0.57 incidents/4-days etc.). Consequently, choosing large values for w will not provide meaningful predictions, since in most cases of w-days there will be at least one terrorist incident occurring. Thus, $w = 2$ seems to be a valid and appropriate choice for the UK case.

The evaluation of the proposed framework is performed using a training, a validation, and a test set, which constitute 75%, 15%, and 10% of the aforementioned dataset, respectively. The performance of all models is evaluated using the weighted ROC-AUC score, the F1-score, Precision, and Recall [27].

Next, the parameters of the evaluated ML and DL models are described. In particular, only the parameters that provide the best performance in the validation set are reported.

- **RF**: a collection of 1000 unpruned DTs are used in the*sklearn* python library based on the formula below:

$$\hat{y}(\mathbf{x}) = \text{mode}\left(\{T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots, T_{1000}(\mathbf{x})\}\right) \tag{4}$$

  where $T_i(\mathbf{x})$ represents the prediction of the i-th tree for the input $\mathbf{x}$.
- **Hybrid**: a collection of 20 weak learners is used to build a hybrid ensemble learning model. More precisely, 5 DT, 5 LR, 5 SVM, and 5 NB classifiers are grouped together to work on a binary classification problem. Let M represent the hybrid ensemble containing DT, LR, SVM, and NB classifiers. For input $\mathbf{x}$, the prediction $\hat{y}(\mathbf{x})$ from the ensemble is computed as the majority vote (or average probability for binary classification) from these models:

$$\hat{y}(\mathbf{x}) = \begin{cases} 1 & \text{if } \frac{1}{20}\sum_{m_i \in M} m_i(\mathbf{x}) \geq 0.5, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

- **SVM**: the default parameters in the *sklearn* python library are used. The decision function for an SVM in a binary classification problem is given by:

$$f(\mathbf{x}) = \text{sgn}\left(\sum_{i=1}^{N} \alpha_i y_i \langle \mathbf{x}_i, \mathbf{x} \rangle + b\right) \tag{6}$$

  where $\mathbf{x}$ is the input vector, N denotes support vector count, α_i symbolises Lagrange multipliers, and y_i

reflects class labels (0, 1). The term $\langle \mathbf{x}_i, \mathbf{x} \rangle$ indicates the inner product between support vector $\mathbf{x}_i$ and $\mathbf{x}$. The bias is given by b and the class is determined via sgn($\cdot$).

- **LR**: the default parameters in the *sklearn* python library are used. The LR model predicts the probability for binary classification as:

$$P(y = 1|\mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{w}^T\mathbf{x}+b)}} \quad (7)$$

where $\mathbf{w}$ is the weight vector, $\mathbf{x}$ the feature vector, and b the bias.

- **GRU**: includes a GRU layer with 25 units and one hidden layer with 1 unit and sigmoid activation function (binary classification). In addition, a grid search was performed over [10, 40] with step 5 to determine the number of hidden units. The operations per layer are as follows: The simplified GRU model for binary classification includes:

$$\begin{aligned} \mathbf{h}_t &= \mathrm{GRU}(\mathbf{x}_t, \mathbf{h}_{t-1}) \\ y_t &= \sigma(\mathbf{w}_{hy}\mathbf{h}_t + b_y) \end{aligned} \quad (8)$$

where $\mathbf{h}_t$ is the GRU hidden state, σ is the sigmoid function, and y_t is the output at time t.

- **MLP**: includes one hidden layer (dense) with 25 units with ReLU activation function [35], while the output layer consists of one unit with sigmoid activation function. In addition, a grid search was performed over [10, 40] with step 5 to determine the number of hidden units. The operations per layer are as follows:

$$\begin{aligned} \mathbf{h} &= \mathrm{ReLU}(\mathbf{W}_h\mathbf{x} + \mathbf{b}_h) \\ y &= \sigma(\mathbf{w}_o^T\mathbf{h} + b_o) \end{aligned} \quad (9)$$

where $\mathbf{x}$ is the input vector, $\mathbf{W}_h$ and $\mathbf{w}_o$ are weight matrices for the hidden and output layer, respectively, $\mathbf{b}_h$ and b_o are bias terms, $\mathbf{h}$ represents the hidden layer activations, ReLU is the activation function for hidden units, and σ denotes the sigmoid function for output.

The Adam optimisation process [36] is used for 20 epochs for the GRU and MLP model.

As explained earlier, three experiments are conducted using the aforementioned predictive models (RF, Hybrid, SVM, LR, MLP and GRU). Each experiment is conducted 10 times per model and then the average values of the evaluation metrics are calculated. In addition, different values of the observation window $dt = 1, \ldots, 15$ (in Equations (2), (3)) are used for each experiment, while the threshold θ is set to $\theta = 0.75$. The results for the best-performing observation window dt (in the validation set) according to the ROC-AUC score are presented.

B. EXPERIMENTAL RESULTS

Table 2 displays the results of using the selected *event* features from the PRF approach (64 in total), from the two other feature selection techniques PFI and RFE (8 and 32 features, respectively), and all event features provided by GDELT (392 in total), to predict terrorist incidents within the next 2-days. The proposed method improves the performance of all models, except the Recall of the baseline method for the RF model, which is equal to 0.780, and the Recall of the RF^{RFE} method, which is the best overall performance regarding the Recall metric at 0.793. However, the RF^{PRF} performs better in 3 out of 4 evaluation metrics with ROC-AUC score 0.615, F1 score 0.734, and Precision 0.737. Moreover, all the ML and DL models using the event features selected by the PRF method (except GRU^{PRF}) outperform all other baseline models, while this does not hold for the other feature selection techniques (PFI and RFE).

TABLE 2. Evaluation results based on *Event* features.

Method	dt	ROC-AUC	F1	Precision	Recall
RF	11	0.500	0.683	0.608	0.780
Hybrid	3	0.471	0.624	0.648	0.606
SVM	3	0.546	0.547	0.698	0.505
LR	3	0.416	0.513	0.612	0.464
MLP	4	0.476	0.571	0.648	0.530
GRU	7	0.600	0.624	0.722	0.589
RF^{PRF}	1	**0.615**	**0.734**	**0.737**	0.732
Hybrid^{PRF}	1	0.580	0.707	0.712	0.702
SVM^{PRF}	1	0.607	0.707	0.726	0.693
LR^{PRF}	1	0.607	0.707	0.726	0.693
MLP^{PRF}	1	0.563	0.727	0.714	0.752
GRU^{PRF}	13	0.498	0.606	0.663	0.573
RF^{PFI}	4	0.603	0.654	0.728	0.622
Hybrid^{PFI}	3	0.524	0.698	0.684	0.717
SVM^{PFI}	3	0.524	0.698	0.684	0.717
LR^{PFI}	10	0.522	0.689	0.670	0.728
MLP^{PFI}	11	0.555	0.643	0.692	0.615
GRU^{PFI}	11	0.465	0.386	0.622	0.362
RF^{RFE}	15	0.500	0.701	0.629	**0.793**
Hybrid^{RFE}	3	0.525	0.655	0.681	0.636
SVM^{RFE}	11	0.505	0.654	0.660	0.648
LR^{RFE}	11	0.512	0.662	0.665	0.659
MLP^{RFE}	14	0.494	0.597	0.671	0.556
GRU^{RFE}	12	0.520	0.613	0.679	0.577

In Table 3, the results using the selected *theme* features (based on PRF, PFI, and RFE) and all theme features provided by GKG are presented. The proposed PRF approach provides the best performance regarding Precision and Recall in predicting the likelihood of terrorist incidents within the next 2-days. More specifically, the SVM^{PRF} and RF^{PRF} models provide the best performance regarding Precision (0.758) and Recall (0.793), respectively, while the MLP^{PFI} and Hybrid^{PFI} models provide the best performance regarding the ROC-AUC (0.592) and F1 score (0.704). In general, the features derived from the PRF and PFI methods provide the best overall performance. It is also worth mentioning that the selected *theme* features perform worse (except for the Precision metric) than those using only the selected *event* features (see Table 2).

Table 4 presents the results using the event and theme features selected by the three feature selection techniques and all event and theme features provided by the GDELT Event

TABLE 3. **Evaluation results based on *Theme* features.**

Method	dt	ROC-AUC	F1	Precision	Recall
RF	15	0.458	0.619	0.647	0.597
Hybrid	6	0.435	0.592	0.617	0.572
SVM	3	0.428	0.590	0.622	0.565
LR	14	0.505	0.702	0.679	0.738
MLP	14	0.449	0.656	0.634	0.634
GRU	7	0.474	0.517	0.637	0.473
RF^{PRF}	15	0.500	0.701	0.629	**0.793**
$Hybrid^{PRF}$	10	0.486	0.653	0.636	0.673
SVM^{PRF}	15	0.567	0.383	**0.758**	0.379
LR^{PRF}	6	0.481	0.645	0.645	0.645
MLP^{PRF}	6	0.540	0.310	0.727	0.336
GRU^{PRF}	11	0.509	0.640	0.680	0.613
RF^{PFI}	10	0.534	0.683	0.674	0.695
$Hybrid^{PFI}$	11	0.521	**0.704**	0.688	0.717
SVM^{PFI}	4	0.476	0.665	0.643	0.693
LR^{PFI}	11	0.514	0.697	0.675	0.747
MLP^{PFI}	14	**0.592**	0.676	0.729	0.647
GRU^{PFI}	4	0.614	0.646	0.736	0.612
RF^{RFE}	3	0.467	0.376	0.632	0.353
$Hybrid^{RFE}$	11	0.455	0.596	0.629	0.571
SVM^{RFE}	11	0.436	0.474	0.610	0.428
LR^{RFE}	6	0.398	0.565	0.593	0.541
MLP^{RFE}	5	0.451	0.344	0.606	0.329
GRU^{RFE}	9	0.464	0.526	0.625	0.483

TABLE 4. **Evaluation results based on features from both categories (*Events* and *Themes*).**

Method	dt	ROC-AUC	F1	Precision	Recall
RF	11	0.408	0.607	0.580	0.637
Hybrid	3	0.500	0.559	0.666	0.515
SVM	11	0.479	0.586	0.654	0.549
LR	3	0.537	0.586	0.689	0.545
MLP	15	0.580	0.743	0.733	0.758
GRU	6	0.506	0.449	0.664	0.416
RF^{PRF}	2	0.487	0.673	0.604	**0.760**
$Hybrid^{PRF}$	15	0.572	0.715	0.718	0.715
SVM^{PRF}	15	0.528	0.488	0.694	0.448
LR^{PRF}	15	**0.710**	**0.755**	**0.796**	0.735
MLP^{PRF}	15	0.509	0.248	0.692	0.287
GRU^{PRF}	4	0.591	0.663	0.675	0.653
RF^{PFI}	15	0.492	0.618	0.667	0.586
$Hybrid^{PFI}$	6	0.565	0.726	0.714	0.750
SVM^{PFI}	4	0.551	0.701	0.698	0.704
LR^{PFI}	4	0.534	0.695	0.688	0.704
MLP^{PFI}	4	0.519	0.663	0.675	0.653
GRU^{PFI}	2	0.579	0.671	0.706	0.650
RF^{RFE}	6	0.526	0.483	0.680	0.447
$Hybrid^{RFE}$	13	0.469	0.630	0.644	0.617
SVM^{RFE}	13	0.491	0.597	0.658	0.561
LR^{RFE}	15	0.516	0.686	0.683	0.689
MLP^{RFE}	12	0.460	0.582	0.643	0.544
GRU^{RFE}	2	0.547	0.630	0.688	0.600

and GKG databases, respectively. The proposed PRF feature selection method provides the best overall performance in all evaluation metrics. More specifically, the proposed LR^{PRF} model provides ROC-AUC, F1-score and Precision equal to 0.710, 0.755, and 0.796 for $dt = 15$, respectively, while the best Recall score is obtained by the proposed RF^{PRF} model (0.760) for $dt = 2$. Lastly, the results of this work indicate that the proposed LR^{PRF} model using

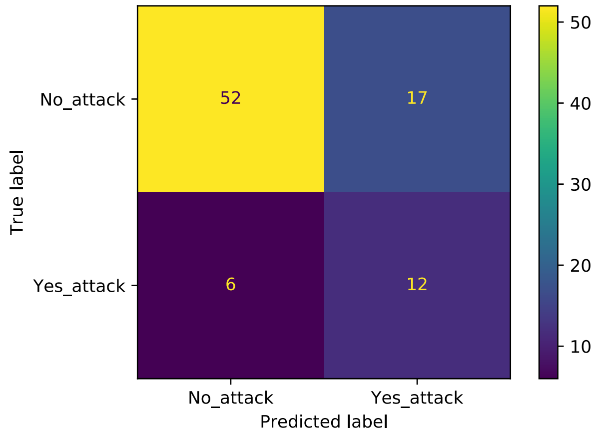


FIGURE 2. **Confusion matrix.**

the selected (and thus considered most informative) event and theme features achieved superior performance to the other approaches (baseline, PFI, RFE) in predicting terrorist incidents. The LR^{PRF} model achieved ROC-AUC score of 0.710 when using the proposed feature selection approach, in contrast to the other two feature selection methods PFI an RFE that provided poor performance.

Overall, the most significant outcome of these experiments is highlighted in Table 4, where the proposed LR^{PRF} model, using only the 134 selected features, achieved the best overall ROC-AUC score equal to 0.71. The confusion matrix for this result is depicted in Figure 2, which shows that the proposed LR^{PRF} model is able to correctly identify 57 out of 69 (75%) 2-day periods where no incident occurred and 12 out of 18 (66%) 2-day periods where an incident did occur. A high ROC-AUC score implies that the model has a low rate of false negatives and false positives, making it more reliable for practical applications. Furthermore, the fact that the LR^{PRF} model outperformed all the other models highlights its superiority in forecasting terrorist incidents and its potential for being used in real-world scenarios for threat assessment and decision support. This result underscores the importance of evaluating models on appropriate sets to accurately gauge their performance and ensure that the most effective models are selected for practical use. Finally, as observed in Table 4, the proposed feature selection outperformed the other feature selection methods, further highlighting the benefits of the proposed approach.

V. DISCUSSION

As the GDELT database encompasses an extensive array of features, feature selection is likely to be useful towards identifying distinctive indicators pertinent to a domain of interest. The experimental results have indicated that the most efficacious feature set has been successfully determined when using the proposed feature selection methodology (PRF), which assesses the predictive capability of individual features, compared to other approaches (PFI and RFE). This is particularly noteworthy given that our case study (UK) includes a sparse number of terrorist incidents, as recorded in the GTD. This relative sparsity of terrorist attack data in the UK and generally in Europe, compared to other parts of the world, makes it challenging to train highly accurate predictive models, as ML and DL models rely on large datasets to perform robust predictions.

Moreover, despite the richness of GDELT's variables in providing insights, they may not always correlate directly with specific predictive outcomes, such as acts of terrorism. This misalignment necessitates a thoughtful and meticulous approach to the selection and refinement of these variables. This is crucial when employing GDELT for specialised predictive analyses, to ensure that the data employed is truly reflective of the phenomena being forecast. The process of refining and aligning GDELT's vast data to specific predictive tasks is therefore not only a matter of data selection, but also one of understanding the complex relationships between data points and real-world events.

Another challenge that cannot be entirely dismissed is the potential inclusion of inaccuracies or misleading information in GDELT, commonly known as 'fake news'. This risk can be considered to be mitigated to some extent by the wide diversity of news sources included in GDELT that are likely to provide a broad spectrum of views and perspectives. Nevertheless, further mitigation strategies could also be considered for filtering fake news; one such plausible strategy could be to review the news sources provided by GDELT and consider only reputable and recognised news outlets with robust policies and tools in place for fighting against the spread of fake news. Such filtering mechanisms will be pivotal in ensuring the integrity and reliability of the data fed into the predictive models, thereby enhancing the credibility of the generated predictive insights.

Additionally, the potential inaccuracies in online data sources may diminish the usefulness of the forecasting model, as predictions in such cases may be skewed and the results may not reflect real-world scenarios. Given the sensitive nature of terrorism and the complexities of the factors leading up to an attack, the proposed approach should be viewed as a supportive tool, rather than a decision-making solution, since it provides insights based on general socio-political indicators and Localized news data, but cannot account for all the intricate and hidden motivations behind terrorist actions.

Overall, this research can assist counter-terrorism efforts by identifying potential risks and also lays a foundation for future work in predictive modelling of terrorist activities. Enhancing the data collection and filtering methods, integrating more diverse data sources, along with advancements in the ML and DL techniques may offer more sophisticated tools for analysing sparse datasets, thus refining the predictive capabilities and reliability of the proposed framework.

VI. CONCLUSION

This work explored the use of GDELT as the primary source of data to predict the likelihood of the occurrence of terrorist incidents and hence terrorism risk. As previously mentioned, predicting the occurrence of terrorist incidents is a field that has not been fully explored, while related efforts have mainly focused on detecting certain characteristics of terrorist incidents. To address this limitation, this work incorporated news data and analysed socio-political and economic instability as additional variables to predict terrorist incidents. In particular, a novel feature selection approach was used, along with a variety of ML and DL models, namely RF, LR, SVM, ensemble, MLP, and GRU, on Localized news data in order to provide terrorist incident forecasting.

The results of the experimental evaluation showed that the LR model trained on 134 selected features provides the best overall performance, achieving the highest ROC-AUC score, F1 score, and Precision. In particular, the proposed feature selection approach has demonstrated superior performance in identifying the most informative variables for predicting terrorism risk. The features retrieved using our method (PRF) have outperformed those selected through established methods, like PFI and RFE. The comparative analysis has revealed the effectiveness of our approach in yielding more accurate predictions of upcoming terrorist incidents and underscores the importance of employing tailored feature selection techniques for terrorist-based predictive tasks.

Although forecasting terrorist incidents and hence terrorism risk is a challenging task, this work investigated the impact of Localized online news in order to track the socio-political unrest of the country of interest, with the goal to predict future terrorist incidents. To summarise, the proposed approach provides a potential likelihood of future terrorist incidents by predicting the potential risk level of the target country based on assessments of socio-political and economic unrest. The objective of this research is to offer additional support to counter-terrorism efforts by providing an additional threat assessment for consideration.

Some possible directions for future work include the detection of fake news on online news sources and the incorporation of additional data sources (e.g., demographics data) that could be indicators of unrest. In particular, the use of a fake news detector could be useful in order to improve the quality and credibility of uploaded online news, thus, ensuring a strong correlation with real events. In particular, it could significantly contribute to the data preprocessing steps, enabling a clearer and more accurate mapping of features to real-world phenomena. This is especially crucial in the domain of event prediction, where the veracity of data directly influences the predictive power of the model.

Finally, extending the data acquisition to encompass a broader range of European countries would provide a more holistic analysis of terror trends across the continent. Such an expansion would not only improve the geographical comprehensiveness, but also allow for a deeper understanding of the regional specificities and cross-border dynamics of terrorism in Europe. By integrating and cross-verifying features from a wider data pool, researchers can achieve more granular insights and enhance the predictive accuracy of potential terror-related activities.

ACKNOWLEDGMENT

This work has received funding from the European Union Research and Innovation Framework Programme under Horizon Europe project PERIVALLON (Grant Agreement No 101073952) and Horizon 2020 project APPRAISE (Grant Agreement No 101021981). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] M. M. H. Khorshid, T. H. M. Abou-El-Enien, and G. M. A. Soliman, ''Hybrid classification algorithms for terrorism prediction in middle east and North Africa,'' *Int. J. Emerg. Trends Technol. Comput. Sci.*, vol. 4, no. 3, pp. 23–29, 2015.

[2] D. Tin, D. G. Barten, H. De Cauwer, L. J. Mortelmans, and G. R. Ciottone, ''Terrorist attacks in Western Europe: A counter-terrorism medicine analysis,'' *Prehospital Disaster Med.*, vol. 37, no. 1, pp. 19–24, Feb. 2022.

[3] START. (2022). *Global Terrorism Database 1970–2020*. Data Retrieved From World Development Indicators. [Online]. Available: https://www.start.umd.edu/gtd

[4] F. Gohar, W. H. Butt, and U. Qamar, ''Terrorist group prediction using data classification,'' in *Proc. Int. Conf. Artif. Intell. Pattern Recognit. (AIPR)*, 2014, pp. 199–208.

[5] A. Sachan and D. Roy, ''TGPM: Terrorist group prediction model for counter terrorism,'' *Int. J. Comput. Appl.*, vol. 44, no. 10, pp. 49–52, Apr. 2012.

[6] M. I. Uddin, N. Zada, F. Aziz, Y. Saeed, A. Zeb, S. A. Ali Shah, M. A. Al-Khasawneh, and M. Mahmoud, ''Prediction of future terrorist activities using deep neural networks,'' *Complexity*, vol. 2020, pp. 1–16, Apr. 2020. [Online]. Available: https://ideas.repec.org/a/hin/complx/1373087.html

[7] 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, Jun. 2022, Art. no. e0270681.

[8] K. Leetaru and P. A. Schrodt, ''GDELT: Global data on events, location, and tone,'' presented at the Int. Stud. Assoc. (ISA) Annu. Conv., San Francisco, CA, USA, 2013. [Online]. Available: http://data.gdeltproject.org/documentation/ISA.2013.GDELT.pdf

[9] N. Levin, S. Ali, and D. Crandall, ''Utilizing remote sensing and big data to quantify conflict intensity: The Arab spring as a case study,'' *Appl. Geography*, vol. 94, pp. 1–17, May 2018.

[10] L. Breiman, ''Random forests,'' *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[11] J. Liu, Q. Huang, C. Ulishney, and C. E. Dumitrescu, ''Comparison of random forest and neural network in modeling the performance and emissions of a natural gas spark ignition engine,'' *J. Energy Resour. Technol.*, vol. 144, no. 3, Mar. 2022, Art. no. 032310.

[12] L. Luo and C. Qi, ''An analysis of the crucial indicators impacting the risk of terrorist attacks: A predictive perspective,'' *Saf. Sci.*, vol. 144, Dec. 2021, Art. no. 105442.

[13] H. Mo, X. Meng, J. Li, and S. Zhao, ''Terrorist event prediction based on revealing data,'' in *Proc. IEEE 2nd Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2017, pp. 239–244.

[14] Y. Feng, D. Wang, Y. Yin, Z. Li, and Z. Hu, ''An XGBoost-based casualty prediction method for terrorist attacks,'' *Complex Intell. Syst.*, vol. 6, no. 3, pp. 721–740, Oct. 2020.

[15] Y. Xia, C. Liu, Y. Li, and N. Liu, ''A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring,'' *Expert Syst. Appl.*, vol. 78, pp. 225–241, Jul. 2017.

[16] G. M. Campedelli, M. Bartulovic, and K. M. Carley, ''Learning future terrorist targets through temporal meta-graphs,'' *Sci. Rep.*, vol. 11, no. 1, p. 8533, Apr. 2021.

[17] L. D. Howell, *International Country Risk Guide Methodology*, vol. 7. East Syracuse, NY, USA: PRS Group, 2011.

[18] WorldBank. (2013). *World Bank Data*. [Online]. Available: https://data.worldbank.org/indicator/

[19] Z. Hou, Y. Zhou, X. Wu, and F. Bu, ''Attention-based spatial–temporal multi-graph convolutional networks for casualty prediction of terrorist attacks,'' *Complex Intell. Syst.*, vol. 9, no. 6, pp. 6307–6328, 2023.

[20] S. Keertipati, B. T. R. Savarimuthu, M. Purvis, and M. Purvis, ''Multi-level analysis of peace and conflict data in GDELT,'' in *Proc. MLSDA 2nd Workshop Mach. Learn. Sensory Data Anal.*, Dec. 2014, pp. 33–40.

[21] Y. Keneshloo, J. Cadena, G. Korkmaz, and N. Ramakrishnan, ''Detecting and forecasting domestic political crises: A graph-based approach,'' in *Proc. ACM Conf. Web Sci.* New York, NY, USA: Association for Computing Machinery, Jun. 2014, pp. 192–196, doi: 10.1145/2615569.2615698.

[22] I. Toure and A. Gangopadhyay, ''Real time big data analytics for predicting terrorist incidents,'' in *Proc. IEEE Symp. Technol. Homeland Secur. (HST)*, May 2016, pp. 1–6.

[23] I. Toure and A. Gangopadhyay, ''Analyzing real time terrorism data,'' in *Proc. IEEE Int. Symp. Technol. Homeland Secur. (HST)*, Apr. 2015, pp. 1–4.

[24] F. Qiao, P. Li, X. Zhang, Z. Ding, J. Cheng, and H. Wang, ''Predicting social unrest events with hidden Markov models using GDELT,'' *Discrete Dyn. Nature Soc.*, vol. 2017, pp. 1–13, May 2017.

[25] L. Huang, H. Yuan, J. Chen, and L. Deng, ''Short-term forecast and analysis of mass incidents based on time series model,'' in *Proc. 14th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD)*, Jul. 2018, pp. 787–791.

[26] D. Galla and J. Burke, ''Predicting social unrest using gdelt,'' in *Proc. Int. Conf. Mach. Learn. Data Mining Pattern Recognit.* Cham, Switzerland: Springer, 2018, pp. 103–116.

[27] A. M. Carrington, D. G. Manuel, P. W. Fieguth, T. Ramsay, V. Osmani, B. Wernly, C. Bennett, S. Hawken, O. Magwood, Y. Sheikh, M. McInnes, and A. Holzinger, ''Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 329–341, Jan. 2023.

[28] K. Leetaru and P. A. Schrodt, ''GDELT: Global data on events, location, and tone, 1979–2012,'' in *ISA Annual Convention*, vol. 2, no. 4. Princeton, NJ, USA: Citeseer, 2013, pp. 1–49.

[29] A. Uyar and F. M. Aliyu, ''Evaluating search features of Google knowledge graph and bing satori,'' *Online Inf. Rev.*, vol. 39, no. 2, pp. 197–213, Apr. 2015.

[30] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, ''Gene selection for cancer classification using support vector machines,'' *Mach. Learn.*, vol. 46, no. 1, pp. 389–422, 2002.

[31] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, ''On the properties of neural machine translation: Encoder–decoder approaches,'' 2014, *arXiv:1409.1259*.

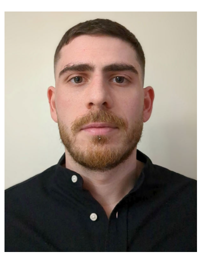
[32] M. W. Gardner and S. Dorling, ''Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences,'' *Atmos. Environ.*, vol. 32, nos. 14–15, pp. 2627–2636, 1998.

[33] ''Global terrorism index 2022,'' Inst. Econ. & Peace, Sydney, NSW, Australia, Tech. Rep., Mar. 2022. [Online]. Available: http://visionofhumanity.org/resources

[34] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, ''Scikit-learn: Machine learning in Python,'' *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[35] A. F. Agarap, ''Deep learning using rectified linear units (ReLU),'' 2018, *arXiv:1803.08375*.

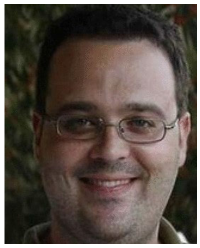
[36] D. P. Kingma and J. Ba, ''Adam: A method for stochastic optimization,'' 2014, *arXiv:1412.6980*.



GEORGIOS KOUTIDIS received the B.Sc. degree in mathematics from the Aristotle University of Thessaloniki, in 2016, and the M.Sc. degree in data science for decision making from Maastricht University, in 2024. He is currently a Research Associate with the Centre for Research and Technology Hellas. His research interests include predictive analytics and forecasting models, with a focus on applying data-driven approaches to enhance decision-making processes.

KOSTAS LOUMPONIAS received the B.Sc. degree in mathematics and the M.Sc. and Ph.D. degrees in statistics from the School of Mathematics, Aristotle University of Thessaloniki, Greece, in 2013, 2015, and 2020, respectively. He is currently a Postdoctoral Researcher with the Centre for Research and Technology Hellas. His research interests include hidden states estimation with censored data, social network analysis, and neural networks.

THEODORA TSIKRIKA received the degree in computer science from the University of Crete, Heraklion, and the M.Sc. degree in advanced methods in computer science and the Ph.D. degree in computer science from the Queen Mary University of London. She has been a Postdoctoral Research Fellow with CERTH-ITI, since 2013, and previously was a Postdoctoral Researcher with CWI, Amsterdam, The Netherlands, from 2007 to 2010; the University of Applied Sciences Western Switzerland, Sierre, Switzerland, from 2011 to 2012; and the Royal School of Library and Information Science, Copenhagen, Denmark, in 2013. She has participated in more than 30 European and national research projects and has co-authored more than 100 publications in refereed journals and international conferences. Her research interests include the intersection of the fields of information retrieval, data mining, artificial intelligence (AI), AI-based multimodal analytics, web search, domain-specific data discovery, web and social media mining, and evaluation, with a particular focus on security and cybersecurity applications.



STEFANOS VROCHIDIS (Member, IEEE) received the Diploma degree in electrical engineering from the Aristotle University of Thessaloniki, Greece, the M.Sc. degree in radio frequency communication systems from the University of Southampton, and the Ph.D. degree in electronic engineering from the Queen Mary University of London, U.K. He is currently a Senior Researcher (grade C) with the Information Technologies Institute, Center for Research and Technology Hellas, Thessaloniki, Greece, and the Head of the Multimodal Data Fusion and Analytics (M4D) Group, Multimedia Knowledge and Social Media Analytics Laboratory. He has edited three books and authored more than 250 related scientific journals, conferences, and book chapter publications. His research interests include multimedia analysis and retrieval, multimodal fusion, computer vision, multimodal analytics, and artificial intelligence, as well as media and arts, and environmental and security applications. He has participated in more than 50 European and national projects (in more than 15 as a project coordinator and the scientific or technical manager) and has been a member of the organization team of several conferences and workshops relevant to the aforementioned research areas.

IOANNIS KOMPATSIARIS (Senior Member, IEEE) was born in Thessaloniki, in 1973. He received the B.S. degree in electrical and computer engineering and the Ph.D. degree from the Aristotle University of Thessaloniki, in 1996 and 2001, respectively. He is currently the Director of the Information Technologies Institute and the Head of the Multimedia Knowledge and Social Media Analytics Laboratory. He is the co-author of 178 journal articles, more than 560 conference papers, and 59 book chapters. He holds eight patents. His research interests include image and video analysis, big data and social media analytics, semantics, human–computer interfaces (AR and BCI), eHealth, and security applications. He is a Senior Member of ACM. He has organized conferences, workshops, and summer schools. He is a member of the National Ethics and Technoethics Committee and an Elected Member of the IEEE Image, Video, and Multidimensional Signal Processing Technical Committee (IVMSPTC). He is an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING.