

# Machine Learning Techniques to Visualize and Predict Terrorist Attacks Worldwide using the Global Terrorism Database

Enrique Lee Huamani<sup>1</sup>, Alva Mantari Alicia<sup>2</sup>, Avid Roman-Gonzalez<sup>3</sup>  
Image Processing Research Laboratory (INTI-Lab)  
Universidad de Ciencias y Humanidades  
Lima, Perú

**Abstract**—Terrorist attacks affect the confidence and security of citizens; it is a violent form of a political struggle that ends in the destruction of order. In the current decade, along with the growth of social networks, terrorist attacks around the world are still ongoing and have had potential growth in recent years. Consequently, it is necessary to identify where the attacks were committed and where is the possible area for an attack. The objective is to provide assertive solutions to these events. As a solution, this research focuses on one of the branches of artificial intelligence (AI), which is the Automatic Learning, also called Machine Learning. The idea is to use AI techniques to visualize and predict possible terrorist attacks using classification models, the decision trees, and the Random Forest. The input would be a database that has a systematic record of worldwide terrorist attacks from 1970 to the last recorded year, which is 2018. As a final result, it is necessary to know the number of terrorist attacks in the world, the most frequent types of attacks and the number of seizures caused by region; furthermore, to be able to predict what kind of terrorist attack will occur and in which areas of the world. Finally, this research aims to help the scientific community use artificial intelligence to provide various types of solutions related to global events.

**Keywords**—Artificial intelligence; decision trees; machine learning; random forest; terrorist attack

## I. INTRODUCTION

The technological progress has benefited millions of people worldwide; being informed about any event is much faster and easier in comparison to previous decades. Thanks to the Internet, communication and social interaction is much more fluid, yet not all are profitable. Since technological growth in the world began to grow a higher degree of dissatisfaction about current events, as explained by the research of [1], which indicates an increase in recent years associated with terrorist attacks and various assaults in the world due to dissatisfaction with the political system. Terrorist attacks, as reported by [2], are considered as such if they are occasioned by political, religious, economic, or social reasons. The same author indicates that in 2007 a total of 2111 attacks were detected, which is almost similar to the peak in 1992; thus, the attacks are reappearing as they were in previous decades. To be able to visualize and predict these types of events, artificial intelligence must be used since this is one of the most modern sciences in charge of creating intelligent algorithms that can learn [3].

Nevertheless, artificial intelligence needs a set of data that can analyze the information and provide assertive solutions. For this reason, in this research, the Global Terrorism Database (GTD) [4], is used since its data will be useful to create classification models such as Decision Tree and the Random Forest to show probabilistic results. Considering artificial intelligence has become essential for the global economy and has brought positive effects on society [5]. Since the study of terrorist attacks can be extended to many areas of knowledge and can contribute to providing strategies to combat them, this paper aims to use Machine Learning to visualize and predict terrorist attacks from 1970 to 2018 (last recorded year), to contribute to the scientific community related to global events.

## II. BACKGROUND

There are some researches related to fighting against terrorism, and a clear example is a case of [6] which carries out an analysis to identify cyberterrorism in social networks following the Russian and Turkish law that determines when it is considered an attacking threat. Some studies used the GTD to show that there is an increase in terrorist attacks, as explained by the research [7], indicating that there are large volumes of data to provide predictive results. This situation can be complemented with the analysis of [8] that conducts a systematic study on the applications of Big Data in the field of counter-terrorism. Another example is the case of [9], which models terrorism using computer science based on the reasoning of Richardson's arms race theory together with elements of the analysis of Peng, Caspar, and Showalter.

## III. METHODOLOGY

AI is intelligence performed by machines. In the field of computational science, an optimal intelligent machine is a versatile agent that perceives its environment and carries out actions that maximize its chances of success in some objective [10]. This research uses the Global Terrorism Database [4], which collects historical information on terrorist attacks from 1970 to 2018. To predict the number of attacks by region and by type of terrorist attack, two classification models are used. Concerning artificial intelligence, these can be divided into several areas of knowledge. For this research, automatic learning, also known as machine learning, is used. Fig. 1 extracted from the study of [11] shows in a general way the knowledge areas of artificial intelligence.

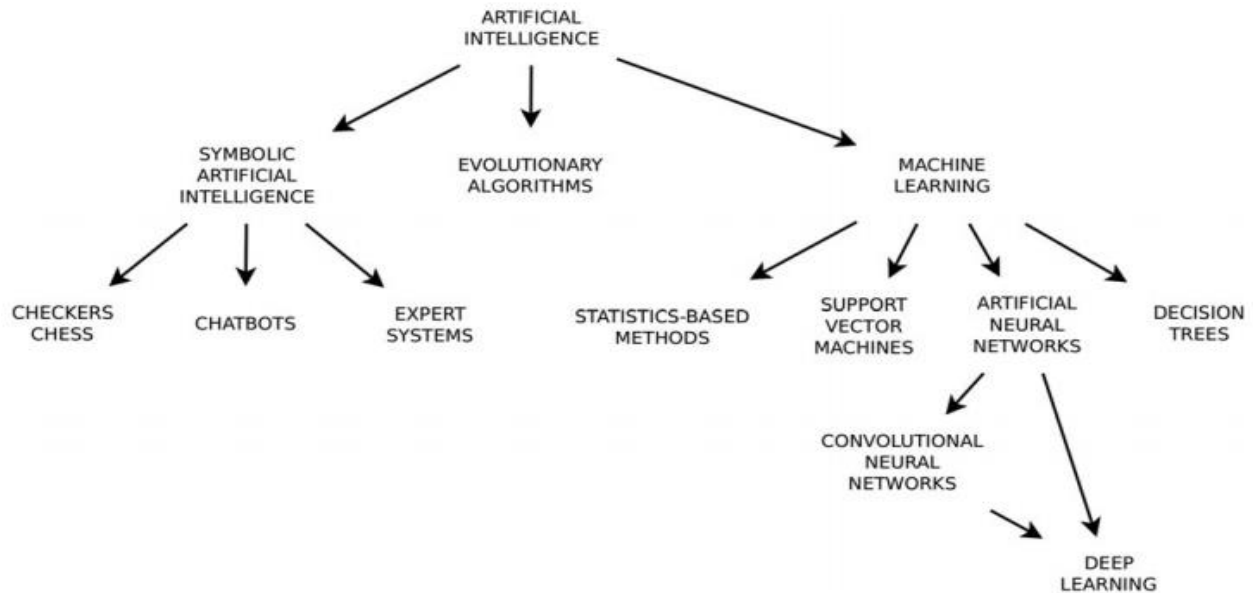


Fig. 1. Artificial Intelligence Domains.

### A. Machine Learning

Also called Automatic Learning, is one of the branches of artificial intelligence, its purpose is the development of techniques to enable computers to learn independently and be able to answer particular questions with great certainty. Machine Learning algorithms were designed and used from the beginning to analyze data sets, and today provides several indispensable tools for intelligent data analysis [12]. These types of intelligent algorithms can be divided into two categories, supervised and unsupervised learning. Table I present a detailed comparison of these two types [3], as well as evidence of the choice of the present research in supervised learning.

TABLE I. SUPERVISED AND UNSUPERVISED LEARNING COMPARISON

	Supervised Learning	Unsupervised Learning
Input Data	Uses Known and Labeled Input Data	Uses Unknown Input Data
Computational Complexity	Very Complex in Computation	Less Computational Complexity
Real Time	Uses off-line analysis	Uses Real Time Analysis of Data
Number of Classes	Number of Classes is Known	Number of Classes is not Known
Accuracy of Results	Accurate and Reliable Results	Moderate Accurate and Reliable Results

### B. Creating Model Steps

For the creation of the Machine Learning models, a methodology must be followed; in this case, it starts from the definition of the objective until the publication of the model. The steps taken to create the models related to the terrorist attacks are described below.

1) *Objective definition:* The objective proposed in this research is to visualize and predict terrorist attacks that have occurred in the world from 1970 to the present decade. To determine if there is an increase in terrorist attacks, the type of attacks, and in which regions they are occurring, using predictive models of classification is sought.

2) *Data collection and comprehension:* This phase of the investigation will be divided into three parts, which are obtaining the information, analyzing and comprehension the data, and data preparation as follows.

a) *Obtaining information:* Information gathering is one of the fundamental parts; if there is no data, the model cannot learn. The GTD information is in a CSV format, which will be extracted using the Python programming language pandas library. This information will be kept in a data frame to go through the statistical analysis.

b) *Data analysis and compression:* To improve data analysis, it is beneficial to have both statistical and graphical measurements to have a global view of how the data behaves. The fields of the GTD data set are analyzed based on the book of [13] that describes and analyzes its metrics in detail and highlights the most significant elements to have an excellent probabilistic model. The data that make up our database are further explained in specific ways with the title GTD dataset.

c) *Data preparation:* When the data is obtained, the information is ensured to be in a correct format so that the algorithm can be fed; the data must be structured so that it can then have training and validation data [14].

3) *Algorithm evaluation:* In this phase, the Machine Learning algorithm is used with the prepared data. A test data set is made to evaluate a range of standard algorithms and select those with the best results. When the favorable models have been selected, it is recommended to train the models with a sufficient amount of data. The mean absolute error serves to

quantify the accuracy of the prediction techniques by comparing, for example, the predicted versus observed values [15]. Concerning the complexity of the tree, you have to have a balance and place an adequate amount of leaves. Not taking care about the information quality, one can obtain data with poor results and obtain an overfitting. This situation represents a little amount of data to validate in our predictive model, or the underfitting, which is to have very over adjusted values, preventing to predict new data with new characteristics. It is necessary to have a balance so that the model gets to have an adequate percentage of possible solutions. There are ways to control the depth of the tree using the Python programming language, among them you have the argument `max_leaf_nodes`, which provides a way to control the over-setting against the lack of adjustment. The more leaves the model is allowed, the more it can be moved from the sub-adjustment area. An average absolute error was used to determine an adequate number of leaves, which is represented by the following equation.

$$mae = \frac{\sum_{i=1}^n abs(y_i - \lambda(x_i))}{n} \quad (1)$$

The equation was applied to both models. In the first one, which is the Decision Tree, a loop was introduced to determine the ideal amount to be used as shown in the pseudocode.

```
-----  
Pseudocode: MeanAbsoluteError()  
-----  
**method to obtain the Mean Absolute Error  
function get_mae (max_leaf_nodes, train_x, val_x, train_y, val_y)  
    model = DecisionTreeRegressor(max_leaf_nodes= max_leaf_nodes,  
    random_states=0)  
    model = fit (train_x, train_y)  
    preds_val = model.predict(val_x)  
    mae = mean_absolute_error (val_y, preds_val)  
return (mae)  
  
**Loop by number of sheets  
foreach max_leaf_nodes in [5, 50, 500, 5000,50000] do  
    my_mae =get_mae (max_leaf_nodes, train_x, val_x, train_y, val_y)  
    print ('Max leaf nodes': (max_leaf_nodes, my_mae))  
end
```

Once the iteration is concluded, the following results are obtained concerning Table II, where it can be seen that the best option is to use a total of 500 leaves in the case of the Decision Tree.

Concerning Random Forest, to create a loop to determine the number of trees can be a very exhaustive job, as the more trees there are, the longer the result time is. Therefore, the Random Forest model is used without a tree delimiter parameter. As a result, with the Decision Tree, we obtained the Mean Absolute Error: 0.4948, which is an ideal value for our prediction as it almost fulfills the midpoint between underfitting and overfitting.

TABLE II. MEAN ABSOLUTE ERROR RESULTS

Max leaf nodes	Mean Absolute Error
5	0.6757
50	0.5468
500	0.4948
5000	0.4634
50000	0.4898
500000	0.4898

4) *Results improvement*: Once the results are obtained, the model can be improved by selecting more features to the data or by setting the parameters of the different algorithms chosen in the model, if the model comes to meet the desired expectations the model is published.

5) *Model publication*: In this last stage, the model is confronted with the real problem. Also, at this stage, it is possible to measure the performance of the model, which forces a revision of the previous steps.

### C. Classification

Two widely used classifiers are employed; the Decision Tree and the Random Forest, which will help be able to predict the required outcomes. Their functionality is explained below.

1) *Decision tree*: The Decision Tree is a modern form of problem decision making [16], it is a type of classification model that is constituted by nodes in which each one of them represents a test of an attribute and a leaf node that provides a classification [17]. In this study, a total of 100 characteristics are classified to determine the prediction of terrorist attacks. However, the limitations of this Decision Tree, as explained in a previous section, is that it is a model that cannot learn new characteristics if they have an over-fit, so other classification models will be considered that will help determine the prediction with a higher degree of accuracy.

2) *Random forest*: Decision Trees present specific difficulties when generating the model, since creating a tree with many leaves can cause an over-fitting that may not be the most appropriate decision. Random trees are, therefore, used to achieve greater assertiveness [18]. Random trees use several trees averaging the final prediction of each tree. With this model, it is possible to have more optimal results [19], in the results section the predictions are given with these two classification models that are widely used in Machine Learning.

### D. GTD Dataset

The Global Terrorism Database (GTD) is an open-source database that contains information on terrorist events around the world from 1970 to 2018 (with annual updates planned for the future). As opposed to many other event databases, the GTD includes systematic data on national and international terrorist incidents that have occurred during this period and currently includes more than 180,000 cases [4]. The GTD data set is used to make predictions for two kinds of categories, which are: Type of terrorist attacks and the number of attacks

per region. Table III shows the types of attacks where the column attacktype1 is the identifier, ATTACKTYPE1\_TXT the name of the kind of attack and the last column the specific description of the attack.

Also, the present research seeks to predict attacks by region. Therefore it is established which countries are being

considered due to the fact that there are countries that have not been systematically registered in the GTD. As shown in Table IV, in the region column the identifier is obtained, which is followed by the txt\_region, the name of the region, and lastly the countries that have been considered.

TABLE III. TYPES OF TERRORISTIC ATTACKS DESCRIPTION

ATTACKTYPE1	ATTACKTYPE1_TXT	Description
1	Assassination	An act intended primarily to murder one or more specific and prominent individuals. It is usually carried out on individuals of some significance, such as high-ranking military officers, government officials, celebrities, etc. It does not involve attacks on non-specific target group members. The killing of a police officer would be an armed robbery, unless there is reason to believe that the perpetrators have targeted a particularly prominent officer for assassination.
2	Armed assault	An attack primarily aimed at causing physical harm or death to humans directly through the use of a firearm, incendiary, or sharp instrument (knife, etc.). It does not cover attacks that involve the use of fists, stones, sticks or other (less lethal) hand weapons. It also includes attacks involving certain kinds of explosive devices in addition to firearms, incendiary or sharp instruments. The subcategories of explosive devices included in this classification include grenades, projectiles, and unknown or other explosive devices that are thrown.
3	Bombing/explosion	An attack in which the primary effects are produced by an energetically non-stable material that rapidly decomposes and delivers a pressure wave resulting in physical damage to the surrounding environment. It may include high or low explosives (including a dirty bomb) but does not extend to a nuclear explosive device that releases fission and/or fusion energy, or an incendiary device in which decomposition occurs at a much slower rate. If an attack involves certain classes of explosive devices in conjunction with firearms, incendiary or sharp objects, then the attack is coded as an armed assault only. The subcategories of explosive devices covered by this classification are grenades, projectiles and unknown or other explosives.
4	Hijacking	An act designed to take control of a vehicle such as an airplane, ship, bus, etc. in order to redirect it to an unscheduled destination, force the release of prisoners, or some other political objective. Getting a ransom payment should not be the sole purpose of a kidnapping, but may be an aspect of the incident as long as other objectives have been declared as well. Kidnappings are distinguished from hostage-taking because the objective is a vehicle, regardless of whether there are people/passengers in the vehicle.
5	Hostage taking (barricade incident)	An act primarily undertaken to achieve a political objective by taking control of hostages through concessions or by interrupting normal operations. Such attacks are distinguished from kidnappings since the incident happens and usually takes place at the target's location with minimal or no intention of keeping the hostages for a prolonged period in a separate underground location.
6	Hostage taking (kidnapping)	An act committed for the purpose of taking possession of hostages so as to attain a political goal by means of concessions or the interruption of normal activities. Kidnappings are different from barricade incidents (the type of attack detailed above) as they involve the transfer and retention of hostages in another location.
7	Facility / infrastructure attack	An act, excluding the use of an explosive, intended mainly to inflict damage on a non-human target, i.e. a building, a monument, a train, an oil pipeline, etc. Such attacks may involve arson and various forms of sabotage (for example, sabotage of a railway is an attack on a facility or infrastructure, even if passengers are killed). Facility/infrastructure attacks may involve acts that are intended to damage a facility, but also harm the surrounding people in an incidental manner (e.g., an arson attack whose primary objective is to damage a building, but which causes injury or death in the process).
8	Unarmed assault	An attack which is primarily intended either to cause physical injury or death to humans in a direct manner using other than an explosive, firearm, incendiary, or sharp instrument (knife, etc.). This occurs since attacks with chemical, biological or radiological weapons are treated as unarmed assaults.
9	Unknown	The attack type cannot be determined from the information available.

TABLE IV. GTD REGION DATA DESCRIPTION

region	txt_region	Países considerados en el GTD
1	North America	Canada, Mexico, United States
2	Central America & Caribbean	Antigua and Barbuda ,Bahamas, Barbados, Belize,Cayman Islands,Costa Rica, Cuba, Dominica,Dominican Republic ,El Salvador, Grenada, Guadeloupe ,Guatemala ,Haiti, Honduras ,Jamaica, Martinique,Nicaragua, Panama ,St. Kitts and Nevis ,St. Lucia, Trinidad and Tobago
3	South America	Argentina, Bolivia, Brazil, Chile, Colombia,Ecuador,Falkland Islands, French Guiana, Guyana,Paraguay,Peru ,Suriname,Uruguay,Venezuela
4	East Asia	,China,Hong Kong,Japan, Macau,North Korea,South Korea,Taiwan
5	Southeast Asia	,Brunei,Cambodia,East Timor,Indonesia,Laos,Malaysia,Myanmar,Philippines,Singapore,South Vietnam,Thailand,Vietnam
6	South Asia	,Afghanistan ,Bangladesh ,Bhutan ,India ,Maldives ,Mauritius ,Nepal ,Pakistan,Sri Lanka
7	Central Asia	Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan
8	Western Europe	Andorra, Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Gibraltar, Greece, Iceland, Ireland, Italy, Luxembourg, Malta, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, Vatican City, West Germany (FRG)
9	Eastern Europe	Albania, Belarus, Bosnia-Herzegovina, Bulgaria, Croatia, Czech Republic, Czechoslovakia, East Germany (GDR), Estonia, Hungary, Kosovo, Latvia, Lithuania, Macedonia, Moldova, Montenegro, Poland, Romania, Russia, Serbia, Serbia-Montenegro, Slovak Republic, Slovenia, Soviet Union, Ukraine, Yugoslavia
10	Middle East & North Africa	Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, North Yemen, Qatar, Saudi Arabia, South Yemen, Syria, Tunisia, Turkey, United Arab Emirates, West Bank and Gaza Strip, Western Sahara, Yemen
11	Sub-Saharan Africa	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, People's Republic of the Congo, Republic of the Congo, Rhodesia, Rwanda, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Swaziland, Tanzania, Togo, Uganda, Zaire, Zambia, Zimbabwe
12	Australasia & Oceania	Australia, Fiji, French Polynesia, New Caledonia, New Hebrides, New Zealand, Papua New Guinea, Solomon Islands, Vanuatu, Wallis and Futuna

#### IV. RESULTS

In this section, the current information of the GTD and the predictive results made with the Decision Tree model and Random Forest are analyzed. This section is divided into the display and results of the classification models.

##### A. Display of the Classification Models

GTD defines a terrorist attack as an attempt by a state actor to achieve a political, economic, or social objective through fear by executing real threats of illegal force and violence. Therefore, these three attributes must be present, according to GTD. Given these attributes, a first overview and discussion of the terrorist attacks that occurred from 1970 to 2018 can be made, as shown in Fig. 2.

As shown in Fig. 2, there has been an increase in terrorist attacks from 2012 to 2018, which, although there has been a reduction in the last three years, is still high in comparison with previous decades. Table V shows the number of attacks by year range.

In the same way, Fig. 3 displays the attacks that have occurred worldwide by region, considering the countries previously detailed in Table IV.

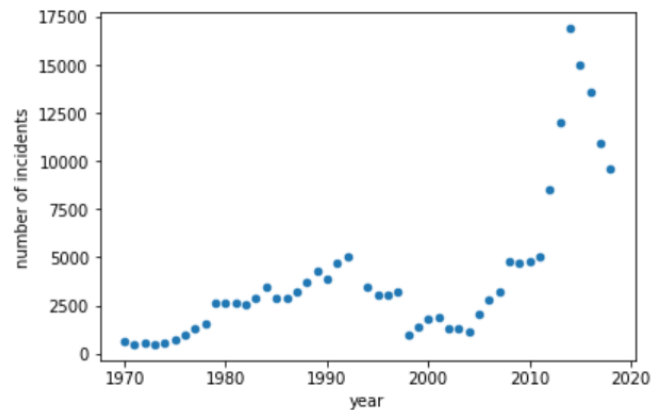


Fig. 2. Terrorist Attacks Occurred from 1970 to 2018.

TABLE V. TERRORIST ATTACKS WORLDWIDE BY YEAR RANGE

Year range	Number of terrorist attacks
1970 - 1980	12575
1980 - 1990	35045
1990-2000	30588
2010 -2018	96570

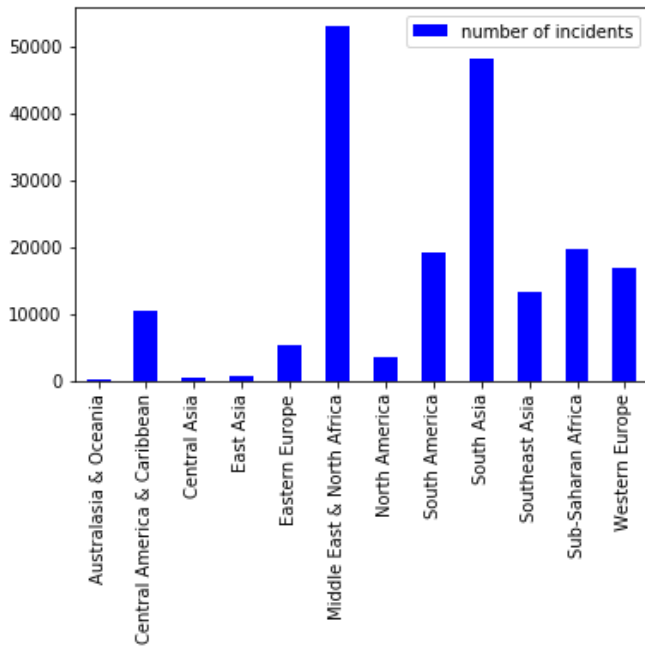


Fig. 3. Quantity of Terrorist Attacks by Region.

Fig. 3 illustrates that there has been a higher number of terrorist attacks in the Middle East & North Africa and South Asia. The investigation by [20] suggests that U.S. involvement in Africa is growing in response to the threat of terrorism brought about by the concerns of foreign corporations to expand their activities on the continent. Research by [21] states that the growth of terrorist attacks in South Asia is due to unemployment, inflation, poverty, and inequality, where income inequality has increased by 1,242%, followed by a population growth rate of 1,125% and political uncertainty of 1,102%. Table VI shows the number of attacks by region in more detail.

As can be seen in Fig. 4 there are more terrorist attacks with bombs and explosions followed by armed assault. The number of attacks occurred by type is given in Table VII.

TABLE VI. TERRORIST ATTACKS WORLDWIDE BY REGION

Region	Number of terrorist attacks
North America	69
Central America & Caribbean	45
South America	1673
East Asia	116
Southeast Asia	7732
South Asia	32476
Central Asia	88
Western Europe	2000
Eastern Europe	2760
Middle East & North Africa	35603
Sub-Saharan Africa	13526
Australasia & Oceania	69

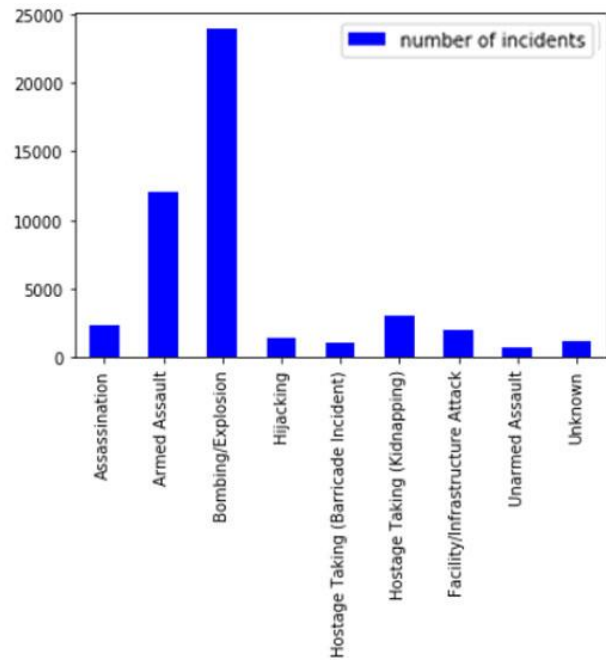


Fig. 4. Quantity of Terrorist Attacks by Type.

TABLE VII. TERRORIST ATTACKS WORLDWIDE BY TYPE

Terrorist attacks types	Number of terrorist attacks
Assassination	20117
Armed assault	45251
Bombing/explosion	91841
Hijacking	688
Hostage taking (barricade incident)	1048
Hostage taking (kidnapping)	12138
Facility / infrastructure attack	11017
Unarmed assault	1096
Unknown	8267

### B. Results of the Classification Models

The models used in this research, as noted in the preceding section, are the Decision Tree and the Random Forest, both focused on supervised learning. The difference between these two models resides in the complexity of the tree. While the Decision Tree tries to ramify all the data, Random Forest uses multiple trees, making the process much more complex, and the results of the predictions are very time-consuming. The results obtained with the Decision Tree are shown below, followed by Random Forest.

1) *Decision tree prediction result:* In this case, the Decision Tree has a total of 500 leaves as it is in the balance of under- and over-utilization, being optimal for the realization of the model. Fig. 5 illustrates a geographical map that determines the number of attacks per region, with the least intense colors having the least attacks and the most intense colors being the largest. Predictive results were obtained concerning this model, as shown in Table VIII, with an accuracy percentage of 75.45% of assertiveness.

Prediction of terrorist attacks types concerning the Decision Tree obtained a 79.24% of accuracy. For this research, this percentage is a very favorable value since making a more significant adjustment can show a percentage almost to 100%, which is an undesirable result for our model to learn with new data. In Table IX, the predictive results of the types of terrorist attacks by using the random forest are presented.

TABLE VIII. TERRORIST ATTACK PREDICTION BY REGION USING DECISION TREES

Region	Number of terrorist attacks
North America	810
Central America & Caribbean	2543
South America	4971
East Asia	0
Southeast Asia	2993
South Asia	12736
Central Asia	0
Western Europe	3646
Eastern Europe	2398
Middle East & North Africa	14053
Sub-Saharan Africa	3697
Australasia & Oceania	19

TABLE IX. PREDICTION OF TYPES OF TERRORIST ATTACKS WITH DECISION TREES

Terrorist attacks types	Number of terrorist attacks
Assassination	445
Armed assault	14660
Bombing/explosion	23599
Hijacking	828
Hostage taking (barricade incident)	2348
Hostage taking (kidnapping)	3222
Facility / infrastructure attack	953
Unarmed assault	845
Unknown	966

2) *Random forest prediction result:* Regarding the Random Forest, an assertiveness percentage of 89.544% was obtained, which in all tests is the most appropriate value to consider in the Random Forest model. Fig. 6 shows a geographical map that determines the number of attacks per region, where the least intense colors have the least attacks, and the most intense colors have the most attacks. Table X shows the results of the number of attacks per region.

Likewise, in conjunction with the types of terrorist attacks carried out with Random Forest, it was possible to obtain a percentage of 90.414% assertiveness. Table XI shows the results obtained.

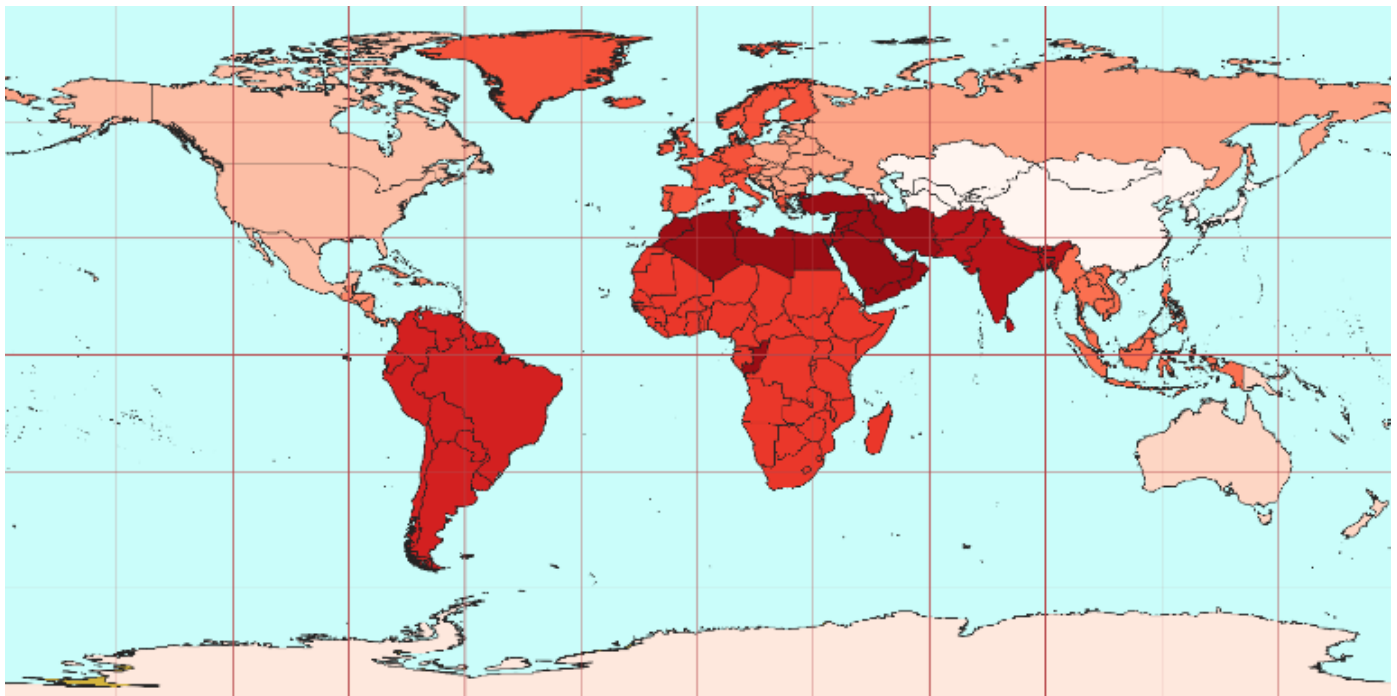


Fig. 5. Geographical Map of Terrorist Attack Prediction by Region using Decision Trees.



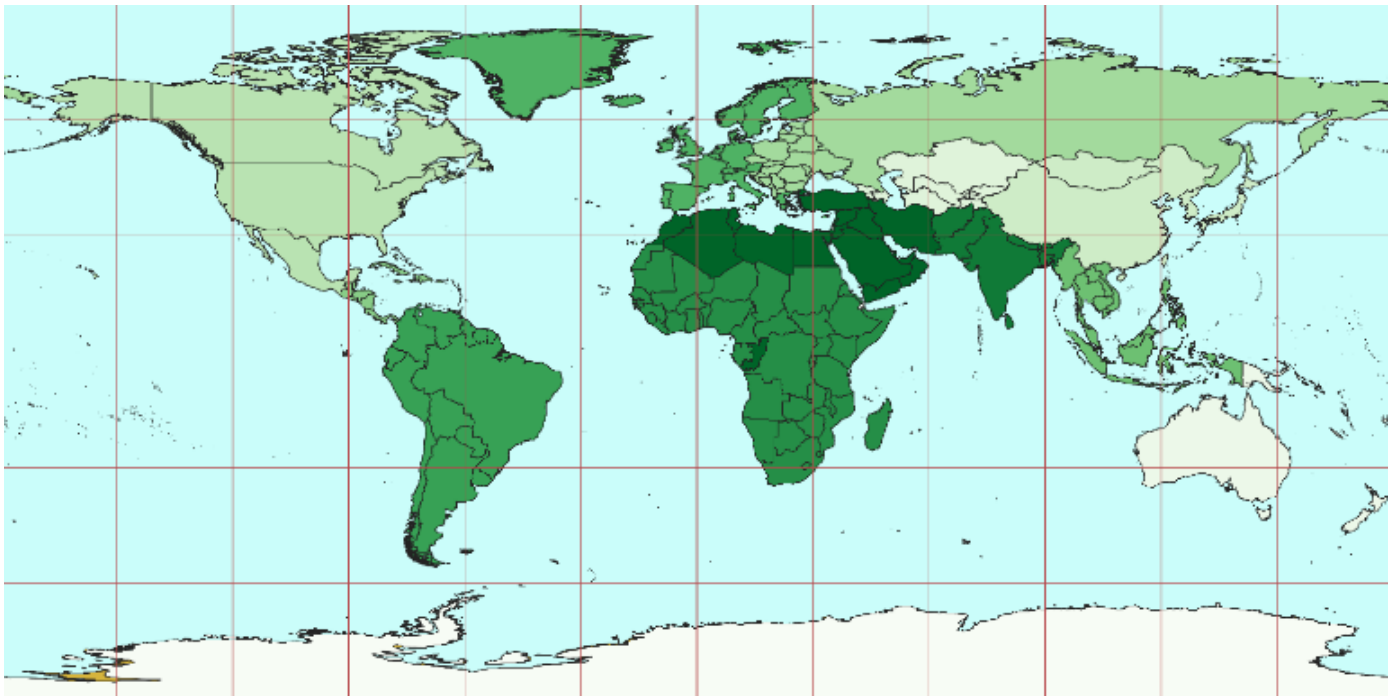


Fig. 6. Geographical Map of Terrorist Attack Prediction by Region using Random Forest.

TABLE X. TERRORIST ATTACK PREDICTION BY REGION USING RANDOM FOREST

Region	Number of terrorist attacks
North America	848
Central America & Caribbean	2599
South America	4819
East Asia	225
Southeast Asia	3368
South Asia	12027
Central Asia	156
Western Europe	4236
Eastern Europe	1331
Middle East & North Africa	13290
Sub-Saharan Africa	4994
Australasia & Oceania	73

TABLE XI. PREDICTION OF TYPES OF TERRORIST ATTACKS WITH RANDOM FOREST

Terrorist attacks types	Number of terrorist attacks
Assassination	2410
Armed assault	12097
Bombing/explosion	23909
Hijacking	1381
Hostage taking (barricade incident)	1114
Hostage taking (kidnapping)	3070
Facility / infrastructure attack	1999
Unarmed assault	733
Unknown	1153

## V. DISCUSSION AND CONCLUSIONS

Terrorist attacks are among the causes of national instability. A clear understanding of how this event is occurring will help us to conduct more in-depth investigations. The focus of future research will be on performing a quantitative analysis of the countries in each region to conduct further research. Other future work to be done is the use of Big Data techniques for sentiment analysis, which will extract information from social networks to determine possible threats of cyber terrorism. Thus the investigation would use large volumes of data. As explained by the research of [22], Big Data offers improved solutions for high amounts of information. To be able to use this type of architecture, the work implemented in the year 2019 of [23] will be employed to provide predictions utilizing a total of 28 computers working in parallel.

Through this research, it is possible to conclude that the use of Machine Learning techniques was able to visualize and predict terrorist attacks. The results section shows that there has been a considerable growth in terrorist attacks since 2010 and that due to the classification models, it was possible to determine the probability of which region and type of attack may occur. Concerning the number of attacks by region, it was obtained that there is a probability that they will happen in the Middle East & North Africa and followed by South Asia. Regarding the types of attacks, there is still the probability that bombs and explosions are involved, followed by armed assault. The results have been successfully achieved by using the historical data collected from the GTD. The models that were made through Decision Trees and Random Forest give the same probabilistic results from 75.45% to 90.414% of assertiveness. These results demonstrate that the techniques of Machine Learning are ideal for contributing to research related to world events.



REFERENCES

- [1] S. Kalaiarasi, A. Mehta, D. Bordia, and Sanskar, "Using global terrorism database (GTD) and machine learning algorithms to predict terrorism and threat," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 1, pp. 5995–6000, 2019.
- [2] G. LaFree, "The Global Terrorism Database (GTD): Accomplishments and challenges," *Perspect. Terror.*, vol. 4, no. 1, pp. 24–46, 2010.
- [3] S. Rusell and P. Norvig, *Inteligencia Artificial*, vol. 2, no. 6. 2007.
- [4] Start, "Global Terrorism Database." [Online]. Available: <https://www.start.umd.edu/gtd/>.
- [5] X. Ren and Y. Chen, "How Can Artificial Intelligence Help with Space Missions - A Case Study: Computational Intelligence-Assisted Design of Space Tether for Payload Orbital Transfer under Uncertainties," *IEEE Access*, vol. 7, pp. 161449–161458, 2019.
- [6] M. B. Bicak and D. Bogdanova, "Fighting Cyber Terrorism: Comparison of Turkey and Russia," *Int. Congr. Big Data, Deep Learn. Fight. Cyber Terror. IBIGDELFT 2018 - Proc.*, pp. 98–101, 2019.
- [7] T. Xia and Y. Gu, "Building terrorist knowledge graph from global terrorism database and wikipedia," *2019 IEEE Int. Conf. Intell. Secur. Informatics, ISI 2019*, pp. 194–196, 2019.
- [8] S. Nie and D. Sun, "Research on counter-terrorism based on big data," *Proc. 2016 IEEE Int. Conf. Big Data Anal. ICBDA 2016*, 2016.
- [9] T. Hackworth and P. Treleaven, "Modelling terrorism," *2007 IEEE Int. Conf. Networking, Sens. Control. ICNSC'07*, no. April, pp. 255–260, 2007.
- [10] D. L. Poole, A. Mackworth, and R. G. Goebel, "Computational Intelligence and Knowledge," *Comput. Intell. A Log. Approach*, no. Ci, pp. 1–22, 1998.
- [11] F. Galbusera, G. Casaroli, and T. Bassani, "Artificial intelligence and machine learning in spine research," *Jor Spine*, vol. 2, no. 1, p. e1044, 2019.
- [12] I. Konenko, "Machine learning for medical diagnosis: History, state of the art and perspective," *Artif. Intell. Med.*, vol. 23, no. 1, pp. 89–109, 2001.
- [13] Start, *Global Terrorism Database Codebook*, no. October. 2019.
- [14] D. Cleaning, "Data Preparation For Data," no. March 2012, pp. 375–381, 2007.
- [15] L. Frías-Paredes, F. Mallor, M. Gastón-Romeo, and T. León, "Dynamic mean absolute error as new measure for assessing forecasting errors," *Energy Convers. Manag.*, vol. 162, no. February, pp. 176–188, 2018.
- [16] S. Patil and U. Kulkarni, "Accuracy prediction for distributed decision tree using machine learning approach," *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, vol. 2019-April, no. Icoei, pp. 1365–1371, 2019.
- [17] S. S. Gavankar and S. D. Sawarkar, "Eager decision tree," *2017 2nd Int. Conf. Conver. Technol. I2CT 2017*, vol. 2017-Janua, pp. 834–840, 2017.
- [18] W. Deng, Y. Guo, J. Liu, Y. Li, D. Liu, and L. Zhu, "A missing power data filling method based on improved random forest algorithm," *Chinese J. Electr. Eng.*, vol. 5, no. 4, pp. 33–39, 2020.
- [19] F. Guo et al., "What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests," *Int. J. Wildl. Fire*, vol. 25, no. 5, pp. 505–519, 2016.
- [20] S. T. Barnes, "Global flows: Terror, oil & strategic philanthropy," *Rev. Afr. Polit. Econ.*, vol. 32, no. 104–105, pp. 235–252, 2005.
- [21] G. Akhmat, K. Zaman, T. Shukui, and F. Sajjad, "Exploring the root causes of terrorism in South Asia: everybody should be concerned," *Qual. Quant.*, vol. 48, no. 6, pp. 3065–3079, 2014.
- [22] A. J. J. Mary and L. Arockiam, "Jen-Ton: A framework to enhance the accuracy of aspect level sentiment analysis in big data," *Proc. Int. Conf. Inven. Comput. Informatics, ICICI 2017*, no. Icici, pp. 452–457, 2018.
- [23] E. L. Huamani, P. Condori, and A. Roman-Gonzalez, "Implementation of a Beowulf Cluster and Analysis of its Performance in Applications with Parallel Programming," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 8, pp. 522–527, 2019.