National College of Ireland

Evaluating the Impact of Drone Strikes on Terrorism Dynamics

MSc Research Project
MSc in Data Analytics

Dixita Yadav
Student ID: x23175796

School of Computing
National College of Ireland

Supervisor: Aaloka Anant

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Dixita Yadav
Student ID:	X23175796
Programme:	MSc in Data Analytics **Year:** 2024
Module:	MSc Research Project
Supervisor:	Aaloka Anant
Submission Due Date:	12th Dec 2024
Project Title:	Evaluating the Impact of Drone Strikes on Terrorism Dynamics
Word Count:	5469 **Page Count:** 17

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Dixita Yadav

Date: 12-12-2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only

Signature:	
Date:	
Penalty Applied (if applicable):	

Evaluating the Impact of Drone Strikes on Terrorism Dynamics

Dixita Yadav
x23175796

Abstract

Terrorism is a threat to the world that results in thousands of casualties, assaults, overnight political shifts, downfalls in economies and much more. Governments and its defence bodies are trying their best to counter against terrorism by adopting latest technological tools such as UAV's (Unmanned aerial vehicle) i.e. Drone strikes, to attack terrorist occupied regions. It has emerged as a best tool of counterattack against terrorism while involving minimal risk of military personnel. It is debatable on the fact that how much of retaliation is observed on a successful drone strike and does drone strikes affect the frequency of terrorist attacks in that region. To study this, research is carried out by using two different open-source datasets i.e. Global Terrorism Database (GTD) and Drone Wars Data. This research aims to evaluate the impact of drone strikes on terrorism to measure the changes in terrorist activities and their motives. Post evaluation an ensemble technique (Voting Classifier) is used on the overlayed data using geospatial proximity and modelled using SVC and Random Forest as base models and Logistic regression as meta model. The implementation helps to understand the classification of severity of the drone strikes.

Keywords: Terrorism, Drone Strike, Voting Classifier, Logistic regression, Random Forest, SVC, Confusion Matrix

1 Introduction

1.1 Background and Motivation

Terrorism, any unlawful practice by an individual or a group of people to achieve a personal, political, ideological or religious goal. Over the last decades, lakh of innocent people is ruthlessly killed that raises questions of safety, stability and security. Fight against terrorism is a challenge for the government and its defence bodies with the key focus to neutralize the threat before it attacks. UAVs (Unmanned Aerial Vehicles) known as drones (Cronin, A.K.,2013), equipped with smart surveillance, advance and precised targeting attributes have emerged one of the critical tools in countering terrorism and neutralizing threats well before hand. It holds the unique ability to reach and strike in remote areas without risking the lives of pilots, provide real-time information of terrorist movements, etc. On the bright side where drones' strikes have emerged to be a successful strategy to eliminate terrorist groups and their planned attacks. It is debatable on the fact that the same drone strikes can cause civil causalities (Martin, C.,2018), which will act as a fuel for terrorist retaliation and their recruitment agenda. Now, this raises concern about the impact of drone strikes on terrorism dynamics and makes it important to dig deep and understand the correlation of drone strikes and change in behaviour of terrorists.

While the previous research has predicted drone strikes, next possible terrorist attack but surprisingly none of them combined both and that created a gap of not exploring the impact on terrorism dynamics i.e. post the drone strike carried out in targeted zone, is there any retaliation observed? Or is there any decrease in terrorist attacks?
Hence, this research aims to fill this gap by using two datasets from different open-source website i.e. Global Terrorism Database (GTD) and Drone Wars Data. By doing so this research contributes to leveraging the effectiveness of drone strikes in impacted regions by understanding the behaviour of terrorist groups and help defence bodies to refine their strategies to counter terrorism.

1.2 Research Question

How Machine Learning can be used to study the intensity and frequency of terrorist activities, by understanding the effect of drone strikes on terrorism dynamics?

1.3 Objectives

The aim of this research is to analyse the past drone strikes and terrorist attacks data available at Drone wars[1] and GTD[2] respectively to evaluate the impact of drone strikes on terrorism dynamics. An investigation on the aftereffects of the drone strikes in the terror attacked regions are studied. The drone strikes in general are expected to have an implicit nature on terrorism. However, it is possible that the situation could be escalated, and to understand and identify the obvious and hidden patterns an ensemble machine learning technique (Voting Classifier) is used. It starts with merging both the datasets using geospatial proximity (latitude and longitude coordinates) and analysing the terrorist attacks and drone strikes that with in the same timeline and within the same regions. The ensemble technique (Voting Classifier) is modelled using SVC and random forest and logistic regression as the techniques. By doing so, the factors contributing to the terrorism dynamics can be identified. This implication finds its place at the government and intelligence agencies to draw conclusions on the effectiveness of drone strikes in targeted zone and make informed decisions to their counter against terrorism.

1.4 Structure of Paper

This section highlights the general layout of the research paper and is briefly discussed below.

- **Section 2:** This section covers review of the work related to predict terrorist attacks and drone strikes by building machine learning and data mining models. This section is further divided into subsections as per the mentioned criteria.

- **Section 3:** This section explains the methodology used to achieve the goal of this research. It includes the detailed description of data gathering, data cleaning, data analysis, implementation and interpreting the result.

[1] https://dronewars.github.io/data/

[2] https://www.start.umd.edu/download-global-terrorism-database

- **Section 4:** This section discusses the methods, frameworks and foundational architecture of proposed research and the related requirements.

- **Section 5:** This section describes the methodology implementation. It includes the description of tools and language used to implement the research. It also, includes the transformed data, output, model development and questionnaires administered details.

- **Section 6:** This section provides detailed review of methods executed and their results. Also, it is further subdivided in subsection as per each experiment. Discussion section is also included which describes each methods outcome.

- **Section 7:** This section discusses the conclusion and the possible future work that can be done on this topic.

2 Related Work

Using highly advanced technology tools and equipment's to counter-terrorism operations has become a hot topic, especially when it comes to drone strikes. Usage of drone strikes is highly debatable as it directly influences the terrorist's activities, recruitment and border relations of the regions where drones' strikes are conducted. In the past years the concept of how precisely the drone strikes is targeted to destroy terrorist areas has gained a lot of attention, as it directly relates to the strategic behaviour followed by the terrorist groups post drone strikes. To get more insights on the same, several research and studies have been conducted and related literature review is given below, that focuses on the impact of drone strikes on dynamics of terrorism and highlighting the specificity of the drone attacks and terrorist adaptions to it.

2.1 Review of research work done on Machine Learning Approach for Enhancing Defence Against Global Terrorism

The researchers (Kanika Singh; Anurag Singh Chaudhary; Parmeet Kaur; 2019) analysed and predicted region and country of the terrorist attack by using six different Machine learning methods i.e. Logistic Regression, Decision Tree, K-Nearest, Linear Discriminant Analysis, Naive Bayes and SVM. Researchers utilized the GTD dataset to analyse and train the models on several variables like attack type, weapon type, regions, target, etc. and achieved the maximum accuracy of 82%, in predicting region and country of attack. But the model failed to capture the complexities of terrorist activities, which leaves the space for possible errors in predictions.

The scope of this study can be improved and expanded by utilizing optimizing algorithms and deep learning methods like transformers or LSTM which results in better recognition of the patterns, hence producing better accuracy. Also, multiple datasets can be integrated to reduce the biasness.

2.2 Review of research work done for prediction of civilians killing in future drone attacks

The researchers (Sarmad Hameed; Faraz Junejo; Mehtab Anwar Yousuf Zai; Imran Amin; 2018) performed a predictive analysis to forecast the total deaths that would die in an upcoming drone strike in Northern area of Pakistan i.e. Federally Administered Tribal Areas (FATA) by using Artificial Neural Network (ANN) method. The study is carried out by collecting and using the data from June 2004 to April 2017, sourced from Pakistan Body Count (PBC) website. 576 drone strikes are recorded and no of people killed including terrorists is 3560 out of which 2539 were innocent citizens and the drone attacks have further provoked and encouraged terrorism in Pakistan (Mahmood, R. and Jetter, M., 2019.) This study is conducted by considering fields like city, date, longitude, latitude, no of strikes , total no of people died in event, etc and successfully found a correlation between the civilians deaths and the frequency of strikes, which aligns with the point that, drone strikes can lead to change in terrorist behaviour and can result an increase in militant recruitment. The scope of this study can be improved by introducing a greater number of parameters for better accuracy. For example, the parameters like political statements, military, etc. are directly proportional to drone strikes. Hence, including these variables will result in producing more accurate results.

2.3 Review of research work done for predicting UAV (Unmanned Aerial Vehicle) strikes in Pakistan by using Neural Networks

In order to minimize the negative effect of illegal drone strikes and save lives of thousands of innocent civilians which can be potentially killed, researchers (Komal Zahid; Ushna Nafees; Shazia Parveen; Uzma Afzal; 2019) trained predictive model using Pakistan UAV strikes data collected from two open source websites. The study is done by using LSTM and RNN to build predictive models by achieving good accuracy with least RMSE value of 6.317852.

The scope of the study can be improved by taking the datasets with a greater number of variables and size. Therefore, richer datasets can be integrated to improve the predictive analysis. Also, if Neural Networks methods can be combined with other machine learning methods such as ensemble methods, might fetch better outputs.

2.4 Review of research work done for Data Mining Model Frameworks for GTD (Global Terrorism Database)

The researchers (Naman Thakur, Satnam Singh and Abhishek Kumar Pathak; 2022) focused and developed a statistical model in python that analysis the real-time terrorism data. To develop this model, Global Terrorism Database is used which holds records of terrorist attack since 1970. Detailed analysis of the model is carried out by using various python libraries like pandas, seaborn and plotly. This study is specifically done to check terrorism trends in the Punjab, India. The built model is a smart early warning system that predicts terrorism related challenges.

The main motivation behind the research is to assist the government and its defence bodies in making better decisions by generating resourceful insights through their mode which will highlight the red-zones or high-risk zones and can be attacked by terrorists.

2.5 Review of research work done in analysing the Emotions, Terrorist Threat, and Drones

The researchers (Kerstin Fisk, Jennifer L. Merolla, and Jennifer M. Ramos; 2019) studied the psychological and emotional factors of people supporting drone strikes against terrorists due to increase in threats to their safety. The research particularly focused and later contended that emotional reaction, especially anger has played an important role in supporting drone strikes as individuals reacted in anger when exposed with the terrorism threats and hence ended up reacting aggressively and supporting drone-strikes. The data collected for this research is from three different nations i.e. Turkey, France and United States via online surveys. Based on the data collected the researchers claimed three hypothesis and ran a robustness test for all and concluded that terror threat made an impact on people and got two major responses i.e. Anger and Fear. And the volume of people who showed anger are more and hence supported Drone Strikes probabilities.

(Chappelle, W., Skinner, E., Goodman, T., Swearingen, J. and Prince, L., 2018) studied the effects on people behaviour post drone strikes by conducting a relative risk analysis. The results were as expected, people showed negative reactions where Relative risk value achieved was 1.94, concluding that civilians witnessing drone attacks are prone to retaliate.

This research acted as a bridge and filled the gap between the public policies and emotional psychology, giving a strong message to the government that people support military tools like drone strikes to counter-terrorism. Although there are limitations to this study as other emotions like sympathy, anxious, etc are not given importance which can influence the decision of using drones. Also, the study collects sample of three different countries, resulting in variations in emotional biasness.

The scope of the study can be improved by studying and including more emotions and using sentiment analysis.

After reviewing and critically analysing all the previous related work that predicts the future terrorist attacks and drone strikes by building various machine learning and ensemble techniques, there still room for further development. Surprisingly, digging deep it is found that there is no dedicated research which combines drone strikes and its impact on terrorism. The research focuses on integrating the drone strike data with the terrorism metrics analysing how drone strikes affect frequency, intensity and type of terrorist activities. This provides a better integrated data approach in comparison with other techniques. One of the main differences from the other studies **(**Garcia-Bernardo, J., Dodds, P.S. and Johnson, N.F., 2016**), (**Kim, S.S., 2023.**)** and **(**Aljohani, M., Mukkamalai, R. and Olariu, S., 2024**)** is that the focus is on the effectiveness, intensity, frequency and success ratio of drone strikes where the analysis is done on the general trend of the terrorism dynamics

Therefore, this laid the foundation for this research and encouraged to try to use machine learning approach. It starts with collecting the data, cleaning it, handling missing values, performing EDA (Exploratory Data Analysis), selecting important features to run the models and finding the results. Later, performed an ensemble technique (voting classifier), consisting of two classifiers (Random Forest, SVC) and Logistic Regression whose individual predictions are feed as one to run the final model (voting classifier) as shown in Figure 1.

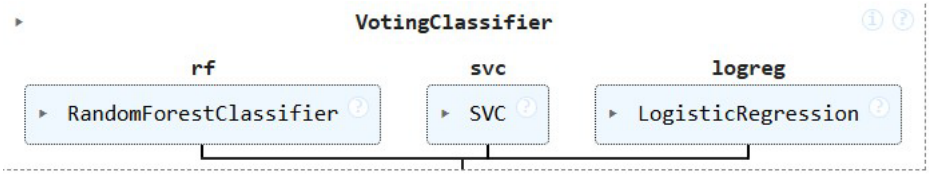

Figure 1: Voting Classifier Pipeline

3 Research Methodology

The data is collected, and it is prepared to understand the data frame of the both the datasets. Once the data frame is understood, the data is cleaned, relabelled and the missing records are handled across. Once the missing records are handled, the both the data are merged using geospatial proximity. The output dataset is not put through another set of cleaning before modelling. The modelling consists of testing and training dataset that is put through an ensemble technique of Random Forest and SVC and logistic regression as base models, which then is taken as an output from voting classifier. Figure 2 gives an overview of the Model Architecture.

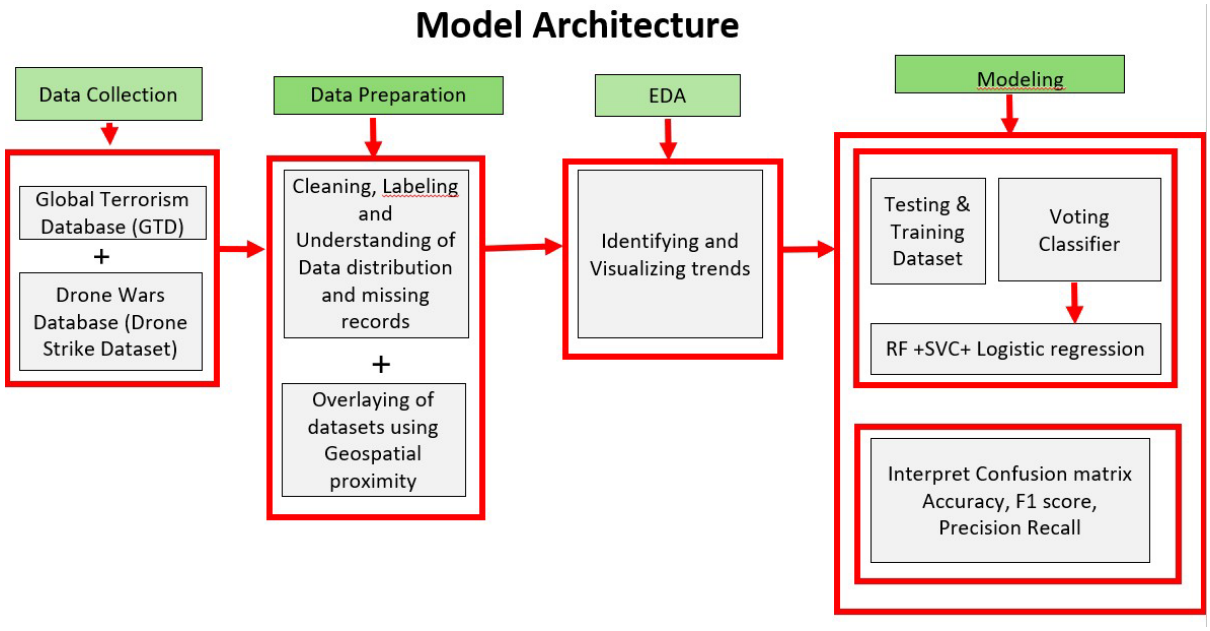

Figure 2: Model Architecture

3.1 Datasets Overview

To kick start with the research finding, two datasets from different open-source website i.e. Global Terrorism Database (GTD)[1] and Dronewars[2] is extracted in xlsx formats. GTD is maintained at University of Maryland by the START (National Consortium for the Study of Terrorism and Responses to Terrorism) and is enriched with terrorism-incidents information across globe from 1970 to 2020. Similarly, Drone Wars Database provides the data collected by Bureau of Investigative Journalism (BIJ) on United States drone strikes in Afghanistan, Yemen, Pakistan and Somalia. Each dataset is discussed below.

3.3.1 Global Terrorism Database

GTD dataset contains 209706 x135 records which provides information of terrorist attacks. Each event in GTD hold following detailed information.

- **Date**: Day, Month and Year of the incident.
- **Location**: Latitude, Longitude, Country, Region and City of the incident.
- **Attack Type**: Type of attack chosen by terrorist. For example, Hijacking, Bombing, Assassination or Assault.
- **Target**: Precise information of who is targeted. For example, Employee, U.S. embassy, name of the person.
- **Target Type**: What kind of target is picked by terrorist. For example, Civilians, Government bodies or Private Business.
- **Weapon Type**: What sort of weapon used for attack. For example, Firearms, Rifle or Bomb.
- **Casualty**: Total number of people Killed and Wounded.
- **Group**: Terrorist group responsible for attack. For example, Left-Wing Militants,Armed Revolutionary Independence Movement (MIRA) or Eritrean Liberation Front.
    - **Motive**: behind the attack. For example, Political, Ideological or Religious.

This dataset is picked to clearly understand when and where a terrorist attack happened.

3.3.2 Drone Wars Database

Drone Wars Data contains 1696 x 17 records which provides information of drone strikes. Each record in dataset includes following details.

- **Date**: Date on which drone strikes occurred.
- **Location**: Latitude, Longitude, Country where drone strike happened.
- **Casualty**: Detailed information of casualties count. For example, maximum number of people killed, maximum civilian reported killed or Maximum number of children killed.
- **President**: President who claimed the responsibility of drone strike. For example, Obama, Trump or Bush.
- **Strike Id**: unique value for each drone strike.

Among the information about the dataset that is mentioned above, only the above-mentioned columns are considered for the research project.

3.2 Data Pre-processing

Before running the model, the datasets are analysed and ensured that they are in the proper format. As discussed in the previous section, it is to be noted that a very limited and relevant features are considered for the cleaning.

The data taken consists of columns that signifies the description of the record and does not provide any absolute information or value to the model. Such columns are ignored since these variables are repetitive and do not serve any purpose of the model. A separate label matrix is followed as mentioned in the previous section to systematize the usage of the same labels used for classification.

Following the objective that the two datasets are to be merged, the plan is to clean the column to merge with the datasets create an overlayed data and re clean the overlayed data. This way only necessary information is obtained, and loss of information or introduction of duplicates can be prevented. The latitude, longitude and the dates of attacks were considered the crucial points of attacks and so the missing records in these columns are clearly, because imputing values would not be ethical.

From the drone wars data, the date column is separated into day, month and year. A separate column known as Casualties is created that consists of the sum of Maximum total people killed by the drone + minimum reported injured + Other Non-civilians killed. The reason to have this column is to understand the overall impact created by the drones, rather than having individual information on the kills.

Now the dataset is ready to be merged. Geospatial Proximity is a technique used to overlay multiple data over each other. This is mainly used to understand temporal outcomes. A buffer of 50 km radius is configured, and the datasets are merged over the latitude and longitude. A separate column “Drone_Strike_Occured” is created to understand that based on the latitude and longitudinal information as to whether the drone strikes have occurred or not. 1 being occurred and 0 being not occurred. Post merging a new data frame is created and is now used as the overlayed dataset which is runs the model.

The overlayed dataset is the final dataset that shall be used for modelling. Since both the columns are merged, the repetitive records and the duplicates are dropped which reduces the data frame to 9957x44 records. A specific set of columns to retain are selected as predictors which are considered for the model. The columns being 'Day_left', 'Month_left', Years_left', 'Region', 'AttackType’, ‘Killed', 'Wounded', 'Target','Group’, ‘Target_type', 'Weapon_type','Motive’, ‘Strike ID', 'President', 'Most Specific Location', 'Most Specific Lat/Long', 'Casualties', 'Drone_Strike_Occurred'. The data types of these columns are checked, and all the non-integer columns are transformed.

The overlayed data is governed by the location to run the model and so, the records that does not have the specific latitude and longitude co-ordinates are dropped. The columns ‘Killed', 'Wounded’, ‘Target’ and ‘motive’ have columns with NA and unknown, which are handled by replacing them with 0,0, unknown and public respectively.

4 Design Specification

4.1 Assumptions made based on Feature Engineering

Based on the feature engineering done and discussed in previous section, following assumptions are made.

- Only the relevant features that captures enough data on the drone strikes and the terrorism are captured to predict the outcomes.
- The features have been redundant are not used to avoid unnecessary model complexity
- The terrorist activities and drone strike become consistent over the period.
- The dependence of drone strikes and the time lag between pre and post drone strikes are considered are expected to be captured appropriately.
- No Variables that add significance to either of the datasets are included
- The data filtered for the model, has sufficient diverse content to capture complex patterns without over-fitting.
- Drone strikes are not directly proportional to terrorist activities.

4.2 Modelling

In continuation to the data preparation the overlayed data is split into testing and training set and the same is modelled using voting classifier. The voting classifies is the best method for this research problem because the random forest helps in addressing the overfitting the SVC handles the non-linearity of the classifiers extending the boundaries and Logistic regression to identify complex patterns. With voting classifier “Hard Voting” technique is used to classify the classes based on the majority of probabilities to produce a better-informed prediction.

4.3 Metrices

- **Accuracy** measures the overall proportion of correctness (predicting severity of the terrorist attack post drone strike) among the classifiers.
- **Precision** to measure the overall proportion of collectively identified positive cases (identifying retaliation attack of drone struck areas) among the classifiers out of all the categorized positive cases

- **Recall** measuring the actual proportion of collectively identified positive cases (total number of drone trike performed) overall in the taken sample.
- **F1 score** to realize the trade-off between actual terror attack occurred because of drone strike and the actual terror attack that happened not because of drone strike people to that of the terror attacks that has no connection with the drone strike. This also helps in understanding over classification if any, that is ranking of more appropriately affected drone strike regions than terror attack regions leading to better prediction.
- **Confusion Matrix** to measure the classifiers that are not falsely classified as Retaliated (True Negatives) and falsely classified as terror attacks which are retaliation because of drone strikes (False Negatives). The above situation of classification is pretty much considered “impact of drone strike”.

5 Implementation

5.1 Tools and Technologies Used for Research

To achieve the desired outcome, following software and library resources were used.

- **Operating System**: 64-bit Windows OS, x64-based processor
- **Processor**: Intel(R) Core (TM) i5-7200U CPU @ 2.50GHz 2.71 GHz
- **Programming Language**: Python
- **Integrated Development Environment (IDE)**: Jupyter Notebook
- **Python Libraries/Modules**: pandas, geopandas, shapely, sklearn, matplotlib, seaborn.

5.2 Implementation of Geospatial Proximity

In order to find relationship between terrorist attacks and drone strikes data, geospatial analysis (Medina, R. and Hepner, G., 2008) is performed. Both the datasets are transformed into GeoDataFrames and buffer zones close to 50 km from the equator around drone strikes are created. Later, both the datasets are integrated by performing spatial join, resulted in locating terrorist events that occurred near the drone strikes locations.

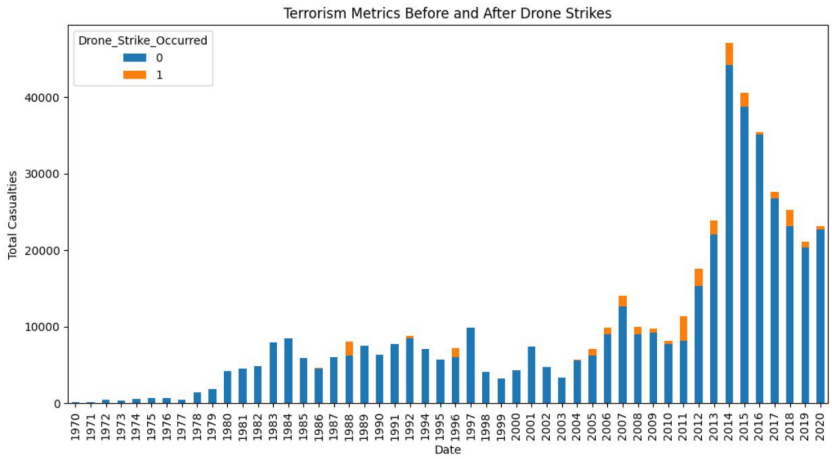

Figure 3: Overlayed dataset of GTD and Drone Strikes

From Figure 3, the overall casualties based on the impact of the drone strikes is taken as represented by amber colour as drone strike conducted, there has been very less that among the casualties the casualties owing to drone strike has been less and the trend seems to be almost same for over a decade. With this information, the overlayed data is led into modelling to classify the severity of the drone strikes.

5.3 Implementation of Voting Classifier

Unnecessary columns are dropped, and the dataset is divided into two segments i.e. 80 (training): 20 (testing). A machine learning pipeline that follows an ensemble voting classifier technique is implemented by combining two individual classifying techniques i.e. Random Forest and SVC and one regression i.e. Logistic Regression. The predictions of RF, SVC and Logistic Regression is combined into one and feed as a single input to the Voting classifier as training data. To get a crisp and informed prediction rather than just a probability number, Hard Voting method is used to train the voting classifier.

Based on the number of casualties, the target feature is modified into grouped class such that if the number killed is less than 10, then the severity is low whilst the number killed is between 10 and 20 then the severity is medium and anything over 20 is considered severe. This grouping is done per isolated incident of terrorist attack after the drone strike.

```
Accuracy: 0.9895424836601308
Classification Report:
              precision    recall  f1-score   support

         Low       0.99      1.00      0.99      2212
    Moderate       0.96      0.52      0.68        42
      Severe       1.00      0.90      0.95        41

    accuracy                           0.99      2295
   macro avg       0.98      0.81      0.87      2295
weighted avg       0.99      0.99      0.99      2295
```

Figure 4: Classification report of the model

From the above Figure 4, it can be observed that the accuracy of the model is 99 percent which means that the model has performed 99 per cent of the times correctly.
However, the following considerations are to be made,

- **Low** class has a precision, recall and F1-Score close to 1.00 that means that the model has "Low" class dominating the dataset.
- **Moderate** class has a very mediocre number of correctly identified instances that is only 62 %, this is probably because the instances is only 42.
- **Severe** class makes close to correct predictions and so it can be seen that the trade off with F1- Score is 95% meaning that the model predicts this class perfectly.

6 Evaluation

The following experiments were conducted, and the results were analysed by plotting confusion matrix and checking their performance factors A comprehensive discussion of the outputs is discussed below.

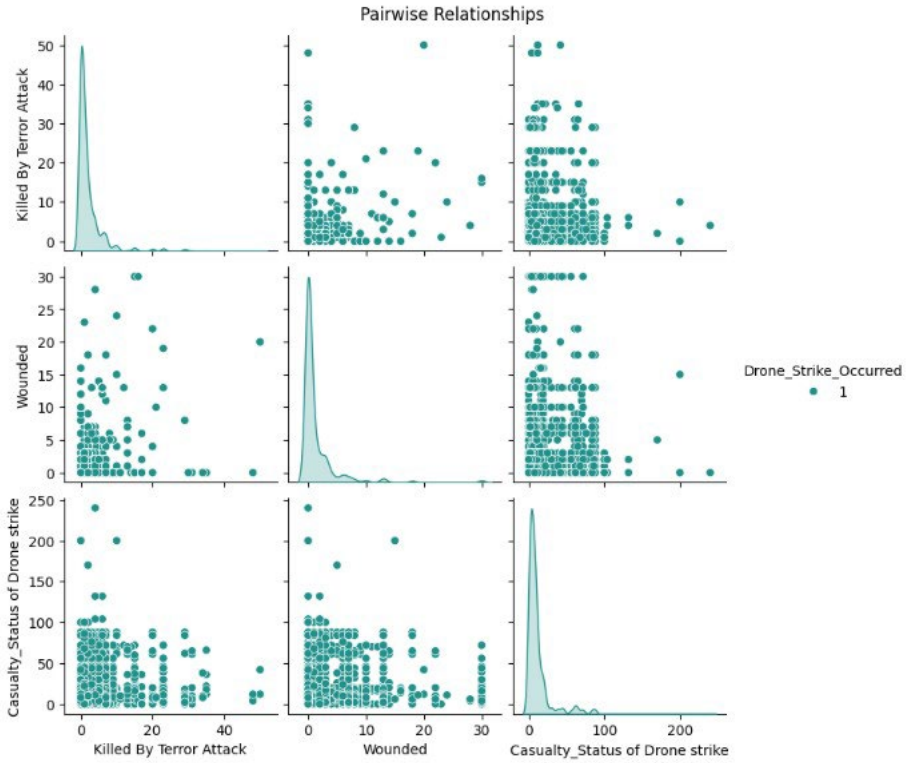

Figure 5: Pair plot of terrorist attack and drone strike

The pair plot from the above figure 5 tells the tale with respect to the presence or absence of the drone strikes.

- **The killed by terror attack** has the longest tail meaning that after most of the drone attacks only fewer casualties are found during the terror retaliation.

- **The Casualty status of Drone strikes** indicates that whenever there are less drone strikes the damage associated tends to be larger.
- **Wounded** seems to have been the highest whenever there is a drone strike but don't always result in casualties. This means that the event of drone strikes is always linked to combined outcomes with wounded and the casualty status of the drone strikes whereas the non-drone strikes always seem to result in more in terms of killed by terror attack.

The visual patterns suggest that non-drone strikes are skewed resulting in contrasting and heavily concentrated drone attacks having broader range of impacts. The following Figure 6 indicates the motive and the region of retaliation (amber colours i.e. green dots - low severity, yellow dots - medium severity and red dots - severe severity) for different regions on different timelines soon after drone strikes as shown in Figure 7 - Drone strikes severity in Pakistan, Figure 8 - Drone strikes severity in Yemen and Figure 9 - Drone strikes severity in Somalia.

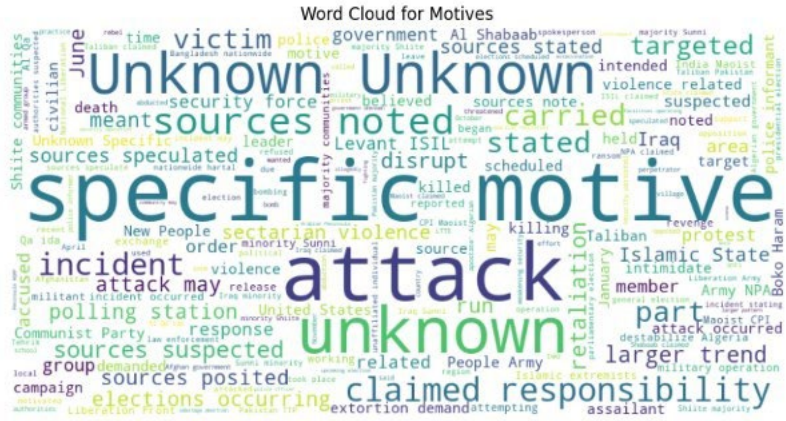

Figure 6: Motive analysis post drone strikes

Figures 7, 8 and 9 are generated by plotting a map that showcases the drone attack severity across Pakistan, Yemen, and Somalia based upon the overlayed data as discussed in the "Data Pre-processing" section.

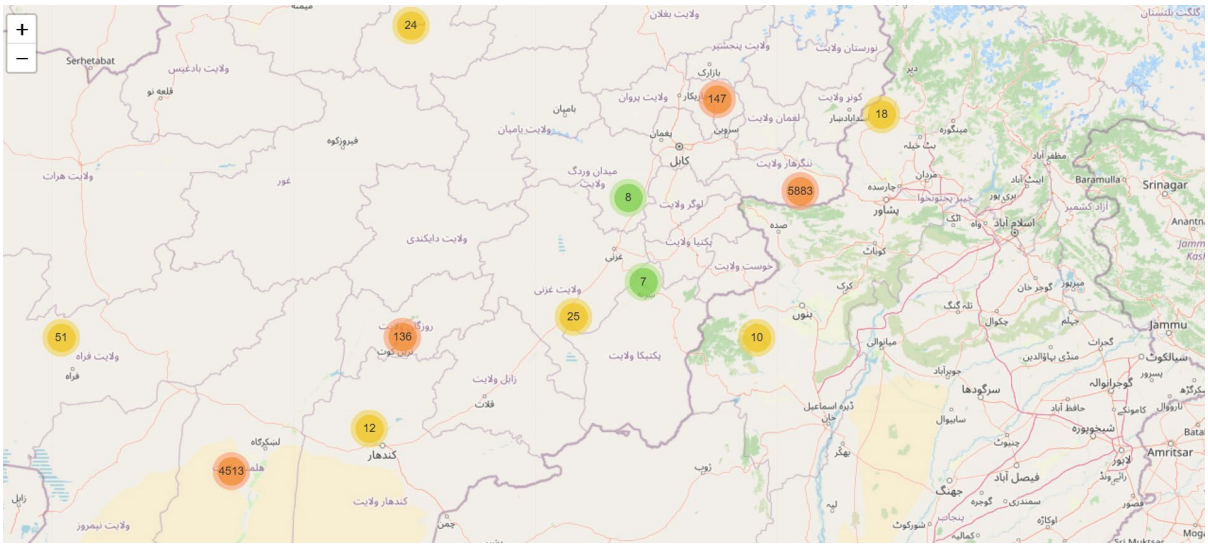

Figure 7 - Drone strikes severity in Pakistan

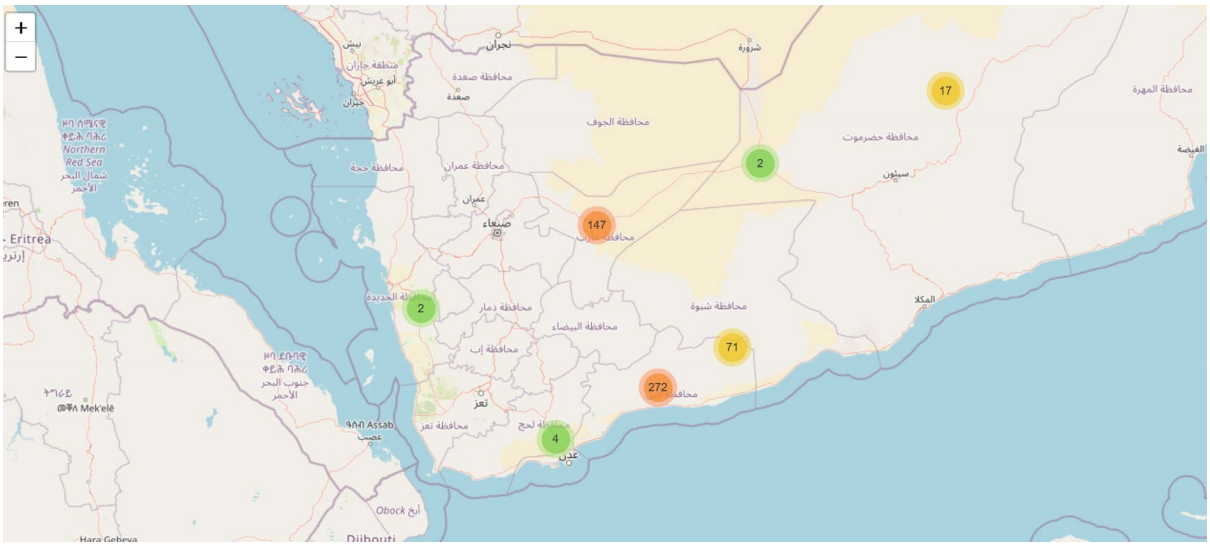

Figure 8 - Drone strikes severity in Yemen

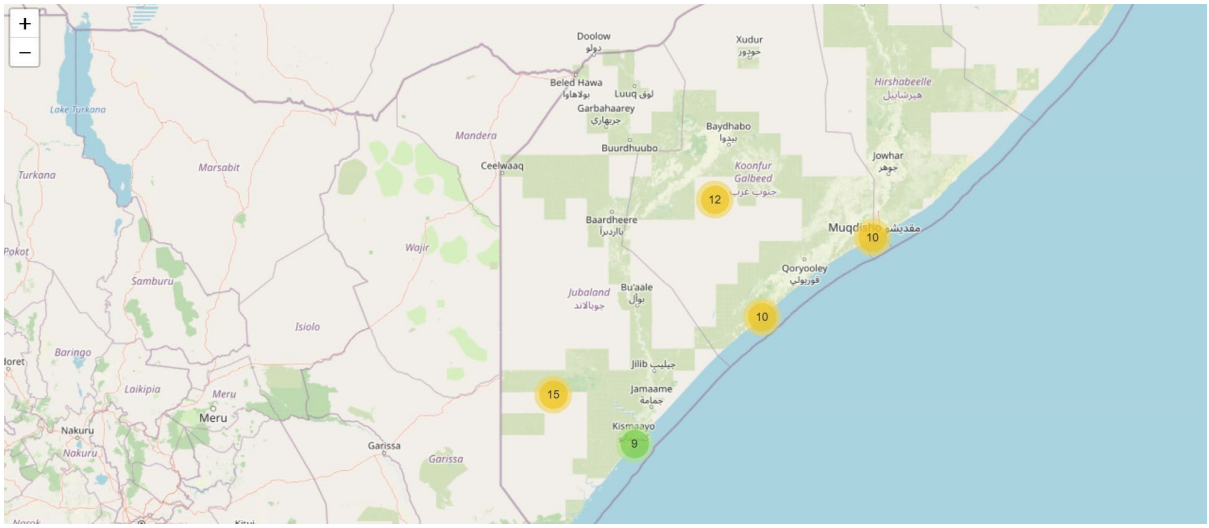

Figure 9 - Drone strikes severity in Somalia

6.1 Evaluation of Confusion matrix

To check the number of correct and incorrect predictions, confusion matrix it plotted, as shown in Figure 10.

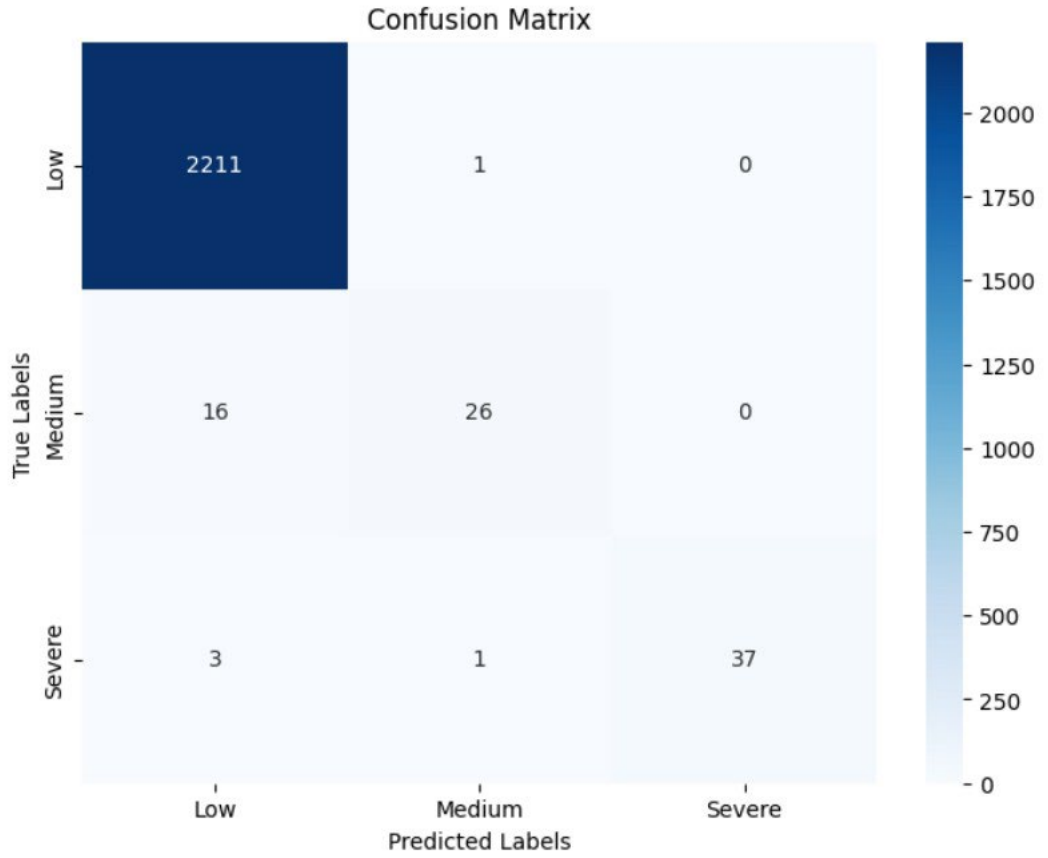

Figure 10: Confusion Matrix

It consists of three classes i.e. “Low” - less attacks after drone strikes, “Medium” - some retaliation by terrorists after drone strikes and “Severe” - massive retaliation after drone strike. Following are the evaluation results of each class.

- **Low**: It is clear from the above matrix that 2211 incidents were correctly categorized as “Low” i.e. true positives. Only 1 incident was incorrectly categorized as “Medium” and 0 incidents were incorrectly categorized as “Severe”.

- **Medium**: In total 16 incidents of medium were incorrectly categorized as “Low”. 26 incidents were correctly categorized as “Medium” and 0 incidents were incorrectly categorized as “Severe”.

- **Severe**: Total 3 incidents were incorrectly categorized as “Low”. Only 1 incident was incorrectly categorized as “Medium” and 37 incidents were correctly categorized as “Severe”.

The above numbers clearly conclude that “Low” is highly and correctly predicted class among the three, where the model predicted almost negligible incorrect predictions. But, on the other hand, the model didn’t perform well for “Medium” class as there are lot of incidents (16 out of 42) which were categorized incorrectly. Model performed well for “Severe” class in

comparison to “Medium” class, only 4 incidents were predicted incorrectly out of 41, which is decent.

6.2 Discussion

With 99.08% accuracy, the classification report states that model (voting classifier) performed well. But the accuracy got skewed due to imbalance in class, as most data is classified in “Low” class. Therefore, confusion matrix is plotted to understand the performance of the model. It is clearly visible that “Low” severity attacks are more in number, which indicates that terrorist groups might do small attacks with minimal damage to avoid being detected or it also suggests that drone strikes have suppressed the terrorist attacks, resulting in low severity.

The Medium severity events are highly misclassified, indicating that the model failed to recognize the incidents where some retaliation happened after the drone strikes but categorized in Low severity. This could be dangerous because what if the terrorist groups change their strategy and switch to complex ideologies like sleeper cells attacks or some coordinated attack to overcome the effects of drone strikes.

There are some severe attacks detected by the model that indicates, few terrorist groups may lead to high casualty incidents.

7 Conclusion and Future Work

By exploring the research question: How do drone strikes impact terrorism dynamics, this study successfully achieved a high accuracy of 99.08 by classifying the severity of terrorist attacks post drone strikes. There are majorly three key takeaways of this research, which is discussed below.

- The presence of low-severity incidents is more, which leads to the conclusion that there is a suppression in large-attacks planned by terrorist organization post drone strikes.
- Difficulty in classifying the medium severity incidents indicates evolving strategies planned by terrorist organizations.
- Few high-impact incidents remain, highlighting the retaliation by the terrorist organizations.

The findings of this research indicated that drone strikes suppressed the frequency of attacks but at the same time gave the tactical adaptions in terrorists. As, there is always room for improvement. This study has laid foundation for future work which are discussed below.

- By studying and understanding the behaviour of individual or groups of terrorist organization post drone strikes can fetch more tailored interventions.
- Time-series models can be used to decode the pattern and severity check before, during and post drone strike.
- By using real time drone strikes data with terrorism data and applying more advanced ML models can fetch real time data for better interpretations.

References

Cronin, A.K., 2013. Why drones fail: When tactics drive strategy. Foreign Aff., 92, p.44. Martin, C., 2018. A means-methods paradox and the legality of drone strikes in armed conflict. In Legal and Ethical Implications of Drone Warfare (pp. 38-71). Routledge

Singh, K., Chaudhary, A.S. and Kaur, P., 2019, August. A machine learning approach for enhancing defence against global terrorism. In 2019 Twelfth International Conference on Contemporary Computing (IC3) (pp. 1-5). IEEE.

Hameed, S., Junejo, F., Zai, M.A.Y. and Amin, I., 2018, November. Prediction of civilians killing in the upcoming drone attack. In 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS) (pp. 1-5). IEEE.

Mahmood, R. and Jetter, M., 2019. Military intervention via drone strikes.

Zahid, K., Nafees, U., Parveen, S. and Afzal, U., 2019, February. Using Neural Network to Predict Unmanned Aerial Vehicle Strikes in Pakistan. In 2019 International Conference on Engineering and Emerging Technologies (ICEET) (pp. 1-6). IEEE.

Thakur, N., Saini, S.S. and Pathak, A.K., 2022, October. Data Mining Model Framework for GTD (Global Terrorism Database). In 2022 International Conference on Cyber Resilience (ICCR) (pp. 1-5). IEEE.

Fisk, K., Merolla, J.L. and Ramos, J.M., 2019. Emotions, terrorist threat, and drones: Anger drives support for drone strikes. Journal of Conflict Resolution, 63(4), pp.976-1000.

Chappelle, W., Skinner, E., Goodman, T., Swearingen, J. and Prince, L., 2018. Emotional reactions to killing in remotely piloted aircraft crewmembers during and following weapon strikes. Military Behavioral Health, 6(4), pp.357-367.

Medina, R. and Hepner, G., 2008. Geospatial analysis of dynamic terrorist networks. In Values and violence: intangible aspects of terrorism (pp. 151-167). Dordrecht: Springer Netherlands.

Kim, S.S., 2023. Predicting and Preventing Drone Terrorism Using Random Forests Machine Learning Algorithm. 한국방사성폐기물학회 학술논문요약집, *21*(1), pp.402-402.

Aljohani, M., Mukkamalai, R. and Olariu, S., 2024. Autonomous strike uavs for counterterrorism missions: Challenges and preliminary solutions. *arXiv preprint arXiv:2403.01022*.

Garcia-Bernardo, J., Dodds, P.S. and Johnson, N.F., 2016. Quantitative patterns in drone wars. *Physica A: Statistical Mechanics and its Applications*, *443*, pp.380-384.