Computer and Systems Engineering Department

Faculty of Engineering, Alexandria University

Graduation Project submitted in partial fulfilment of the B. Sc. Degree

July 2019

Towards Terrorist Group Prediction in Middle East and North Africa

Authors:
Fady Nabil Yacoub
Khaled Fahmy Kassem
Mark Mamdouh Salama

Supervisors:
Prof. Dr. Mohamed S. Abougabal
Prof. Dr. Marwan Torki

Acknowledgement

We are very thankful to Prof. Dr. Mohamed S. Abougabal, and Prof. Dr. Marwan Torki who not only served as our supervisors, but also guided and encouraged us throughout the project, and guided us with great dedication, never accepting less than our best efforts, and allowing for the completion and success of this project.

Abstract

From the start of the 21st century terrorism has been an influencing factor on the economy, tourism and general life aspects of citizens in Middle Eastern and North African Countries (MENA).

The project aims to present a unique solution for the terrorist group prediction. Terrorist group predicting consists is identifying the terrorist group responsible for a certain terrorist attack using a set of common features between each terrorist attack.

An analytical approach is presented to monitor the effects of freedom, happiness and economic prosperity of the attacked countries on identifying the terrorist group. Also, these metrics are used to analyse the nature of a terrorist attack in Middle East and North Africa Countries.

Table of Content

List of Figures

List of Tables

List of Acronyms

Symbol	Meaning	Page Number
MENA	**M**iddle **E**ast and **N**orth **A**frica	**III**
kNN	**k N**earest **N**eighbor	**18**
NB	**N**aive **B**ayes	**19**
DT	**D**ecision **T**ree	**19**
RF	**R**andom **F**orest	**20**
SVM	**S**upport **V**ector **M**achine	**21**
TP	**T**rue **P**ositive	**22**
TN	**T**rue **N**egative	**22**
FP	**F**alse **P**ositive	**22**
FN	**F**alse **N**egative	**23**
GTD	**G**lobal **T**errorism **D**atabase	**27**
GeoEPR	**Geo**-referencing **E**thnic **P**ower **R**elations	**33**
GDP	**G**ross **D**omestic **P**roduct	**34**
NaN	**N**ot **a N**umber	**42**
sklearn	**S**ci**k**it **Learn**	**44**
pandas	**P**ython **D**ata **A**nalysi**s** Library	**44**
numpy	**Num**erical **Py**thon	**44**

Chapter 1

Introduction

1.1 General

Terrorism has been present since antiquity as a perceived method to accomplish a political goal, ideological or religious change in the latter part of the 20th century and into the 21st century.

Terrorism is a difficult term to define, it has no global unified definition till that moment, though it has traits which defines it. The utilization of violence for an ideological gain whether that be political, religious, economic or a general ideological aim can identify terrorism well.

This subset of gains and goals is achieved through influencing public fear in certain countries, communities or races. Especially, the congested communities that contain a spectrum of races and religions and enormous population with low quality of education and high radical polarizing.

Also, terrorism goals may be achieved through affecting countries that have strong economies and wide touristic base in order to affect the main key points that their economy is based on.

Terrorism can be characterized as being adaptive as it always adopts new technologies, methods, tactics and techniques to achieve the desired goals.

Also, some terrorist attacks appear to be disjoint and unrelated to other terrorist attacks. While that appears to be true, the truth is somehow deeper than what appears to naked eye.

1.2 Motivation

Being youth in the twenties of their age who are living in the Middle East one of the most community congested areas and one of the most severely affected regions by terrorism and terrorist attacks all over the globe, we decided to utilize modern technology and our knowledge and abilities to help in making the world generally and our region specifically a more peaceful place for mankind to live and prosper in.

Several machine learning techniques have been applied for terrorist groups prediction problem.

In this study, some traits of attack are identified where it is supposed that each group has its unique fingerprint attacks which allows to point a finger towards the group responsible for the attack and accuse it with the crime.

1.3 Scope of Work

The project follows the typical machine learning process, starting by data preprocessing and ending by testing and obtaining results, passing through features selection and training classifiers.

Data preprocessing is step 1 in the project where it starts from choosing the best datasets suiting the project followed by data cleaning, duplicates removal, non-assigned values compensation or removal. It also includes selecting the tuples of countries that are located in the Middle East and North Africa region.

Features Selection depends mainly on selection of the unique characteristics of each attack like attack type, weapon type and type of targeted victim. Besides, some features related to the country of the attack such as number of ethnic groups, happiness and freedom scores are used to have multiple perspectives on the nature of terrorist attack.

Several classifiers were trained by different combinations of features in order to identify the combination features that obtain maximum results and to be certain of the results obtained by these classifiers.

Finally, a comparison is held between the results obtained in this study and the results obtained by studies which are concerned with the problem of terrorist group prediction.

1.4 Organization of the Report

The report is organized into 6 chapters. In chapter 2, a background about terrorist group prediction techniques will be discussed, and related work and applications which tried to solve the problem will be presented in section 2.5.

In chapter 3, suggested features will be illustrated. Design, implementation and development process and tools will be introduced in chapter 4.

Chapter 5 will include results, comparison with related work results and discussion. Finally, chapter 6 will contain the conclusion and future work of terrorist group prediction and suggestions for terrorism issues data applications.

Chapter 2

Background and Related Work

2.1 Introduction

In chapter 1, a general description of the problem was presented, the motivation to work in a terrorism related topic and terrorist groups prediction, and the scope of work were introduced.

In this chapter, a background about data science and big data is discussed in section 2.2. Also, multi-class classification is presented in section 2.3. Classifiers and Performance measures used in are mentioned in sections 2.4 and 2.5 respectively. Finally, a summary of the related work to solve tasks related to terrorism generally and terrorist group prediction specifically is presented in section 2.6. The need to extend the related work is discussed in section 2.7. Finally, the chapter is concluded by section 2.8.

2.2 Data Science and Big Data [1, pp 2 - 22]

Data Science is an emerging field by itself, whereas Business Intelligence is an explanatory field, data science is an exploratory field.

Data Science is the intersection of:

- Machine Learning: investigates how computers can learn (or improve their performance) based on data. [2, pp 21 - 28]
- Domain Knowledge: The domain in which the application is going to be applied.

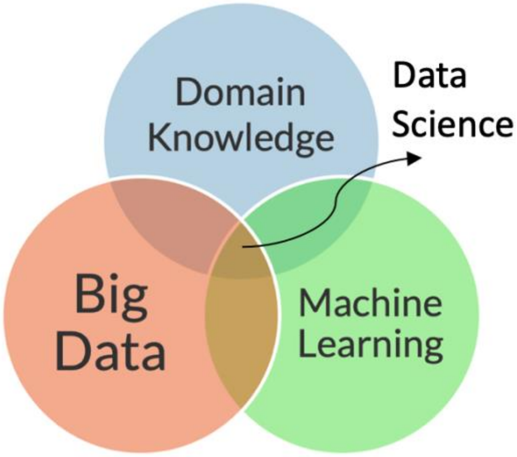


Figure 2.1: Data Science Venn Diagram

Example: Natural Language Processing, terrorist group prediction.

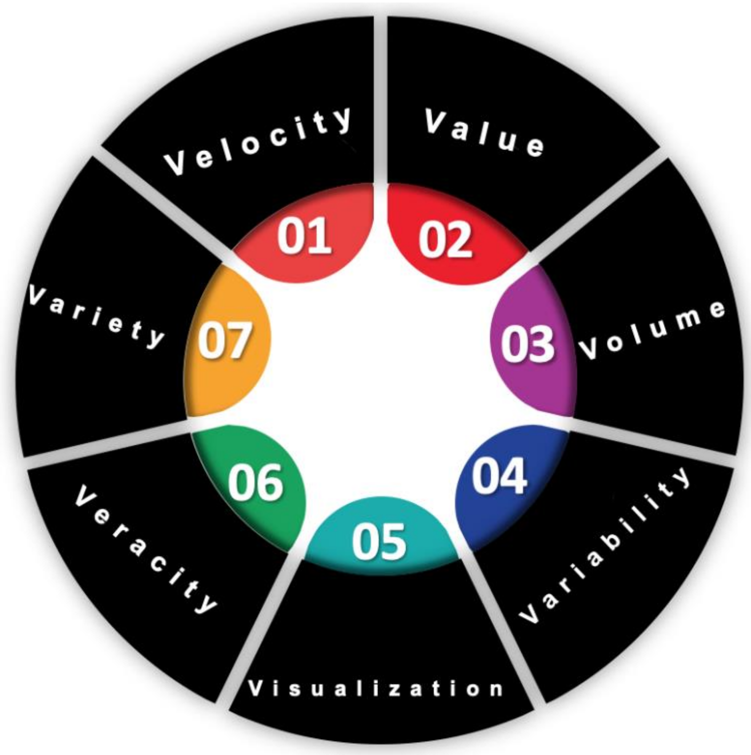


Figure 2.2: Characteristics of Big Data

- Big Data: A huge amount of data that is characterised by seven Vs which are:

1. Velocity: The speed of which the data is generated.
2. Value: Just having Big Data is of no use unless it is turned into value.
3. Volume: The size of the data is very huge.
4. Variability: The data whose meaning is constantly changing.
5. Visualization: The data in manner that is readable and accessible.
6. Veracity: The trustworthiness of the data in terms of accuracy.
7. Variety: The different types of data.

2.3 Multi-Class Classification [2, pp 430 - 432]

Instead of binary classification where a data point is classified to one of 2 classes which is mostly 1 or 0, multi-class classification classifies a point into one of 3 or more classes.

In this project 5 classes are used for each suspected terrorist group to predict from.

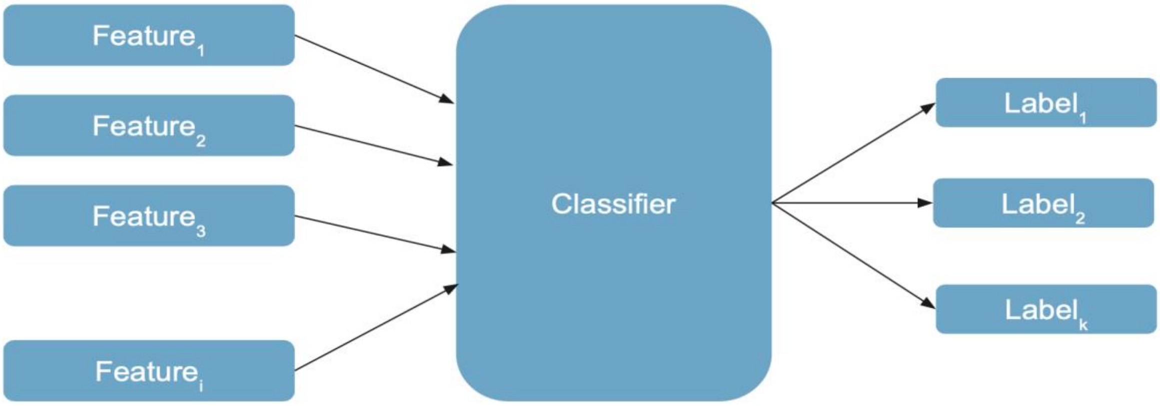


Figure 2.3: Multi-class classification illustration diagram.

2.3.1 All versus All

In this approach it learns a classifier for each pair of classes where a classifier is trained using tuples of the two classes it should discriminate.

For example, having classes A, B and C, the data point is classified to be either A or B or C.

Naive Bayes, Decision Tree and k Nearest Neighbor are examples of that type.

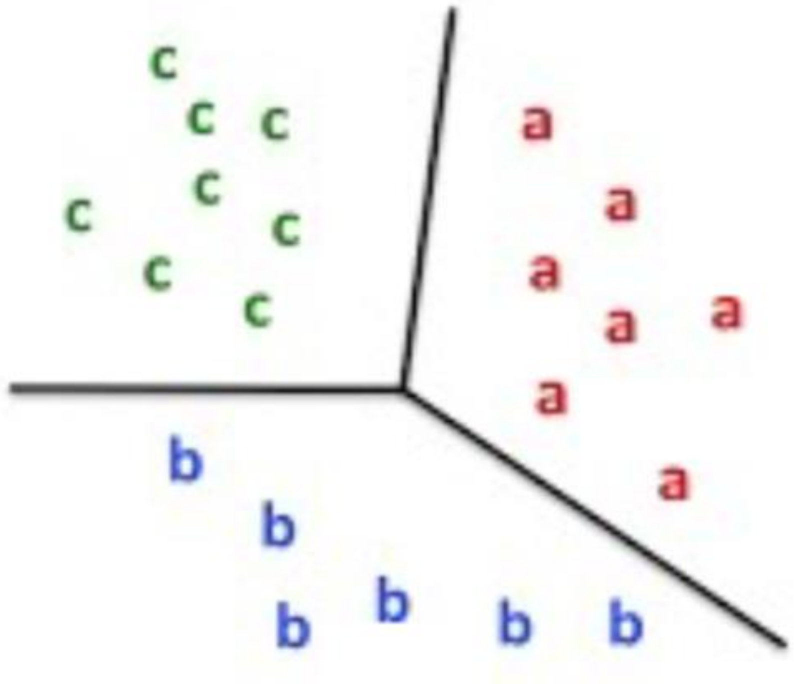


Figure 2.4: Multi-class classification, all versus all technique.

2.3.2 One versus all

It can be also considered Multi-Class Classification using Binary Classification where having classes A, B and C, the data point is classified for each one individually. If we considered class A for example the data point is either A or Not A.

SVM is one of the most common examples of this type.

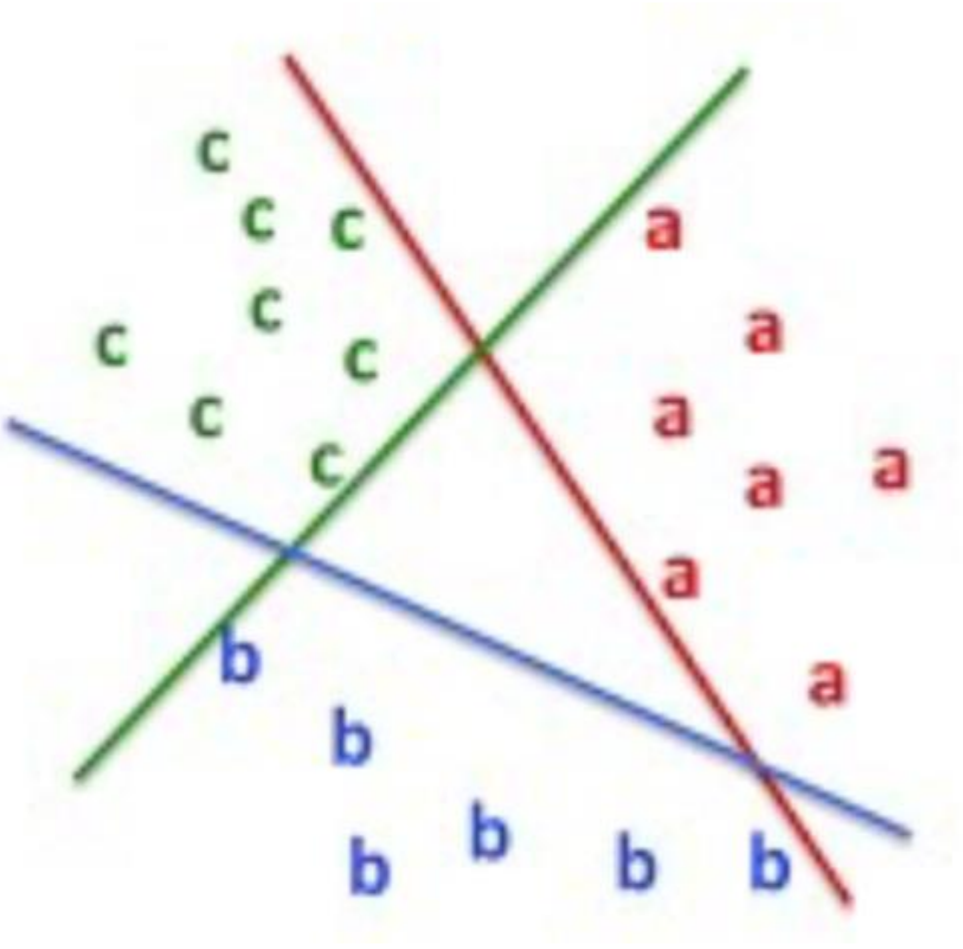


Figure 2.5: Multi-class classification, one vs all technique.

2.4 Classifiers [2]

2.4.1 k Nearest Neighbor (kNN) [2, pp 423 - 425]

A supervised machine learning algorithm which mainly relies on distance calculation between data points and the point required to determine its class.

kNN assumes the point being assigned to the class most common among its k nearest neighbors, where k is defined initially before the algorithm works.

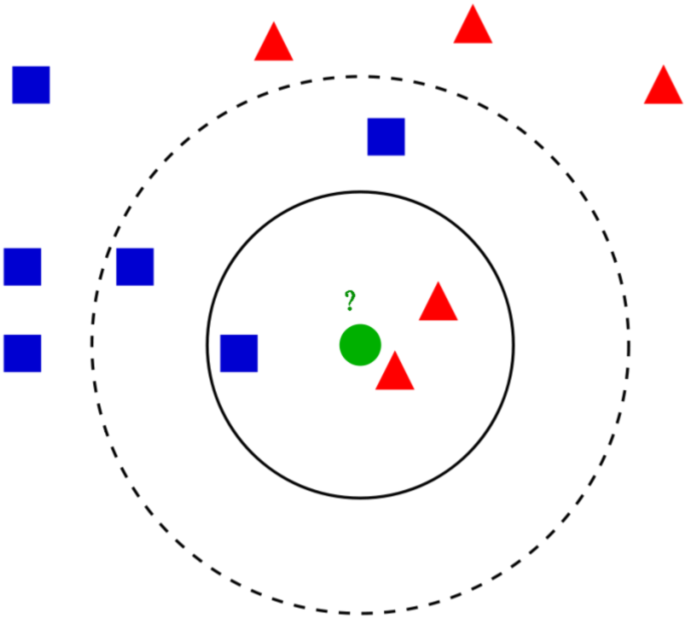


Figure 2.6: kNN illustration. Photo from[3].

2.4.2 Naive Bayes (NB) [2, pp 351 - 355]

A probabilistic classifier based on applying Bayes' theorem,

$$P(Y \mid X) = \frac{P(X \mid Y)P(Y)}{P(X)}. \quad (2.1)$$

Class conditional independence is a naive assumption made so that the algorithm is simplified, thus it is named “Naive”,

$$P(X_1, X_2, \ldots, X_d | Y_j) = P(X_1|Y_j)P(X_2|Y_j)\ldots P(X_d|Y_j). \quad (2.2)$$

In other words, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

2.4.3 Decision Tree (DT) [2, pp 355 - 359]

A decision tree is a flowchart-like structure in which each internal node represents a feature, the branch represents a decision rule and each leaf node represents the outcome label.

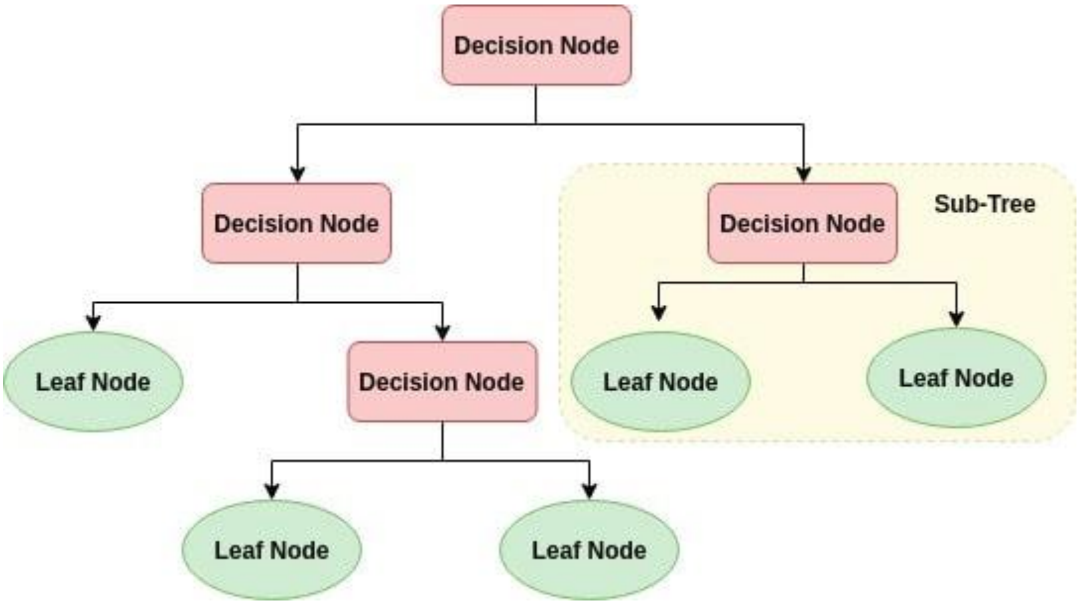


Figure 2.7: Decision Tree. Photo from [4]

For a decision tree, the decision boundaries are parallel to axes since a test condition involves a single attribute at a time.

Decision tree is also suitable for multi-class classification.

2.4.4 Random Forest (RF) [2, pp 382 - 383]

Random Forest is an ensemble learning algorithm where it is formed by a bunch of Decision Trees trained on:

1. Randomly (with replacement) selected data from the training set (Bagging).
2. Random selection of features for each tree in the forest.

A majority vote is done over the results of all the trees in the forest and a decision is made accordingly.

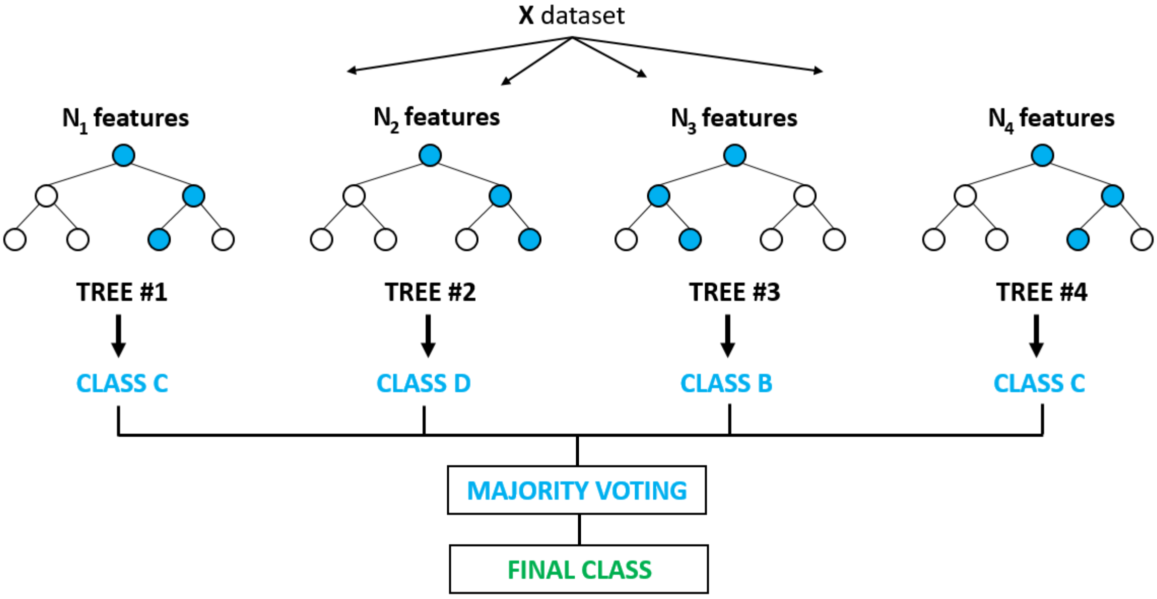


Figure 2.8: Random Forest. Photo from [5]

2.4.5 Support Vector Machine (SVM) [2, pp 408 - 415]

A linear binary classifier that aims to find a hyperplane that maximizes the margin between data points of the 2 classes.

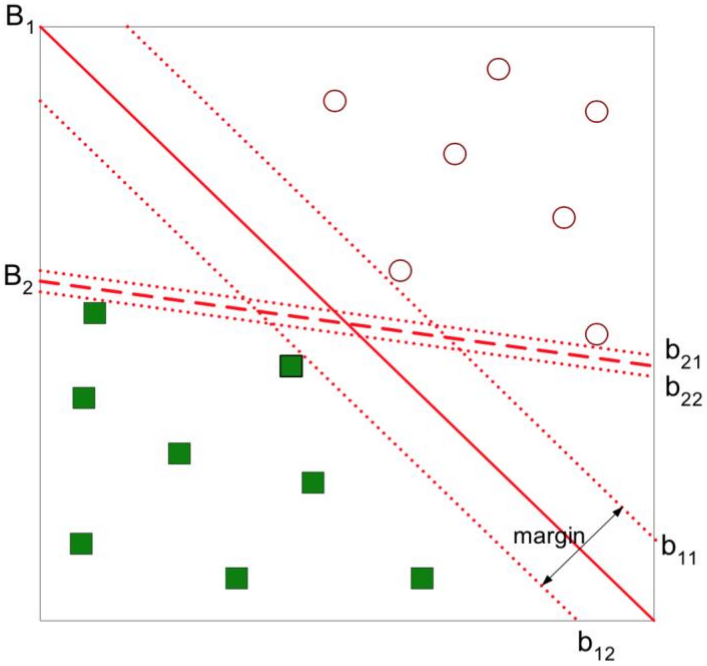


Figure 2.9: SVM, aiming to choose the hyperplane with maximum margin.

Although, SVM is a linear classifier, mathematical kernels can be used to transform the data to an n-dimension in which it is linearly separable.

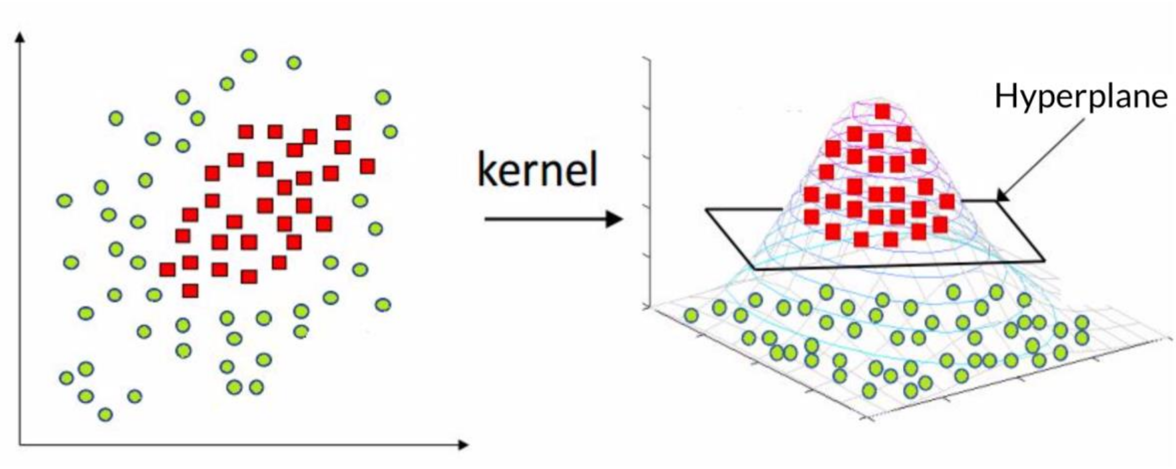


Figure 2.10: Using kernels for dimension transformation in SVM.

Being a binary classifier applied in a multi-class classification problem like terrorist group prediction, SVM follows the One vs All or One vs the Rest method for multi-class classification.

2.5 Performance Measures [2, pp 364 - 369]

2.5.1 Accuracy

Accuracy refers to how obtained values are close to the real values or it is the ratio of true predicted to the whole data.

Accuracy is more descriptive most commonly in cases of balanced class, for example (in 100 samples, 20 sample for each of 5 classes)

Accuracy equation is as follows:

$$Accuracy = \frac{TP+TN}{all} . \qquad (2.3)$$

2.5.2 Precision

Precision is how close two or more measurements are to each other. It expresses the ratio of correctly predicted (True Positive) values to the ratio of predicted as positive.

The equation of precision can be expressed as follows:

$$Precision = \frac{TP}{TP+FP}. \qquad (2.4)$$

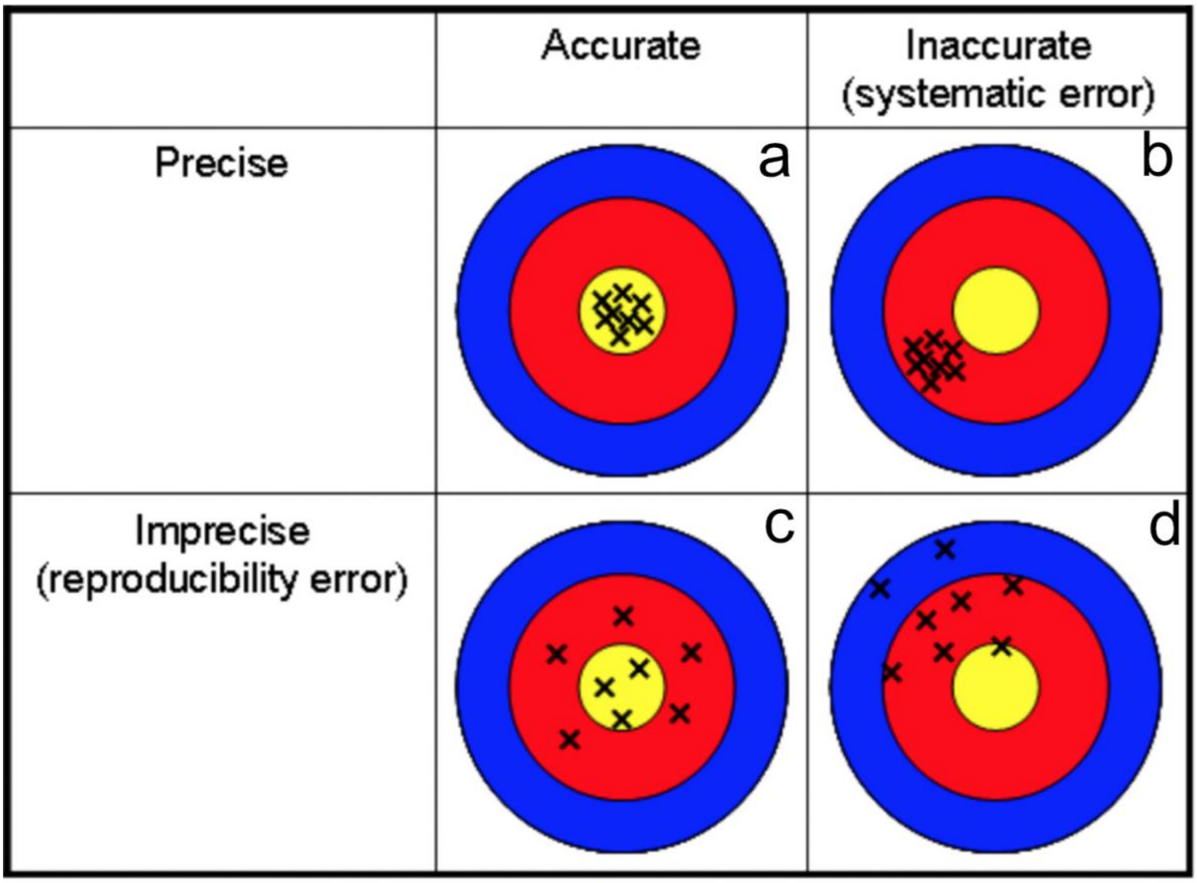


Figure 2.11: Illustrating accuracy and precision. Photo from [6]

Having precise results (compact) even in case of inaccurate results as in (b) add figure number helps to figure out the trend in errors and fix it once and for all, therefore the results go from inaccurate and precise as in (b) to accurate and precise as in (a).

2.5.3 Recall

The ratio of correctly predicted positive observations to the all observations in actual class.

Its equation can be expressed as follows:

$$Recall = \frac{TP}{P} = \frac{TP}{TP+FN}. \tag{2.5}$$

2.5.4 F-Measure or F-Score

It is a weighted measure of precision and recall together, where it can be expressed as

$$F_\beta\ Score = \frac{(1+\beta^2)*Precision*Recall}{\beta^2*Precision+Recall}. \tag{2.6}$$

As β increases recall weight (importance) increases and vice versa.

F-measure is used mainly in case of unbalanced classes.

The most common F-measure is F_1-measure which gives equal importance for both precision and recall.

The equation of F_1-Score is expressed as follows:

$$F_1\ Score = \frac{2*Recall*Precision}{Recall+Precision}. \tag{2.7}$$

2.5.5 Micro Average and Macro Average [7]

All of the performance measures mentioned are mainly explained on binary classes. Since this project covers multi-class classification problem, little tweaks will be applied on the previously explained performance measures to work for the multi-class classification.

Two types of average exist, the first one is called micro average which (in case of precision) instead of dividing the TP of one class, it sums all the true positives of all classes and calculate its ratio to summation of total true positives and total false positive. Recall is calculated using a similar way.

Since its operations are on a low-level scale (true positive, false positive and false negative) it is called "micro".

The second type is macro average which is a normal average where (in case of precision) it divides the sum of all values by the number of classes.

	Precision	Recall
Micro Average	$\frac{\Sigma TP_i}{\Sigma TP_i + \Sigma FP_i}$	$\frac{\Sigma TP_i}{\Sigma TP_i + \Sigma FN_i}$
Macro Average	$\frac{\Sigma Precision_i}{\#of\ Classes}$	$\frac{\Sigma Recall_i}{\#of\ Classes}$

Table 2.1: Micro Average and Micro Average performance measures.

2.5.6 The equivalence between Accuracy and Micro Average F-Score [7]

In multi-class classification, since at a time one class is considered the positive class and all the other are considered negative and then another class is considered positive and so on, therefore

$$\Sigma FP = \Sigma FN. \quad (2.8)$$

and $\Sigma TP = \Sigma TN.$ (2.9)

From equation (2.8) and (2.9) it can be deduced that both Recall and Precision are equal.

Concerning accuracy,

$$\because Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.10)$$

Using equations (2.8) and (2.9) to substitute in (2.10)

$$\begin{aligned} Accuracy &= \frac{2TP}{2(TP+FP)} \\ &= \frac{TP}{TP+FP} \end{aligned}$$

$$\therefore Accuracy = Recall = Precision. \quad (2.11)$$

Thus, from what was discussed it is clear why could micro average of precision, micro average of recall and accuracy be equal in the case of multi-class classification.

2.6 Summary of Related Work

Four research papers were surveyed to establish the concrete idea of this study.

Paper [8] and paper [9] both dealt with terrorist events locations identification and prediction through creating a grid world map and predicting the potentiality of terrorist events taking place in a certain grid (which represents a set of locations in a certain place in the world) according to several input features such as location, terrorist events history in that location, population density and ethnic diversity… etc.

Each of papers [8] and [9] dealt with different regions to apply their studies on, and different approaches for the solution.

On the other side, papers [10] and [11] dealt with the problem of terrorist group identification and prediction which is the same problem this study is dealing with.

Papers [10] and [11] are common in depending on identifying a virtual fingerprint consisting of the type of attack, country, weapon type and some other features. The significant difference between the two approaches of the two papers is using different regional scopes as paper [10] worked on a worldwide scope while [11] concerned a smaller and more specific region which is Middle East and North Africa region, similar to the scope concerned by this project.

	Paper [10]	**Paper [11]**
Used Datasets	Global Terrorism Dataset [12]	GTD
Region Concerned	Worldwide	Middle East and North Africa
Time Period	1970 : 2012	2009 : 2013

Table 2.2: Comparison between paper [10] and paper [11].

In the following subsections, the classifiers, features and performance measures used in each paper will be presented.

2.6.1 Global Terrorism Database (GTD)[12]

Global Terrorism Database is the world's largest dataset available on terrorism incidents from 1970 till 2017.

Global Terrorism Database is collected under the supervision of terrorism research experts from the University of Maryland. It covers multiple details and several aspects of each terrorism event all over the globe.

It includes 135 attributes including year, month, type of the attack, weapon type, specific location of the event (longitude and latitude) and most important of them all is terrorist group name which is used as the classifier’s label in this study.

GTD is considered the cornerstone of this project since it contains all the attributes needed to build a terrorist group predictor.

2.6.2 Classifiers

		[8]	[9]	[10]	[11]
Supervised Learning	**SVM**	✓	✗	✗	✓
	Deep Learning	✓	✗	✗	✗
	Random Forest	✓	✗	✗	✓
	kNN	✓	✗	✓	✓
	Naive Bayes	✗	✗	✓	✓
	Decision Tree	✗	✗	✓	✓
	decision stump (DS)	✗	✗	✓	✗
Unsupervised Learning	**k-means**	✗	✓	✗	✗
	DBSCAN	✗	✓	✗	✗
	BIRCH	✗	✓	✗	✗

Table 2.3: Summary of machine learning algorithms used in related work.

From the previous table it can be noticeable that papers [8], [10] and [11] have dealt with their problems through various classification techniques while paper [9] was the only one that followed unsupervised learning techniques.

2.6.3 Classifiers Features

	[8]	[9]	[10]	[11]
Latitude	✓	✓	✗	✗
Longitude	✓	✓	✗	✗
Month	✗	✗	✓	✓
City	✗	✗	✓	✓
Country	✗	✗	✓	✓
Weapon Type	✗	✗	✓	✓
Attack Type	✗	✗	✓	✓
Target	✗	✗	✓	✓
Group Name	✗	✗	✓	✓
Year	✗	✗	✗	✓
Region	✗	✗	✗	✓
Provstate	✗	✗	✗	✓
Distance to major navigable lake	✓	✓	✗	✗
Distance to major navigable river	✓	✓	✗	✗
Distance to ice—free Ocean	✓	✓	✗	✗
Average precipitation	✓	✓	✗	✗

Average temperature	✓	✓	✗	✗
Ethnic diversity	✓	✓	✗	✗
Major drug regions	✓	✓	✗	✗
Nighttime lights	✓	✓	✗	✗
Population density	✓	✓	✗	✗
Topography	✓	✓	✗	✗
Transportation site	✗	✓	✗	✗
Religious places	✗	✓	✗	✗
Political places	✗	✓	✗	✗
Catering outlets	✗	✓	✗	✗
Accommodation outlets	✗	✓	✗	✗
Happiness Score	✗	✗	✗	✗
Freedom Score	✗	✗	✗	✗
Economy	✗	✗	✗	✗

Table 2.4: Summary of related work features.

2.6.4 Performance Measures

	[8]	[9]	[10]	[11]
Accuracy	✗	✗	✓	✓
ROC-AUC	✓	✗	✗	✗
F-Measure	✗	✓	✗	✓

Table 2.5: Summary of related work performance measures.

From table 2.5 it can be noticed that accuracy is common between paper [10] and paper [11] while F-measure is common between paper [9] and paper [11].

2.7 Need to Extend the Related Work

Close study of table 2.3 reveals that there is a need to:

1. Study a country's economic and freedom effects on the nature and factors of a terrorist attack.
2. Study the effect of ethnic diversity on the terrorist attacks in the countries concerned with the study.
3. Hybridize the problem of predicting terrorist groups [10] & [11] with the problem of predicting terrorism events locations [8] & [9].

Close study of table 2.4 reveals that there is a need to:

1. Introduce deep learning to the problem of terrorist groups prediction.
2. Add more classifiers and performance measures.
3. Add more datasets.

2.8 Scope of the Project

In section 2.7, the need to extend the related work was discussed. This project focuses on studying the country's economic and freedom effect on the nature of the terrorist attack.

Also, studying the effect of ethnic diversity in each country is taken into consideration in this study. These experiments are evaluated using more than one performance measure.

The rest of needed extensions are left for future work which is discussed in chapter 6.

2.9 Conclusion

In this chapter, a data science and machine learning background were presented and many works related to the terrorist events and groups were discussed. Also, the need to extend the previous work and the scope of this project were discussed.

In the next chapter, the main features of terrorist group prediction and the selected performance measures will be presented. Both are based on the related work presented in this chapter.

Chapter 3

The Proposed Terrorist Groups Prediction

3.1 Introduction

In the previous chapter, a background about terrorist group classification was presented. Also, classifiers and datasets were introduced.

After deep surveillance for the 4 papers mentioned in chapter 2, it was decided to settle on extending paper [10] and [11] work since it was more practical and suitable than re-implementing and extending the work of paper [8] and paper [9] which needed licenses purchase for the geographical map pre-processing problem and datasets.

In this chapter the new datasets adopted in this study are mentioned in section 3.2. In section 3.3 classifiers used are presented followed by the features used to train the classifiers in section 3.4. The tuple of classifier features used as an input for the terrorist group classifier is presented in section 3.5. The chapter is concluded by section 3.6.

3.2 Datasets

3.2.1 Geo-referencing Ethnic Power Relations (GeoEPR) [13]

GeoEPR dataset provides information about every politically relevant ethnic group. Each country is paired with the names of the ethnic groups present in it and some other attributes like the duration an ethnic group is present in a country in case that ethnic group exists no more.

3.2.2 World Happiness Report [14]

“The overall rankings of country happiness are based on the pooled results from Gallup World Poll surveys from 2015-2017”

World Happiness Report contains 10 attributes at the top of them happiness score, freedom score and GDP per capita (Economy) which were used in this study.

The 3 versions of World Happiness Report at years 2015, 2016 and 2017 were used in this study to analyze the change and stability against the time factor in the 3 most recent years to the time of the study.

The following table illustrates the datasets used in this project:

Dataset	Source	Size	Release
Global Terrorism Database (GTD) [12]	National Consortium for the Study of Terrorism and Responses to Terrorism (START), University of Maryland. Global Terrorism Database (GTD), Version 3. https://www.kaggle.com/START-UMD/gtd/home.	181692 Tuples, 135 Columns	2017
Geo-referencing Ethnic Power Relations (GeoEPR) [13]	Center for Comparative and International Studies (CIS), International Conflict Research, ETH Zurich, https://icr.ethz.ch/data/epr/#ed	1471 Tuples, 10 Columns	2018.1
World Happiness Report [14]	Sustainable Development Solutions Network (A global initiative for United Nations). (2017). World Happiness Report, Version 2. https://www.kaggle.com/unsdsn/world-happiness	155 Tuples, 12 Columns	2015, 2016, 2017

Table 3.1: Description of the 3 datasets used in this project.

3.3 Classifiers

The following table shows the classifiers used in the 4 papers and the classifiers used in this study:

		[8]	[9]	[10]	[11]	**Proposed**
Supervised Learning	**SVM**	✓	✗	✗	✓	✓
	Deep Learning	✓	✗	✗	✗	✗
	Random Forest	✓	✗	✗	✓	✓
	kNN	✗	✗	✓	✓	✓
	Naive Bayes	✗	✗	✓	✓	✓
	Decision Tree	✗	✗	✓	✓	✓
	decision stump (DS)	✗	✗	✓	✗	✗
Unsupervised Learning	**k-means**	✗	✓	✗	✗	✗
	DBSCAN	✗	✓	✗	✗	✗
	BIRCH	✗	✓	✗	✗	✗

Table 3.2: The proposed classifiers for terrorist group prediction.

Five classifiers are proposed to be used in this study which are:

- Support Vector Machine
- Random Forest
- K Nearest Neighbor
- Naive Bayes
- Decision Tree

Since this project is an extension on paper [10] and [11] work, the 5 chosen classifiers are similar to those used in paper [11] whereas classifiers of paper[10] are a subset of those used in paper[11] except decision stump (DS).

3.4 Classifiers Features

The following table illustrates the features used for terrorist group prediction:

	[8]	[9]	[10]	[11]	**Proposed**
Latitude	✓	✓	✗	✗	✗
Longitude	✓	✓	✗	✗	✗
Month	✗	✗	✓	✓	✓
City	✗	✗	✓	✓	✓
Country	✗	✗	✓	✓	✓
Weapon Type	✗	✗	✓	✓	✓
Attack Type	✗	✗	✓	✓	✓
Target	✗	✗	✓	✓	✓
Group Name	✗	✗	✓	✓	✓
Year	✗	✗	✗	✓	✓
Region	✗	✗	✗	✓	✗
Provstate	✗	✗	✗	✓	✓
Distance to major navigable lake	✓	✓	✗	✗	✗

Distance to major navigable river	✓	✓	✗	✗	✗
Distance to ice—free Ocean	✓	✓	✗	✗	✗
Average precipitation	✓	✓	✗	✗	✗
Average temperature	✓	✓	✗	✗	✗
Ethnic diversity	✓	✓	✗	✗	✓
Major drug regions	✓	✓	✗	✗	✗
Nighttime lights	✓	✓	✗	✗	✗
Population density	✓	✓	✗	✗	✗
Topography	✓	✓	✗	✗	✗
Transportation site	✗	✓	✗	✗	✗
Religious places	✗	✓	✗	✗	✗
Political places	✗	✓	✗	✗	✗
Catering outlets	✗	✓	✗	✗	✗
Accommodation outlets	✗	✓	✗	✗	✗

Happiness Score	✗	✗	✗	✗	✓
Freedom Score	✗	✗	✗	✗	✓
Economy	✗	✗	✗	✗	✓

Table 3.3: The proposed features for training the classifiers.

Region is not used in this study on contrary to paper [11], since all data tuples are in the same region (Middle East and North Africa) so (after comparing the results before and after adding it) it appeared not to be an effective feature.

Happiness Score, Freedom Score and Economy are not used in any of the previous work as they are added as contribution by this project. These features are taken from World Happiness Report [14] with its releases of 2015, 2016 and 2017.

Economy is expressed by Gross Domestic Product per Capita which is known as GDP per Capita. GDP per capita is a measure of a country's economic output that accounts for its number of people. It divides the country's gross domestic product by its total population. That makes it the best measurement of a country's standard of living. It tells you how prosperous a country feels to each of its citizens.

Ethnic groups diversity can be an effective feature since the Middle East and North Africa contains a lot of diverse ethnic groups.

3.5 Tuple of the Classifier Features (Classifier Input)

Close study of the detailed classification features listed in section 3.3 suggests using the following classification tuple of 18 features as an input for the classifiers:

Month
City
Country
Weapon Type
Attack Type
Target
Year
Provstate
Ethnic diversity
Happiness Score 2015
Freedom Score 2015
Economy 2015
Happiness Score 2016
Freedom Score 2016
Economy 2016
Happiness Score 2017
Freedom Score 2017
Economy 2017

Table 3.4: Features tuple.

3.6 Performance Measures

The following table illustrates the proposed performance measures in this study:

	[8]	[9]	[10]	[11]	**Proposed**
Accuracy	✗	✗	✓	✓	✓
ROC-AUC	✓	✗	✗	✗	✗
F-Measure	✗	✓	✗	✓	✓

Table 3.5: Proposed performance measures.

It was preferred to use both accuracy and F-measure in this study. Since the data is unbalanced, it was necessary to use F-measure to avoid misleading results.

On the other side, accuracy was used for several reasons:

1. To compare the results of this study with paper [10] that uses the accuracy as the only performance measure.
2. Still holds in monitoring the effects of the new features introduced in this study as explained in detail in chapter 5
3. To emphasize on the importance of using the suitable performance measure with the suitable data.

3.7 Conclusion

In this chapter the datasets, classifiers, main features used for terrorist groups classification and the tuple of the classifier features were introduced. At the end of the chapter, the performance measures used in this study were presented.

In the next chapter, the project design and implementation will be discussed.

Chapter 4

Design and Implementation

4.1 Introduction

In chapter 3, the features of the terrorist group classifier were presented.

In this chapter, the design of the project's typical machine learning process is explained in section 4.2. The main tools used in development, coding and presenting results are presented in section 4.3. The chapter is concluded by section 4.4.

4.2 Architectural Design

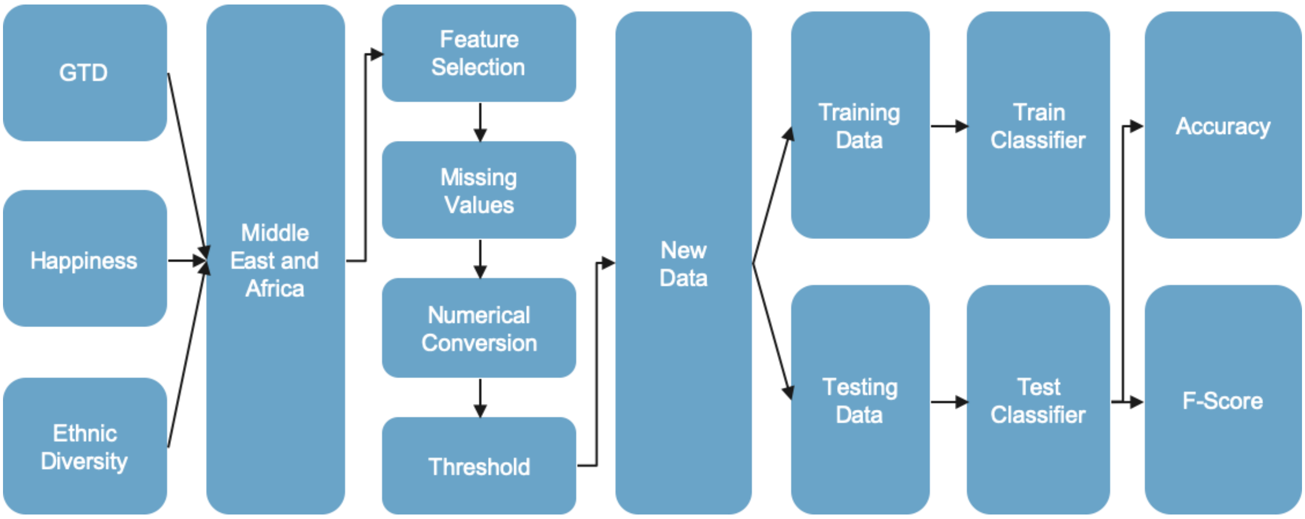


Figure 4.1: The project's architectural design.

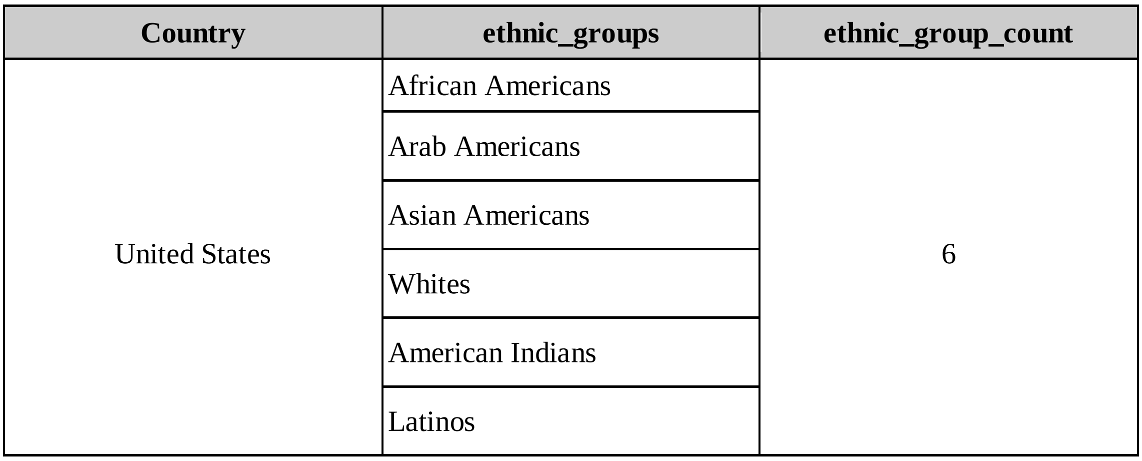
4.2.1 Data Preprocessing and Cleaning

As mentioned in section 3.2 and as what appears in the diagram (Figure 4.1) 3 datasets are used where Global Terrorism Database is the cornerstone of this project.

GTD, World Happiness Report data and GeoEPR are joined on the common country where the Non Middle Eastern nor North African countries are dropped from the dataframe.

All attributes are dropped from the dataframe except those which were mentioned in section 3.4 and the class label (Terrorist Group).

GeoEPR data (ethnic groups in each country) was used as an ethnic group count instead of using the ethnic groups names as shown in the following table:



Country	ethnic_groups	ethnic_group_count
United States	African Americans	6
	Arab Americans	
	Asian Americans	
	Whites	
	American Indians	
	Latinos	

Table 4.1: GeoEPR ethnic groups count example.

As an experimental choice, it was decided to use the data starting from year 2000 for the following reasons:

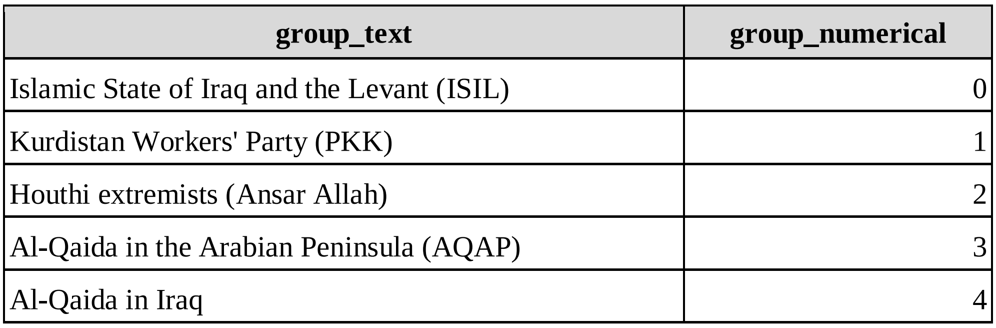
1. To emphasize on more recent events from the start of the 21st century instead of training with old data from 1970 to 2000 which might not be so useful since most of the terrorist groups of that period disappeared and the terrorist attacks style changed a lot due to the technological revolution.
2. To emphasize on the events after the September 2001 attacks which influenced and affected the global security and the behavior of many terrorist groups.

Like most or certainly all of the data collected over a long period of time, GTD had some few tuples with missing attributes or “not a number” (NaN) values. A decision was made to drop the tuples containing missing or NaN values to avoid having an outlier or misleading values in case of applying one of the well known values compensation techniques.

4.2.2 Numerical Conversion

Categorical attributes (i.e strings) like city, country … etc. and the class label (group name) where converted to numerical value in order to be able to train the classifiers.

The following table is an example for numerical conversion for textual values:



group_text	group_numerical
Islamic State of Iraq and the Levant (ISIL)	0
Kurdistan Workers' Party (PKK)	1
Houthi extremists (Ansar Allah)	2
Al-Qaida in the Arabian Peninsula (AQAP)	3
Al-Qaida in Iraq	4

Table 4.2: Numerical conversion of categorical data.

4.2.3 Thresholding

Having 636 terrorist groups that executed terrorist events in the MENA region from the interval between 2000 to 2017 indicates having 636 class labels to predict from which is a very large number, especially when some groups were not involved in more than tens of events, which is few compared with those who were involved in hundreds or maybe thousands of them.

Following the same process of thresholding followed in [10] and [11], 5 terrorist groups where only selected from the 636. Those 5 groups were responsible for more than 90% of the total attacks in the region at that interval of time.

The selected groups are the same mentioned in table 4.2 at section 4.2.2.

4.2.4 Training Data and Testing Data

As mentioned in section 4.2.1, data from the year 2000 to 2017 was used in this study.

The total number of tuples after performing filtration and preprocessing is around 10,000 tuples, the details of training and testing data are as mentioned in the following table:

	Number of Tuples	Percentage	Time Interval
Training Data	7842	~80%	[2000, 2016]
Testing Data	1648	~20%	2017

Table 4.3: Training and testing data details.

4.3 Tools Used

4.3.1 Programming Language Used

Python3 programming language was used as the main language for this project development.

4.3.2 Development Environment: Google Colaboratory [15]

“Colaboratory is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud.”

Google Colab was used as the development environment from the beginning of the project as it facilitates importing external python libraries such as pandas and numpy.

4.3.3 External Libraries

The following libraries were used:

- Scikit Learn (sklearn) [16]: data-mining library used to provide implementation for different machine learning algorithms including supervised and unsupervised learning.
- Pandas [17]: “an open source library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.”

- Numpy [18]: a fundamental package for scientific computing with Python.

4.3.4 Github

The project's colab notebook was integrated with a repository on GitHub from day 1 of the project to ensure version control and facilitate development for the team members.

4.3.5 Google Sheets

A special program by Google for tables creation, mathematical calculations and data visualization.

Excel had a great role in the project where it was used for result visualization as in section 5.3 and section 5.4 and data calculations through the whole project.

4.4 Conclusion

In this chapter, the architectural design of the project and the tools used for development were discussed.

In the next chapter results obtained and results comparison with previous work will be explained.

Chapter 5

Results and Comparisons

5.1 Introduction

In the previous chapter, implementation and design of the project were discussed side by side with the tools used in development.

In this chapter, the contribution of this project over previous work is highlighted in section 5.2. The results of testing the classifiers are presented in section 5.3. In section 5.4 an analysis for the obtained results is held followed by comparison with previous work results. The chapter is concluded by section 5.5.

5.2 Highlighting Contribution Over Previous Work

	This Project	**Paper [10]**	**Paper [11]**
Used Datasets	GTD, World Happiness Report & GeoEPR	GTD	GTD
Region Concerned	Middle East and North Africa	Worldwide	Middle East and North Africa
Time Period	2000 : 2017	1970 : 2012	2009 : 2013
Training Interval	2000 : 2016	-	Using 10-Fold Cross Validation
Testing Interval	2017	-	No specific time interval

Table 5.1: Full comparison between this project and the related work.

One of the main goals of this project is to avoid and overcome the weakness of the previous studies to solve the same type of terrorist group prediction problem.

On one hand paper [10] a larger scope for the study was followed since the region concerned with that study was the whole world. The weakness is that a very large time interval was used in the study which is indeed the whole GTD dataset till the time of the study (2012).

On the other hand no full details of training and testing were mentioned in the paper.

In paper [11] a period of about 4 years was used in that study, although it was the most recent to its date of publish (which is a positive point), a deduction can be made that it is not a sufficient interval to train and test the classifiers.

Also, using 10 Fold Cross Validation and random test set breaks a main feature in this kind of problems which is the time factor respection as it is not efficient to train the classifier by random tuples from 2013, 2010 and 2012 then test it with data from 2009, 2012 and 2013.

In this project, the weaknesses in the mentioned papers were overcome:

- Suitable time interval was selected to perform the study on [2000, 2017] where it is was divided into training [2000, 2016] and testing [2017].
- In paper [9] their goal was to predict the terrorist event in the upcoming years given the previous ones. So, this method was followed in this project but applied on a slightly different problem.
- New datasets with new features were introduced to the problem of terrorist group prediction.
- Similar to paper [11], the terrorist events in a more tight region (Middle East and North Africa) was studied.

In the following sections results obtained from the combinations of features from Global Terrorism Database and World Happiness Report (GTD + Happiness) are presented.

5.3 Results

Every classifier and its results will be discussed individually.

5.3.1 k Nearest Neighbor

kNN		
	Micro Average F Score / Accuracy	Macro Average F Score
GTD (Baseline)	87.00%	67.00%
GTD + Happiness	96.00%	69.00%
GTD + Ethnic Groups	91.00%	63.00%
GTD + Happiness + Ethnic Groups	96.00%	69.00%

Table 5.2: kNN Micro Average F-score/Accuracy and Macro Average F-score for all features combinations.

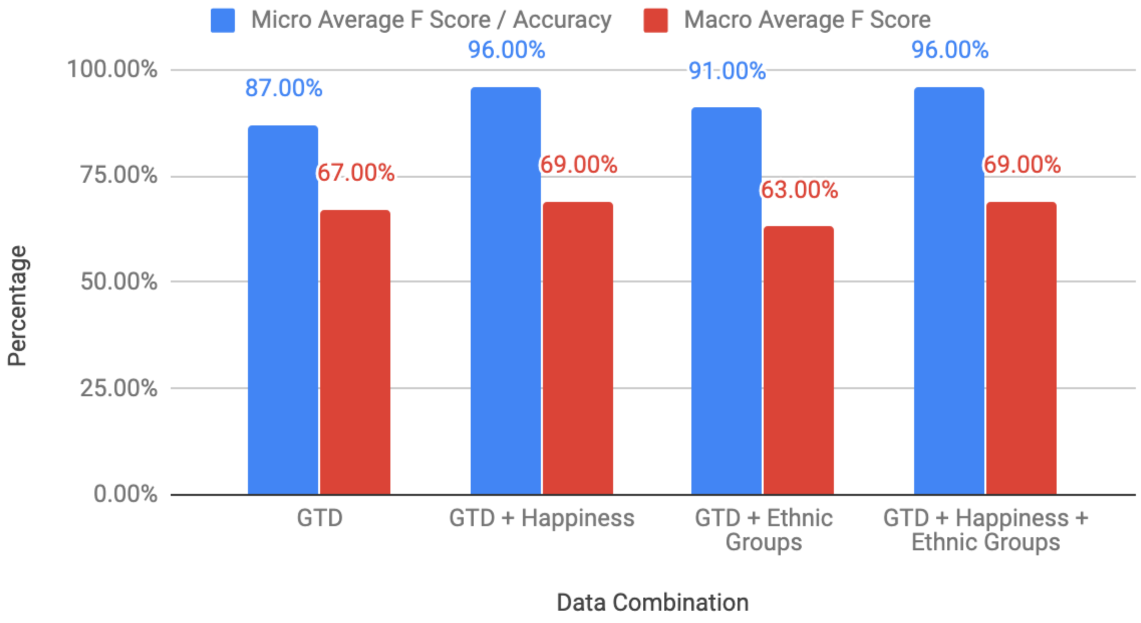


Figure 5.1: kNN results bar chart with percentage on the y-axis and datasets combinations on the x-axis.

5.3.2 Naive Bayes

Naïve Bayes		
	Micro Average F Score / Accuracy	Macro Average F Score
GTD (Baseline)	34.00%	40.00%
GTD + Happiness	37.00%	42.00%
GTD + Ethnic Groups	33.00%	39.00%
GTD + Happiness + Ethnic Groups	37.00%	42.00%

Table 5.3: Naive Bayes Micro Average F-score/Accuracy and Macro Average F-score for all features combinations.

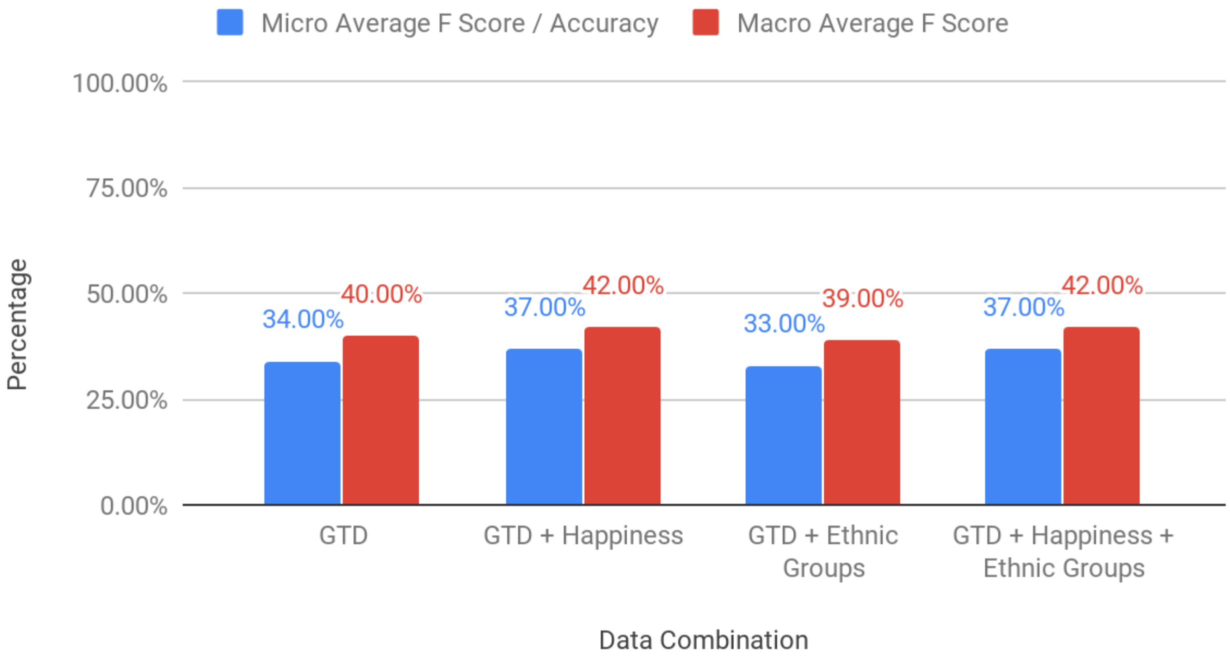


Figure 5.2: Naive Bayes results bar chart with percentage on the y-axis and datasets combinations on the x-axis.

Naive Bayes classifier makes a very strong assumption on the independence of features given the output class, due to this, the result can be very bad.

5.3.3 Random Forest

Random Forest		
	Micro Average F Score / Accuracy	Macro Average F Score
GTD (Baseline)	95.00%	72.00%
GTD + Happiness	94.00%	72.00%
GTD + Ethnic Groups	94.00%	71.00%
GTD + Happiness + Ethnic Groups	94.00%	72.00%

Table 5.4: Random Forest Micro Average F-score/Accuracy and Macro Average F-score for all features combinations.

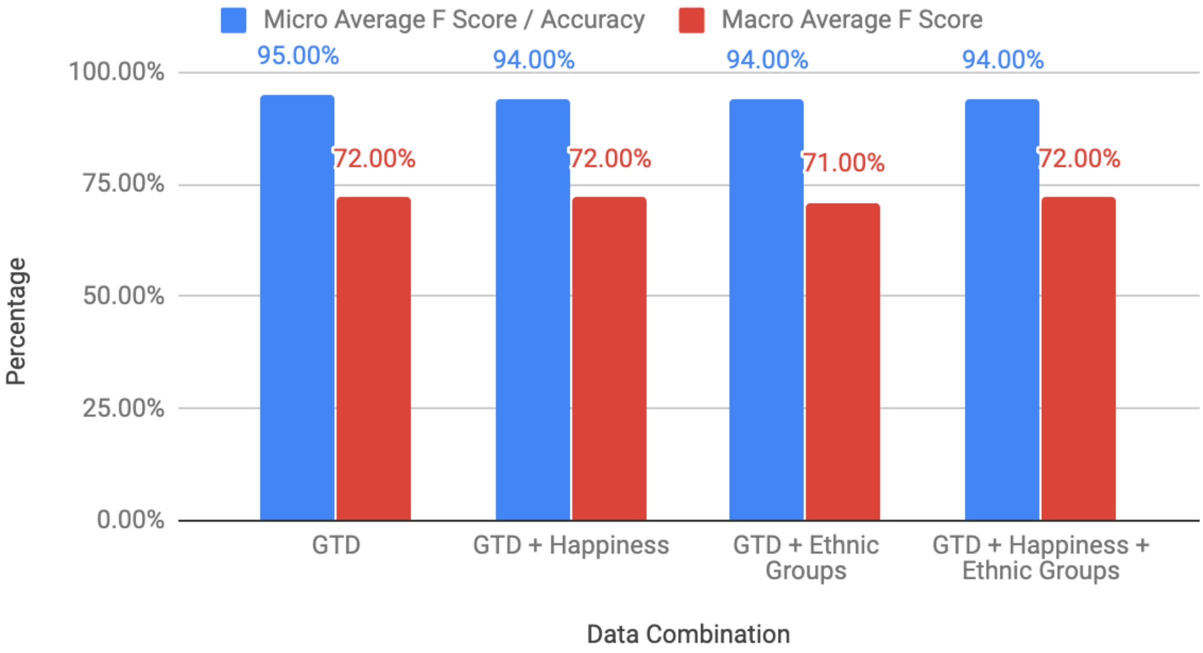


Figure 5.3: Random Forest results with percentage on the y-axis and datasets combinations on the x-axis.

5.3.4 Decision Tree

Decision Tree		
	Micro Average F Score / Accuracy	Macro Average F Score
GTD (Baseline)	91.00%	68.00%
GTD + Happiness	90.00%	69.00%
GTD + Ethnic Groups	91.00%	70.00%
GTD + Happiness + Ethnic Groups	90.00%	69.00%

Table 5.5: Decision Tree Micro Average F-score/Accuracy and Macro Average F-score for all features combinations.

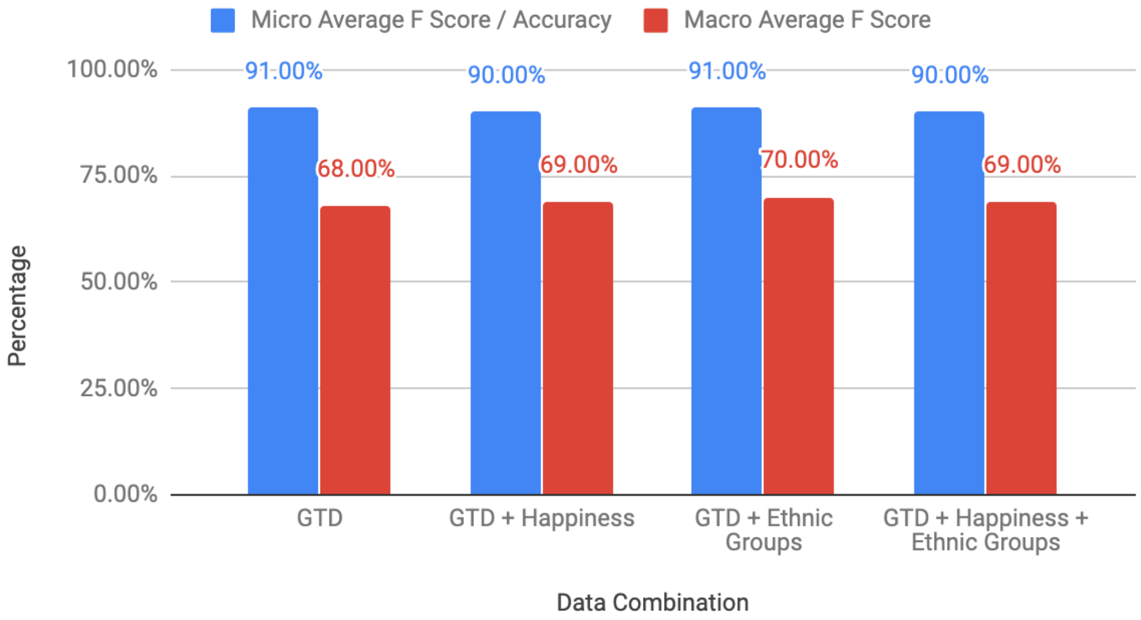


Figure 5.4: Decision Tree results bar chart with percentage on the y-axis and datasets combinations on the x-axis.

5.3.5 Support Vector Machine

SVM		
	Micro Average F Score / Accuracy	Macro Average F Score
GTD (Baseline)	93.00%	77.00%
GTD + Happiness	95.00%	81.00%
GTD + Ethnic Groups	93.00%	77.00%
GTD + Happiness + Ethnic Groups	95.00%	81.00%

Table 5.6: SVM Micro Average F-score/Accuracy and Macro Average F-score for all features combinations.

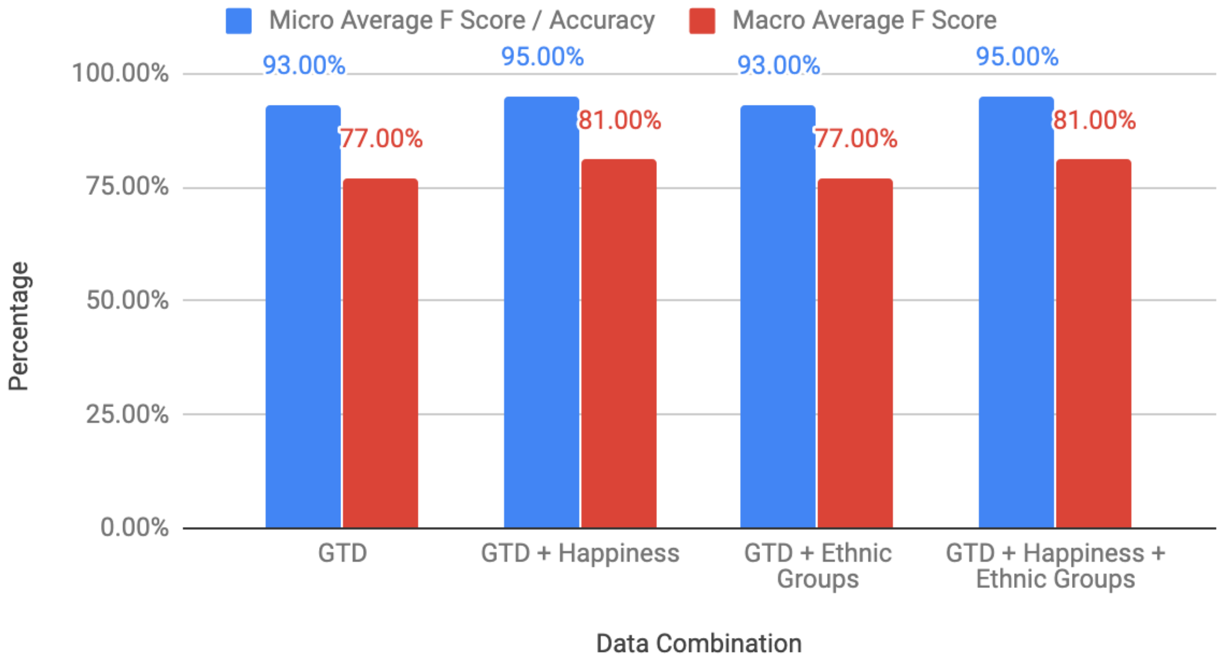


Figure 5.5: SVM results bar chart with percentage on the y-axis and datasets combinations on the x-axis.

5.4 Results Analysis and Comments

It is quite noticeable that most of the classifiers got high scores for Micro and Macro Average F-Scores.

The positive effect of using World Happiness Report features (freedom, happiness and economic effect) on the baseline (GTD features only) is noticeable, since the scores have increased by 2.61% on average after freedom score, happiness score and economic state are added to the study. Although it is a slight increase, it opens the door towards more contributions in the same terrorist group prediction problem through including similar metrics, which will be discussed later in Chapter 6.

While features from World Happiness Report have an effect on the baseline results, the effect of ethnic diversity was not noticeable, which is a logical result to obtain since ethnic races are almost similar in the Middle East and North Africa Region with no such diversity like the United States or India for example.

This can be shown in the following table:

	kNN		**NB**		**RF**		**DT**		**SVM**	
	Accuracy	F Score	Accuracy	F Score	Accuracy	F Score	Accuracy	F Score	Accuracy	F Score
(GTD + Happiness) - (GTD)	9.28%	2.00%	2.61%	2.00%	-0.30%	0.00%	-0.30%	1.00%	1.76%	4.00%
(GTD + Ethnic Groups) - (GTD)	4.31%	-4.00%	-1.03%	-1.00%	-0.36%	-1.00%	0.36%	2.00%	0.00%	0.00%
(GTD + Happiness + Ethnic Groups) - (GTD)	9.65%	2.00%	2.61%	2.00%	-0.36%	0.00%	-0.55%	1.00%	1.76%	4.00%
	Green labeled cells are of highest values in their columns									

Table 5.7: Effects of the new features introduced in this study on the baseline.

It is easily noticeable that in 4 out of 5 classifiers, World Happiness Report features are more involved in the positive change in results.

5.4.1 Accuracy Comparison with Previous Work

	kNN	NB	RF	DT	SVM
This Project (GTD + Happiness)	96.11%	37.07%	94.29%	90.23%	95.02%
Paper**[10]**		92.75%	83.43%	84.97%	

Table 5.8: Accuracy comparison between this project and paper [10].

For paper [10], only Naive Bayes, Random Forest and Decision Trees are common with this project. It is significantly noticeable that in 2 out of 3 classifiers the accuracy results are better by around 10%.

	kNN	NB	RF	DT	SVM
This Project (GTD + Happiness)	96.11%	37.07%	94.29%	90.23%	95.02%
Paper**[11]**	54.45%	64.85%	51.98%	67.32%	56.93%

Table 5.9: Accuracy comparison between this project and paper [11].

In 4 out of 5 classifiers the results of this project are very significantly higher than the accuracies of paper [11] by on average 40%.

	kNN	NB	RF	DT	SVM
This Project (GTD + Happiness)	69.00%	42.00%	72.00%	69.00%	81.00%
Paper**[11]**	52.10%	58.50%	47.90%	64.50%	54.00%

5.4.2 F-Score Comparison with Previous Work

Table 5.10: F-score comparison between this project and paper [11].

In 4 out of 5 classifiers the results of this project are significantly higher than the results of paper [11].

5.5 Conclusion

In this chapter, the results of training the classifiers, result analysis and comments were discussed, followed by an in-depth comparison with the related work from previous studies which was presented at the end of the chapter.

In the next chapter, a full conclusion for the project is presented, side by side with suggested future work which is expected to positively affect the results.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In chapter 1, the project's main idea was illustrated which is a predictor which aims to identify the terrorist group based on some traits of the country of the attack and of the attack itself.

In chapter 2, the need to extend the previous work was presented in section 2.6. A need to study the effects of economy, ethnic diversity and happiness on the nature of terrorist attack was presented and conducted with results in chapter 5.

Also, the need to hybridize the work of papers [8] and [9] and this project with the work of papers [10] and [11] are left as future work which is discussed in this chapter.

From the results presented in chapter 5, it is obvious that taking happiness, economics & freedom scores into consideration positively affects the performance of the classifiers.

Also combining GTD and World Happiness Report features together with using the suitable time interval (21st Century data) and testing on recent years positively affects the performance of the classifiers.

The work conducted by this project helps the authorities to tighten the bound on expecting the suspect terrorist group through relying on more reliable work.

Also, the work conducted by this project permits future work to rely on a reliable and well organized work for further extensions.

6.2 Future Work

In the previous section, the implemented features were stated. In this section the new features, classification techniques and extensions for the problem that could be added will be introduced.

6.2.1 Future Work Related to Classification Techniques

Deep learning can be introduced to terrorist group prediction problem.

6.2.2 Future Work Related to the Terrorism Issue

The problem of predicting terrorist groups responsible for a terrorist event like papers [10], [11] and this project could be hybridized with the problem of predicting terrorist events locations [8] and [9]. This hybridization will produce a predictor able to predict both potential terrorist group and potential event location.

6.2.3 Future Work Related to Sources of Data

Twitter can be used as a source for big data, since tweets and reactions during massive terrorist attacks can be collected and analysed.

6.2.4 Future Work Related to Features Used

Since the happiness, freedom, economic impact and ethnic diversity had had a significant effect on the results of the predictor. Similar factors could be added to studies dealing with similar type of problems. (ex Democracy Index, Education Statistics … etc).

Bibliography

[1] David Dietrich, Barry Heller & Beibei Yang (EMC Education Services). (2015). Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. (3rd Edition). [Online]. Available: http://index-of.co.uk/Big-Data-Technologies/Data%20Science%20and%20Big%20Data%20Analytics.pdf [July 3, 2019].

[2] Jiawei Han, Micheline Kamber and Jian Pei. (2011). Data Mining Concepts and Techniques. (3rd Edition). [Online]. Available: http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf [June 27, 2019].

[3] “k-nearest neighbors algorithm.” Internet: https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm, [June 27, 2019].

[4] A. Navlani. “Decision Tree Classification in Python.” Internet: https://www.datacamp.com/community/tutorials/decision-tree-classification-python, Dec. 28, 2018 [June 27, 2019].

[5] “Random Forest Classifier – Machine Learning” Internet: https://www.globalsoftwaresupport.com/random-forest-classifier-bagging-machine-learning/, Feb. 23, 2018 [June 27, 2019].

[6] Peter Kellman and Michael S Hansen. (2014, January). “T1-mapping in the heart: accuracy and precision.” Journal of Cardiovascular Magnetic Resonance. [Online]. 16(2), pp. 6. Available: https://jcmr-online.biomedcentral.com/articles/10.1186/1532-429X-16-2 [July 26, 2019].

[7] “Why are precision, recall and F1 score equal when using micro averaging in a multi-class problem?” https://simonhessner.de/why-are-precision-recall-

and-f1-score-equal-when-using-micro-averaging-in-a-multi-class-problem/, July 19, 2018 [June 27, 2019].

[8] Ding F, Ge Q, Jiang D, Fu J, Hao M (2017) Understanding the dynamics of terrorism events with multiple-discipline datasets and machine learning approach. PLoS ONE 12(6): e0179057. https://doi.org/10.1371/journal.pone.0179057

[9] Xun, Zhang & Jin, Min & Fu, Jingying & Hao, Mengmeng & Yu, Chongchong & Xie, Xiaolan. (2018). On the Risk Assessment of Terrorist Attacks Coupled with Multi-Source Factors. ISPRS International Journal of Geo-Information. 7. 354. 10.3390/ijgi7090354

[10] Gohar, Faryal & Haider, Wasi & Qamar, Usman & Publications, SDIWC. (2014). Terrorist Group Prediction Using Data Classification.

[11] Tolan, Ghada & H. M. Abou-El-Enien, Tarek & M. H. Khorshid, Motaz. (2015). HYBRID CLASSIFICATION ALGORITHMS FOR TERRORISM PREDICTION in Middle East and North Africa. International Journal of Emerging Trends & Technology in Computer Science (IJETTCS). 4. 23-29.

[12] National Consortium for the Study of Terrorism and Responses to Terrorism (START), University of Maryland. “Global Terrorism Database (GTD)” Internet: https://www.kaggle.com/START-UMD/gtd/home. [April 2019]

[13] Center for Comparative and International Studies (CIS), International Conflict Research, ETH Zurich. GeoEPR, Internet: https://icr.ethz.ch/data/epr/#ed, December 11, 2018 [April 2019].

[14] Sustainable Development Solutions Network. “World Happiness Report” Internet: https://www.kaggle.com/unsdsn/world-happiness. [April 2019]

[15] “Welcome to Colaboratory!” Internet: https://colab.research.google.com/notebooks/welcome.ipynb [June 26, 2019].

[16] “Scikit-learn Machine Learning in Python.” Internet: http://scikit-learn.github.io/stable [June 26, 2019].

[17] “Python Data Analysis Library.” Internet: https://pandas.pydata.org [June 26, 2019].

[18] NumPy Developers. “NumPy.” Internet: https://www.numpy.org [June 26, 2019].