BEHAVIOR ANALYSIS BASED PROTEST EVENT DETECTION

DAVRANIŞ ANALİZİ TABANLI PEOTESTO OLAYLARI TESPİTİ

MAHMOUD HOSSEIN ZADEH

PROF. DR. İLYAS ÇİÇEKLİ
Supervisor

Submitted to
Graduate School of Science and Engineering of Hacettepe University
as a Partial Fulfillment to the Requirements
for the Award of the Degree of Doctor of Philosophy
in Computer Engineering

June 2023

ABSTRACT

BEHAVIOR ANALYSIS BASED PROTEST EVENT DETECTION

Mahmoud HOSSEIN ZADEH

Doctor of Philosophy, Department of Computer Engineering

Supervisor: Prof. Dr. İlyas ÇİÇEKLİ

June 2023, 115 pages

Purpose: Social unrest is a phenomenon that occurs in all countries, both developed and poor. The only difference is in the cause of a social unrest and it is mostly economic in underdeveloped countries. The occurrence of protest and the role of social networks in it have always been debatable topics among researchers. Protest Event Analysis is important for government officials and social scientists. Here we present a new method for predicting protest events and identifying indicators of protests and violence by monitoring the content generated on Twitter.

Methods: By identifying these indicators, protests and the possibility of violence can be predicted and controlled more accurately. Twitter user behaviors such as opinion share and event log share are used as indicators and this study presents a new method based on Bayesian logistic regression algorithm for predicting protests and violence using Twitter user behaviors. According to the proposed method, users' event log share behaviors which include the rate of tweets containing date and time information is the reliable indicator for identifying protests. Users' opinion share behaviors which include hate-anger tweet rates is also best for identifying violence in protests.

Results: A research database consists of tweets generated on the BLM (Black Lives Matter) movement after the death of George Floyd. According to information published on acleddata.com, protests and violence have been reported in various cities on specific dates. The dataset contains 1414 protest events and 3078 non-protest events from 460 cities in 37 U.S. states. Protest events include 1414 protests in the BLM movement between May 28 and June 30 among which 285 were violent and 1129 were peaceful. We tested our proposed method on this dataset and the occurrence of protests is predicted with 85% precision. It is also possible to predict violence in protests with 85% precision with our method on this dataset.

Conclusion: According to the research findings, the behavior of users on the Twitter social network is a reliable source for predicting incidents and violence. This study provides a successful method to predict small and large-scale protests, different from the existing literature focusing on large-scale protests.

Keywords: Protest Detection, Event Detection, Social Behavior, Social Media, Bayesian Logistic Regression, Machine Learning

ÖZET

DAVRANIŞ ANALİZİ TABANLI PEOTESTO OLAYLARI TESPİTİ

Mahmoud HOSSEIN ZADEH

Doktora, Bilgisayar Mühendisliği

Tez Danışmanı: Prof. Dr. İlyas ÇİÇEKLİ

Haziran 2023, 115 sayfa

Amaç: Protesto olayları, hem gelişmiş hem de fakir tüm ülkelerde meydana gelen bir olgudur. Tek fark, bu toplumsal huzursuzluğun nedenlerindedir. Az gelişmiş ülkelerde protestoların nedenleri ekonomiktir. Protestoların ortaya çıkışı ve bunda sosyal medyanın rolü, araştırmacılar arasında her zaman tartışmalı bir konu olmuştur. Protesto Olayı Analizi, hükümet yetkilileri ve sosyal bilimciler için önemlidir. Bu çalışmada, Twitter'da üretilen içeriği izleyerek protesto olaylarının gerçekleşeceğini tahmin etmek ve protesto ve şiddet göstergelerini belirlemek için yeni bir yöntem sunulmuştur.

Yöntemler: Bu göstergeler belirlenerek, protestolar ve şiddet olasılığı daha doğru bir şekilde tahmin ve kontrol edilebilir. Fikir Paylaşımı ve Tarih ve Zaman davranışları gibi Twitter kullanıcı davranışları gösterge olarak kullanılmıştır ve bu çalışma, Twitter kullanıcı davranışlarını kullanarak protestoları ve şiddeti tahmin etmek için Bayes lojistik regresyon algoritmasına dayalı yeni bir yöntem sunmaktadır. Önerilen yönteme göre, kullanıcıların tarih ve zaman bilgilerini içeren tweet'lerin oranını içeren Tarih ve Zaman davranışları, protestoları belirlemede güvenilir bir göstergedir. Kullanıcıların nefret-öfke tweet oranlarını içeren Fikir Paylaşımı davranışları da protestolardaki şiddeti tespit etmek için başarılıdır.

Bulgular: Bir araştırma veri tabanı, George Floyd'un ölümünden sonra BLM (Black Lives Matter) hareketi üzerine oluşturulan tweet'lerden oluşuyor. Acleddata.com'da yayınlanan bilgilere göre, çeşitli şehirlerde belirli tarihlerde protestolar ve şiddet olayları yaşandığı bildirilmiştir. Veri seti, 37 ABD eyaletindeki 460 şehirden 1414 protesto olayı ve 3078 olaysız günü içeriyor. Protesto olayları, BLM hareketinde 28 Mayıs ile 30 Haziran arasında 285'i şiddetli ve 1129'u barışçıl olmak üzere 1414 protestoyu içermektedir. Önerdiğimiz yöntemi bu veri seti üzerinde test edilmiştir ve protestoların oluşumu %85 doğrulukla tahmin edilmiştir. Yine bu veri setindeki yöntemimizle, protestolardaki şiddeti %85 doğrulukla tahmin etmek mümkündür.

Sonuç: Araştırma bulgularına göre, Twitter kullanıcıların davranışları, olayları ve şiddeti tahmin etmek için güvenilir bir kaynaktır. Bu çalışma, büyük ölçekli protestolara odaklanan mevcut literatürden farklı olarak, küçük ve büyük ölçekli protestoları tahmin etmek için başarılı bir yöntem sunmaktadır.

Anahtar Kelimeler: Protesto Belirleme, Olay Belirleme, Sosyal Davranış, Sosyal Medya, Bayes Lojistik Regresyon, Makine Öğrenimi

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Dr. İlyas ÇİÇEKLİ and my co-supervisor Asst. Prof. Dr. Gönenç ERCAN for their continuous support of my Ph.D. study and related research, patience, encouragement, and immense knowledge. They have always been there for me to provide encouragement and helpful feedback on my research.

It is necessary to give special thanks to Prof. Dr. İlyas ÇİÇEKLİ. His sincere and supportive behavior was very impressive.

Besides my supervisors, I would like to thank the rest of my thesis committee members, Prof. Dr. Ferda Nur ALPASLAN and Asst. Prof. Dr. Burcu CAN, for their insightful comments and motivation, but also for the tough question, which incented me to widen my research from different views.

Last, I would like to thank my family, especially my most beloved wife Roya and my son Ali, who has given me the greatest motivation to finish my dissertation. During my dissertation, her positive energy was constructive.

CONTENTS

Page

ABSTRACT i

ÖZET iii

ACKNOWLEDGEMENTS v

CONTENTS vi

TABLES ix

FIGURES x

ABBREVIATIONS xi

1. INTRODUCTION 1
   1.1. Purpose of the Thesis 1
   1.2. Contributions 2
   1.3. Outline of the Thesis 3

2. BACKGROUND INFORMATION 5
   2.1. Overview 5
   2.2. Behavior Analysis 5
      2.2.1. Definition of Behavior 5
      2.2.2. Types of Behaviors in Social Media 7
         2.2.2.1. Individuals Behaviors 7
         2.2.2.2. Collective Behaviors 8
      2.2.3. User Behaviors in Twitter 9
         2.2.3.1. General Information Sharing 10
         2.2.3.2. Opinion Sharing 11
         2.2.3.3. Call for action 11
         2.2.3.4. Event log sharing 12
   2.3. Social Media 13
      2.3.1. Definition of Social Media 13

2.3.2. Twitter 15
2.3.2.1. About Twitter 15
2.3.2.2. The Importance of Twitter in Social Media Research 16
2.3.3. Social Media Mining 17
2.3.3.1. Definition 17
2.3.3.2. Applications 17
2.3.3.2.1. Psychology 18
2.3.3.2.2. Business 18
2.3.3.2.3. Political 18
2.3.3.2.4. Sport 19
2.3.3.2.5. User Profile 19
2.4. Social Media for Events 20
2.4.1. Event 20
2.4.2. Types of Event 22
2.4.2.1. Scientific Event 22
2.4.2.2. Personal Event 22
2.4.2.3. Trade Event 22
2.4.2.4. Crisis Event 23
2.4.2.4.1. Disaster Event 23
2.4.2.4.2. Protest Event 23
2.4.3. Event detection from Twitter 24
2.4.3.1. Detection type in Event Detection 25
2.4.3.2. Detection method in Event Detection 26
3. SOCIAL MEDIA FOR EVENT PROTEST 28
3.1. Protests 28
3.2. Social Media Role in Protests 28
3.3. General Architectural View of Protest Event Detection 30
3.3.1. Data preprocessing 30
3.3.2. Feature Extraction 32
3.3.2.1. Structural 32
3.3.2.2. Temporal 33
3.3.2.3. Spatial 33
3.3.2.4. Semantic 33

3.3.3. Detection models 33
3.4. Related Works on Social Media in Protests 36
3.5. Protest Event Detection Applications 43
4. PROTEST EVENT ANALYSIS BASED ON TWITTER'S USER BEHAVIORS 47
4.1. Overview 47
4.2. Datasets 47
4.2.1. Occupy Wall Street Dataset 48
4.2.2. Emotion Analysis Dataset 49
4.2.3. BLM Dataset 49
4.3. User Behaviors and Protests 57
4.3.1. Relationship between Twitter user behaviors and protest day 57
4.3.2. Recognition of EventLogShare 59
4.3.3. Emotion Analysis on Tweets 61
4.4. The Proposed Prediction Method Based on User Behaviors 64
5. EXPERIMENTAL RESULTS 69
5.1. Protest events 72
5.2. Protest types 77
6. DISCUSSION, CONCLUSION AND THOUGHTS FOR THE FUTURE 83
6.1. Online Protest Event Detection 84
6.2. The Open Dataset Problem 85
6.3. Small and large-scale protests 85
6.4. Limitations of study 85
6.5. Conclusion and Thoughts for the Future 86
REFERENCE 89
APPENDICES 104
A - Publication Derived From Thesis 104

TABLES

Page

Table 2.1 Social media types and examples of each type 14

Table 2.2 Statistical information of SNNs users 15

Table 3.1 Preprocessing methods & tools 32

Table 3.2 Comparison of PED methods in details 35

Table 3.3 Protest detection features, methods and performances.................................. 42

Table 4.1 Datasets of study and their role 48

Table 4.2 Emotion dataset specifications 49

Table 4.3 The information is available in the acleddata.com dataset............................ 52

Table 4.4 The final dataset 53

Table 4.5 The states with the most frequent protests 54

Table 4.6 The cities with the most frequent protests... 56

Table 4.7 The relationship between the day of the protest and the behavior of twitter users.. 58

Table 4.8 The accuracy of behavior detection of eventlogshare 60

Table 4.9 The detection accuracy of eventlogshare behavior after adding custom sutime rules 61

Table 4.10 The relationship between the day of the protest and the behavior of twitter users.. 61

Table 4.11 Sentiments and emotions.. 62

Table 5.1 Variables of confusion matrix 69

Table 5.2 Protest event classification accuracy, precision, recall, and F1 score 73

Table 5.3 A more detailed table of the best features combination results in a protest event 74

Table 5.4 Protest types classification accuracy, precision, recall, and F1 score 79

Table 5.5 A more detailed table of the best features combination results in protests type 80

FIGURES

Page

Figure 2.1 User behaviors in twitter 9
Figure 2.2 The location of event detection in information retrieval. 21
Figure 2.3 The overall process of event detection.. 24
Figure 2.4 The subgroups of event 25
Figure 3.1 News related to the shutdown of internet in protests 29
Figure 3.2 General architectural view of PED 31
Figure 4.1 Stages of converting tweet ids into csv file with the given specifications..... 50
Figure 4.2 The number of tweets by day 51
Figure 4.3 The number of protests and protest type in the dataset 54
Figure 4.4 Behavior detection steps of eventlogshare 59
Figure 4.5 Overview of the presented method 65
Figure 5.1 Protest events classification accuracy, precision, recall, and F1 score 74
Figure 5.2 The fitted sigmoid curve and the decision boundary for time info 75
Figure 5.3 The fitted sigmoid curve and the decision boundary for date info 76
Figure 5.4 The fitted sigmoid curve and the decision boundary for number of tweet 77
Figure 5.5 Violent events classification accuracy, precision, recall, and F1 score 80
Figure 5.6 The fitted sigmoid curve and the decision boundary for hate-anger tweets ratio……………..………………………………………………………......81

ABBREVIATIONS

PEA	: **P**rotest **E**vent **A**nalysis
PED	: **P**rotest **E**vent **D**etection
BLM	: **B**lack **L**ives **M**atter
BLR	: **B**ayesian **L**ogistic **R**egression
SM	: **S**ocial **M**edia
SMM	: **S**ocial **M**edia **M**ining
SNUBA	: **S**ocial **N**etwork **U**sers **B**ehavior **A**nalysis
SNNs	: **S**ocial **N**etworking **S**ites
SMO	: **S**equential **M**inimal **O**ptimization
OWS	: **O**ccupy **W**all **S**treet
$\mathbf{X^2}$	: **C**hi-square **I**ndex
Fo	: **F**requency **O**bserved
Fe	: **F**requency **E**xpected
TP	: **T**rue **P**ositive
FP	: **F**alse **P**ositive
FN	: **F**alse **N**egative
TN	: **T**rue **N**egative
GIS	: **G**eneral **I**nfo **S**hare
OS	: **O**pinion **S**hare
ELS	: **E**vent **L**og **S**haring
SA	: **S**entiment **A**nalysis

1. INTRODUCTION

1.1. Purpose of the Thesis

Protest Event Analysis is important for government officials and social scientists. Here we present a new method for predicting protest events and identifying indicators of protests and violence by monitoring the content generated on Twitter.

Analysis of events using social media data has been a topic of interest for researchers for many years. Various types of events, including personal, public, and protest events, have been analyzed in previous studies, and different results have been reported. In this thesis, protest events are studied as part of social unrest.

Protests as a form of social unrest occur vastly in different countries. Contrary to popular belief, protests are not limited to less developed countries and are not solely driven by economic reasons. Examples such as the Yellow Vest protests in France and the Black Lives Matter (BLM) movement in the United States demonstrate that protests have diverse underlying causes and can occur even in developed countries. One of the primary objectives of this study is to investigate the effectiveness of social media networks in real-world events. There are various conflicting opinions about the impact of social media networks and the extent of their influence. Another goal of this thesis is to analyze and measure the effectiveness of social media networks in predicting protest events.

Previous studies on PED (Protest Event Detection) have been based on the classical classification approach. Related works in predicting protests have shown that these studies mainly focus on increasing prediction accuracy. However, identifying and predicting protest events requires indicators that can be used for analyzing and forecasting such occurrences. This thesis identifies these indicators for the BLM protests in the United States. Based on the results obtained, the rate of DateInfo and TimeInfo tweets relative to all tweets is critical in predicting important protest events.

Violence during protests imposes significant financial and human costs on countries and individuals. Protests can be categorized into peaceful and violent events. Peaceful events have lower news value and do not entail financial and human costs for countries. Predicting violent events can be valuable for governments and help prevent very high costs in protests. One of the primary objectives of this thesis is to provide a methodology for identifying violence in the BLM movement in the United States. It was discovered through analyzing tweets published after George Floyd's death that monitoring the rate of hate-anger tweets can predict violent incidents.

Events of protests can be divided into small-scale and large-scale events. Small-scale events occur in small cities with a low population, while large-scale events happen in more populous cities with the participation of hundreds and thousands of people. One of the most significant shortcomings in related works is the lack of a method for simultaneously predicting small and large-scale protest events. Given the extent of the BLM movement in the United States, this thesis presents a successful method for predicting protest events in both small and large-scale events. For the first time, an accurate and acceptable method for simultaneous prediction of protest events on both small and large-scale has been proposed.

This thesis aims to find a practical and accurate solution for analyzing social unrest in small and large-scale events. Based on a comprehensive review of related works, insufficient attention to protest indicators was the primary motivation for this study. In this thesis, violence was given special attention as a damaging component in some protests, and the proposed method was also successful in predicting violence in protests. According to the study's findings, protests and violence can be predicted with acceptable accuracy using social network data such as Twitter.

1.2. Contributions

The contributions of the thesis are as follows:

- Forecasting protest events using Twitter users' behaviors. A new method for predicting protests is presented using the Bayesian logistic regression approach and user behaviors.
- Providing a Bayesian logistic regression algorithm to predict protests and identify early indicators to monitor social networks.
- Providing a way to select a feature to predict protests based on Twitter users' behaviors. Twitter users' behaviors (EventLogShare, InformationShare, and OpinionShare) are presented as features for predicting protests and protest type.

- Predicting violent protests using Twitter user's behaviors. In most work, protests are predicted, but the probability of violent protests is not investigated.
- Identifying the most important early indicators of violence in social protests. The most worrying feature of the protests is the probability of violence. This dissertation identifies the most important indicators that specify the probability of violent protests.

1.3. Outline of the Thesis

This dissertation comprises several chapters that discuss various aspects related to PED, behavior definition, and social media, followed by a detailed analysis of social media usage during these events. The following three chapters focus on the main works related to proposing a new method for estimating protests and violence. The organization of the thesis is as follows:

In Chapter 2, the discussion begins with an overview of behavior and its various definitions. Different behaviors on social networks, particularly on Twitter, are then explained in detail. Collective and individual behaviors performed by users on Twitter are elaborated upon in this chapter. Social networks provide a platform used by protesters during events of protests. Twitter has been used as a data source in research due to the publication of user tweets. This chapter discusses social networks and social media mining in detail due to their relevance to the thesis concepts. The concept of events, types of events, and event detection on Twitter are explained in the final section of chapter two.

In Chapter 3, protests are defined as a form of social unrest. The different components of a protest that can be the research subject are outlined. The role of social networks in events has been extensively evaluated in previous studies, and a summary of these evaluations is provided in this chapter. The stages of the proposed methods for predicting protests, including pre-processing, feature identification, and prediction, are explained in this chapter. More than thirty relevant articles are described in various tables in this chapter. By studying this section, readers will become familiar with various works on predicting protests.

Chapter 4 is dedicated to presenting the proposed method based on user behavior. Collecting the dataset for this thesis required a significant amount of time and energy. One of the main challenges in PED research was the lack of comprehensive and publicly available datasets. This issue led to the collection of unique datasets for analysis in all studies in this field. The datasets used in this chapter are described in detail. The presented method includes the identification of spatial and temporal information, as well as an explanation of sentiment analysis as a part of this thesis. The framework of the proposed method and detailed explanations of each section of the method are provided in this chapter.

In Chapter 5, the results obtained from the implementation of the proposed method on the dataset are presented. The results related to predicting protests and violence are presented in separate sections. In Chapter Six, a comparison of the obtained results with related research is presented. The strengths and limitations of the proposed method are explained in detail in the discussions presented. This thesis concludes with the Conclusion and Future Work.

2. BACKGROUND INFORMATION

2.1. Overview

In this chapter, a comprehensive definition of the word “behavior” is provided first. Because the word "behavior" is used in different fields, different meanings are derived from it. What has been investigated in this thesis and based on which a model for identifying protests has been presented is the behavior of users in social networks. Types of user behaviors, including individual and collective behaviors, are explained in this chapter. The four behaviors of users on Twitter are studied in this part of the thesis and explained with examples.

Many researchers have studied social networks due to their great effects on human behavior. In the rest of this chapter, helpful information is provided about social networks and Twitter.

Social media mining is defined, and its applications in psychology, business, politics, sports, and other fields are also explained.

In the final part of this chapter, full explanations about the use of social networks in events are mentioned, and the analysis of different events with social network data in research is described.

2.2. Behavior Analysis

The word "behavior" has many different meanings because it is used in many fields of study. It is necessary to provide a precise definition of this word. In the following section, individual and collective behaviors are described. The types of behavior of users on Twitter are classified and explained in the rest of this section.

2.2.1. Definition of Behavior

The concept of behavior in broad communities extends from social to online, cultural, business, economic, and mobile domains. However, deep, complete, and effective capturing, quantifying, representing, analyzing, measuring, and learning semantics, evolution, networking, sequencing, utility, the effect of collective, individual, and cohort behaviors in the real world cannot be done using comprehensive and systematic theories, systems, and tools [1]. The highly important concept of behavior is seen in the societal, cultural, political, virtual, living, military, economic, and scientific world.

“Behavior” literally means acting or behaving and how any material acts or reacts under specified circumstances [2]. “Behavior” in Wikipedia means the mannerisms and actions made by artificial entities, systems, or organisms together with its environment, including the other organisms or systems around, and also the physical environment. Behavior shows how the system or organism responds to different inputs or stimuli, whether conscious or subconscious, overt or covert, voluntary or involuntary, and internal or external. Therefore, the behavior is very social and ubiquitous. Besides the common terms, including human behaviors, animal behaviors, organizational behaviors, and consumer behaviors, behaviors are ubiquitous and emerge at any time in every place. In the physical world, behaviors are explicit and have been investigated from various perspectives. One social and digitalized life was rapidly developed and deeply engaged due to advanced computing technology, particularly online games, mobile applications, social networks, virtual reality, multimedia information processing, pattern recognition, visualization, machine learning, and social media; there was the emergence of more behaviors in the social and virtual world. In addition, their engagement with the social and virtual world led to the complexity of behaviors in traditional spheres. These areas are dominated by social behaviors. In more classic areas such as living spaces, business, politics, and economics, behaviors become increasingly social. Behaviors in different scenarios and applications are representative of relevant social and non-social relationships, structures, characteristics, and effects. For example, a trader’s behavior in stock markets affects others, as manifested through action properties, trading actions such as placing an order at a certain price, time, and volume for target security. A rich or concrete object, i.e., an individual or a group's behavior, is formed by the actions, presentation, response, and the effect related to the corresponding properties forms [1].

Nowadays, the behavior is a widely applied concept from social sciences to trade, mobile, economics, and culture. In this research, the concept of behavior is limited to the users' behavior in social media.

The growing expansion of social media led to the increased sharing of information [3]. Social media allow users to share whatever they want. Thus, it is said that social media caused to democratize societies because people in society can express their views at a lower price [4]. As a result, users in social media can show different behaviors [5].

For example, sharing, posting, liking, and commenting are examples of social media users' behaviors. Studying all types of social media behaviors, we can classify these behaviors into two individual and collective groups. The individual behaviors include a single user's behavior, and collective behavior includes a group of users. For example, posting a photo of a birthday party or sharing a city's climate is individual behavior. Immigrating from social media to social media or trying to trend a hashtag by a group of users can be regarded as collective behavior.

The user's activities in mass media produce behavioral data like Twitter. These broad and large data indicate users' tendencies, beliefs, views, and relations. These behavioral data allow monitoring the individual and collective behaviors of users by analyzing and observing them. In this section, two types of users' behaviors are discussed in more detail.

2.2.2. Types of Behaviors in Social Media

2.2.2.1. Individuals Behaviors

A user's individual behaviors can include one of the following behaviors [6].

1. User-user behavior

This type of behavior occurs between two users. For example, following a user is a behavior of this type. In this type of behavior, two users react or act as two independent entities.

2. User-entity behavior

This type of behavior occurs between an entity and a user. An entity can be a post on Facebook. Liking a post on Facebook, posting a tweet on Twitter, and sharing an image on Instagram are user-entity behaviors.

3. User-community behavior

This type of behavior occurs between a user and the community. Membership in a LinkedIn group or leaving a Facebook group is of these behaviors. All of these behaviors can be investigated and analyzed by computational methods. Based on these analyses, the users' behavioral patterns can be obtained. For this purpose, the first step is to collect the dataset and then analyze the data and reach the user's behavioral pattern. For example, to analyze the user's behavior toward dating, the user dating pattern can be identified by collecting the user's friends' data. User's behavior pattern can be obtained by investigating characteristics of the user's friends and using the machine learning algorithms. This approach has been raised and used in several studies on data mining. In all of these approaches, the suitable dataset is first collected. Then, features of the dataset are studied. An algorithm such as a neural network, classification, or prediction model is presented based on the dataset features and machine learning. In this regard, a user's behavior patterns in dating can be achieved in mass media. Based on the assessment methods, the veracity of the model is assessed [6].

2.2.2.2. Collective Behaviors

Collective behavior can be done either individually or collectively. Individually, a person shows behaviors spontaneously and without any planning, leading to collective behavior. However, in collective behavior, a group shows collective behavior with planning and clear objectives.

Collective behaviors usually emerge as a reaction to severe accidents or disasters. These reactions emerge based on three concepts of excited emotion, adaptive responses, and response to social violence [6].

The collective behavior formed based on the excited emotion is the most spontaneous type of collective behavior and emerges based on the intense emotions, the definition of logical thinking, and the participants' critical ability. The collective behavior formed based on the adaptive response will emerge in response to the new and ambiguous successes. The third and the most important factor of collective behaviors is social pressure. In collective behavior, there may be clear attitude toward success and how to deal with it. In this study, which is conducted on the analysis of the users' behavior in social media, this definition of collective behavior, i.e., response to social pressure, is widely applied. One of the most important manifestations of collective behavior is the users' collective behavior in social protests. In this field, users use mass media to convey the voice of the protestors, coordinate for the location and time of protest, and make comments about the protests.

According to the available literature, six types of collective behaviors, i.e., communication processes, emerging undifferentiated groups, social interactions and influence, behaviors beyond the traditional culture, group emotions, and responses to problems using social systems. The first four types are applied to use social media to study collective behaviors [7].

Meaningful patterns in individual and social behavior are recognized by extracting and analyzing social media interactions. Social media data can be computed analytically.

Social media are the communication channels through which social effects are expanded. In this thesis, social protests are studied as collective behavior.

2.2.3. User Behaviors in Twitter

Today, social networks are the most important information transfer and sharing tools [8]. Users of these networks produce and share views, beliefs, and subjects in which they are interested. In this regard, the study field has been created in this area, called the Social Network Users Behavior Analysis (SNUBA).

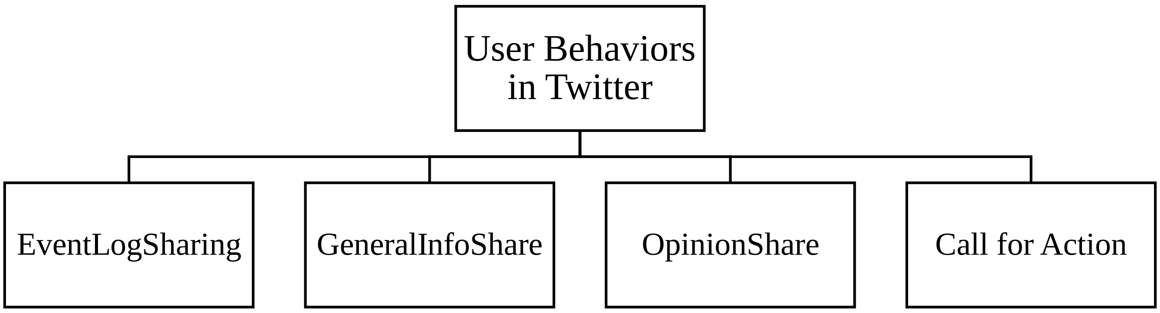


Figure 2.1 User behaviors in Twitter

SNUBA aims to analyze users activities such as sharing, posting, liking, commenting, following, etc. More activities depend on the field in which behavior analysis is done. For example, in a social protest on Twitter, Behavior Analysis means analyzing the user's Tweets content.

Users' behavior in social networks is divided into individual and social behaviors. Users usually create content in subjects such as the social protest of those interested in social behaviors.

With the increasing expansion of social networks in several countries, the protestors use social networks such as Twitter to organize and share their protests. One of the governments' concerns about social networks is to use these networks to organize protests all over the countries.

In this regard, users' behaviors are investigated in the social networks. In the research by Dr. Wang, the users' behaviors in the social networks have been ranked in events such as Occupy Wall Street (OWS) [9]. She investigates the social sciences and the subjects related to social networks, and she has classified users' behaviors in social networks in a global protest. She claims that the users' behaviors in social networks during the Wall Street protests are as follows:

1. Event log sharing
2. Sharing general information about the event
3. Sharing opinion about the event
4. Call for action

She believes that when an event such as a social protest occurs in social networks, users try to show one of the above-mentioned social behaviors. Figure 2.1 shows user behaviors in Twitter.

In the Wall Street protests, which are one of the most important social movements in the USA, the Twitter social network was intensively used by the protesters. The cyberspace users try to react to this event using Twitter. Dr. Wang's view about the users' behaviors in Wall Street protests can be extended to other countries' protests. This research field is a new one and needs further studies [9].

2.2.3.1. General Information Sharing

Another behavior expressed by the social network users, particularly on Twitter in their events and protests, is general information sharing. Many social network users do not have enough information about all events, but the main users should share this following the post. Thus, the main aim of this class of tweets is to share general information about the events. The users who have sympathy for the event and follow it seek to make other silent users follow. Thus, this class of tweets is created and shared by them. General information sharing in different fields attracts many social network users, and a large volume of the content created for the events is classified in this class. Some eyewitnesses create content in this class. Official organizations such as the police department or banks create these tweets or content by sharing information with people.

For example, an English tweet of this behavior is as follows:

“Security cameras in this area b/c big brother can newer enough eyes on us. Arrest during test run for #occupywallstreet”

The main goal of this class is to share general information about the event with the social network's general users.

2.2.3.2. Opinion Sharing

For each event, some users create opinions about that event. These opinions can be positive or supportive, while they can be negative and critical opinions about the event. The third case is the sharing of impartial views about the event.

Besides opinions about the event, the social network users comment on the participants in the event. For example, the comments about an event such as street protests include a positive comment on the Wall Street protest participants or negative or critical comments about the participants in Wall Street protests. The third type is the impartial comment about the event participants. Some English tweets of this behavior are given as follows:

@wenders1022 “I’ve been to PTA meeting of 10 where nothing can get done. Do the #occupywallstreet crowd really think they can strategize with bullhorns?”

“RT @HealthworksFit: #BlackLivesMatter. We stand in solidarity with the Black community near and far against systemic racism and injustice.…,https://t.co/cdedP2V4eP”

“Don’t let these mothafuckas fool you https://t.co/DoUW6SV9Tb”

This behavior in the social networks regarding events such as street protests briefly includes the users' opinions about the event. Opinions can be positive, negative, or impartial.

2.2.3.3. Call for action

Another behavior of social network users on Twitter is related to the tweets created by the users to attract celebrities' attention and traditional media such as radio and television, and other users. The call for action of the participants in social protest is another type of behavior in social networks. For example, call for action includes requesting to send food and blanket for the participants in the Wall Street protest in the following tweet:

@AnonyOps “Any bands want to bring equipment to wallstreet? We’re going to send food/pizza and some jams to protestors. #occupywallstreet”

As observed in this tweet, some users requested for blanket, and they claim that they are sending food and pizza to the protesters. This type of tweet is accompanied by the request of the user who created these tweets. In fact, the main goal of the user creating these tweets is to attract the attention of other users and help the main protestors.

The percentage of these tweets or content creation in social network events regarding protests is the lowest of the other three behaviors, i.e., Event log sharing, general information sharing, and opinion sharing .

2.2.3.4. Event log sharing

This Section relates to the event log sharing behavior of the users in a social protest. One of the users' main behaviors on Twitter in the street protests is that users try to share spatial, temporal, and logistic event log. For example, the protest will be held in front of Wall Street Park on Monday at 6 P.M. Through this content creation, users try to share the event log information. For example, they tweet protest hour and location to attract more people to the location. Some users who cannot attend the location at the due hour for different reasons try to inform others of the time, location, and other logistic information by sharing the information. One of the main behaviors that users show in the street protests is to share time and space event log. English tweets of this behavior are given as examples:

“Tommorow Wall Street will belong to the people, #oc cupywallstreet is this saturday #sep17.”

“*RT @angelxxelyse: A rally/protest is scheduled Friday (tomorrow) 5 p.m. at Peter's Park in the South End Boston. Please consider attending…”*

“RT @HeadOverFeels: On June 6th join the #DoctorWhoBlackout! We're supporting #BlackLivesMatter with this livetweet of THE GHOST MONUMENT a…, https://t.co/6zqPO3gIve”

“RT @MrAndyNgo: Antifa groups in Portland, Ore. have announced a 6 p.m. gathering at Laurelhurst Park. This is a middle class residential ne…”

2.3. Social Media

Social networks are the subject of extensive research and are presented in this social media section. Due to user content-sharing policies on Twitter, the research done on social media is mainly focused on Twitter. Social media mining and its applications are among the other topics discussed in this section.

2.3.1. Definition of Social Media

Social media are web-based applications that allow users to generate, share content, and spend time [10]. They are divided into thirteen different categories depending on the application, and the type of content produced [11]. An overview of social networking sites (SNNs) and examples of each type is presented in Table 2.1.

Microblogs have been able to attract the attention of many users, organizations, and researchers. The reason for this popularity is the provision of communication services with the feature of portability, ease of use, and immediacy, which allow users to respond or disseminate information instantly [12]. Witnesses to an event or people directly related to that event can publish timely information through microblogs.

There are dozens of sites that only some of SNNs could appear in people's daily lives. Twitter, Facebook, YouTube, TikTok and Instagram are the most popular sites. Statistical information on social networking apps is depicted in Table 2.2.

Table 2.1 Social media types and examples of each type

Social Media Types	Examples	
	Names	Website
Microblogs	Twitter Tumblr	twitter.com tumblr.com
Social bookmarking	Pinterest Delicious	pinterest.com delicious.com
Collaborative projects	Wikipedia Mozilla	wikipedia.org mozilla.org
Video sharing	YouTube Vimeo	youtube.com vimeo.com
Forums	IGN Boards Gaia Online	ign.com/boards gaiaonline.com
Enterprise social networks	Social cast Yammer	socialcast.com yammer.com
Business networks	LinkedIn XING	linkedin.com xing.com
Photo sharing	Flickr Instagram	flickr.com Instagram.com
Social gaming	Mafia Wars World of Warcraft	mafiawars.com warcarft.com
Social networks	Google+ Facebook	plus.google.com facebook.com
Blogs	The Huffington Post Boing Boing	huffingtonpost.com boingboing.net
Products/services review	Amazon Elance	amazon.com elance.com
Virtual worlds	Second Life Twinity	secondlife.com twinity.com

As a social networking and microblogging service, Twitter supports geo-tagging locations in their users' posts with images and videos. For this reason, Twitter plays a more important role in the events of the protests and is also used more in studies in this field [13]. The vast majority of data types used to predict protests were textual data from Twitter [14-17]. In addition, Facebook [18], newspapers [18, 19], blogs [19, 20], and Tumblr posts [21] have also been used.

Table 2.2 Statistical information of SNNs users

Online social network	Year	Monthly active users	Most active countries
Twitter	Aug.2022	397 Million	Unites States
Facebook	Aug.2022	2.9 Billion	India
TikTok	Aug.2022	1 Billion	India
Instagram	Aug.2022	1.4 Billion	Unites States

2.3.2. Twitter

2.3.2.1. About Twitter

Twitter is one of the fastest-growing microblogs. As of August 2022, the number of active Twitter users is 397 million, and it is among the top 11 in terms of the number of social media users [22]. Twitter allows users to write their messages in 280 characters. Tweets were originally limited to 140 characters until November 2017 when Twitter increased the character limit to 280 characters [23]. The messages being short is an effective feature for reading and transmitting quickly. Users can write messages in various areas, including their daily activities [24], life events [25], and the latest global and local news and events [26, 27]. Using Twitter, the news can be heard even faster than traditional media. According to empirical studies, Twitter is often the first social media to report news in less than a few seconds about natural events such as earthquakes [26, 28]. Users can use Twitter to have their daily conversations. Regardless of Twitter users' different motivations, this microblog can be considered a source of information.

Twitter users' relationship is asymmetric and forms a directional social network (directional graph) called the follower or follower network [29]. User A follows User B without obtaining permission from B. Twitter, does not limit the number of followers for users. Still, users can configure their account and change it to be accessible only to their followers. Generally, messages are available on Twitter for everyone. The user can read the messages of the users he/she follows through his/her timeline. User A can forward user B messages to his/her followers, which is called Retweet. Retweets are identified by the RT prefix preceded by the original author of the message (B) (RT @ username). A reply is a message that a user sends in response to another user's message. Reply messages start with reply-to @ username.

A message contains Mention if the text of that message has a Username. Users can specify the subject of their messages with the # sign, which is called a hashtag (like #birthday). Twitter provides two types of Application Programming Interfaces (API) for extracting messages. First is the Search API that can restrict the search of messages by location, time, author, hashtag, etc., and second is a Stream API that allows the user to extract tweets online. The availability of Twitter data has motivated researchers to conduct research in various fields and create tools and applications.

Analyzing this rich and continuous flow of user-generated content on social media, especially Twitter, can provide unprecedented valuable information. Even though this amount of information was not obtained by traditional media, it has led to the growth of studies and research in social media analysis. Tweets as a dynamic source of information enable individuals, government agencies, and companies to stay informed of what is happening right now. People are eager to receive new offers, facts, opinions, and news. At the same time, companies are increasingly using Twitter to promote and offer their products and services to customers, maintain credibility, analyze users' feelings by considering their products and competitors, respond to customer complaints, and improve decision making and business intelligence [30]. Twitter is also seen as a tool for predicting stocks in the stock market [31], election results [32] and crime [33], sharing political speeches and events [34], and monitoring tourism activities. Event analysis and identification has been introduced as a controversial topic in Knowledge Discovery and Data mining [35].

2.3.2.2. The Importance of Twitter in Social Media Research

Twitter is one of the most important mass media on the internet. This social medium is highly important in political issues. The word "Twitter Diplomacy" has been invented to use this social medium for diplomatic purposes [36]. The importance of Twitter in the diplomacy world after Donald *Trump*'s *administration* has increased more than ever [37].

Twitter is of special importance in social media research. The reason is that Twitter diplomacy about making its data available to the researchers. Twitter allows you to collect one percent of the available tweets about a hashtag. In this regard, one can conduct several studies on these tweets. However, considering this Twitter diplomacy, the number of research on Twitter is more than that on other social media such as Facebook, Instagram, and Telegram. Twitter is the pioneer of social media in academic research. For this reason, the data generated by the users on Twitter has been used in this research.

2.3.3. Social Media Mining

2.3.3.1. Definition

Considering the broad data published in social media, researchers identify the patterns from the data using data mining methods and algorithms [6]. It has been defined as [6]:

> "The process of extracting and analyzing social media interactions to recognize meaningful patterns in individual and social behavior."

> "The computationally process of social media data which patterns can be compute analytically."

Social Media Mining (SMM) has attracted researchers' attention in different fields such as economics, trade, sociology, computer, and even medicine more than ever. In the past six years, more than 2500 articles have been published in medicine using the data of social networks such as Twitter and Reddit [38, 39]. This shows the importance of data in social networks because these data have been widely applied in medicine as well. These articles present abundant applications.

2.3.3.2. Applications

Since the study focuses on Twitter, this Section reviews the related literature on SMM on Twitter. The researchers have been attracted to social media to predict and analyze current and future status. In this section, applications of SMM are discussed. Since this study focuses on the use of Twitter for prediction, this Section reviews the related literature on the use of Twitter. Several study fields use SMM, such as psychology, sociology, security, health, politics, financial affairs, sports, and other fields.

The methods which are used for SMM include machine learning and Natural language processing techniques. The first group consists of studies that have made predictions using the linear regression model and map features such as the volume of tweets or the score calculated from the sentiment analysis of tweets toward the event's occurrence. The second group is nonlinear models that use nonlinear machine learning algorithms such as support vector machine or logistic regression; time-based methods in which the temporal relevance of features such as the volume of tweets are examined using autoregressive algorithms. Finally, there are the last group of studies in which the models are introduced to solve the problem of event prediction in a specific field (using the exploration of the association rule to discover the most frequent names of players and predict the results of sports matches).

In summary, SMM has been successfully used to detect the user's profile information such as age and gender, and the presented techniques have high accuracy. SMM has been successfully used for psychology, but high accuracy techniques have not been presented for successful use of SMM in financial affairs and politics.

2.3.3.2.1. Psychology

Several works have been done on the use of social media for the prediction of the human psyche. The produced data in social media provide psychologists with the opportunity to identify the psychological model of users. In these articles, individuals' psyche has been analyzed using data mining techniques and machine learning algorithms. For example, to predict the incidence of depression in individuals, a model has been presented based on the individual's Tweets and machine learning algorithms. The main aim of using SMM for psychology is the psychoanalysis of individuals to prevent dangerous situations such as suicide [40-50].

2.3.3.2.2. Business

Social media have been long used for forecasting services and goods sales. The first works on SMM have been related to the prediction of box offices. Using the content and the number of tweets available for the film, one can successfully analyze sales of the film at the box office. The researchers have been interested in the prediction of good sales or the prediction of the stock price. Several works have been done on using Tweets content to predict the stock price in different stock exchanges and the prediction of sales of a special product [31, 51-59].

2.3.3.2.3. Political

Using the produced content in social media for political prediction and analysis is one of the most challenging and important cases for the researchers. Can the content produced on Twitter predict the result of the election?

This question is one of the ten questions which researchers working on political social media analysis try to answer. This content is highly important for the users as one of the members of society who attend countries' political trends. This content includes attitudes, views, and ideas of individuals about the politics of their country. Of course, it is necessary to note that users are more conservative for making political comments. One user can comment on a film fearlessly and easily, but it is difficult to comment on a political authority such as a country's leader. This is more difficult, particularly in countries with limited social freedoms [32, 60-67]. Despite this problem, users' content in social media can almost give a suitable and successful prediction.

2.3.3.2.4. Sport

Data mining methods and machine learning algorithms such as regression and sequential minimal optimization algorithm (SMO) have been used in the related works. Sports is one of the fields which attracted the attention of SMM researchers. Results of the sports matches such as car race, baseball, tennis, American football, basketball, and football have been presented using the content produced in social media. Betting and estimating the sports matches have had a market as large as $500 billion in 2018, which is estimated to increase to $6000 billion by 2026. Considering the largeness of this market, researchers and companies tend to predict the results of these matches. One of the best methods for predicting sports matches is using data mining and machine learning techniques. Several works have been done in this field, but no accurate techniques have been obtained yet [68-70].

2.3.3.2.5. User Profile

The users' profile is important for many trades. Twitter makes use of the users' profile information to present targeted advertising to the customers. In this study field, researchers try to identify the users' profile information such as age and gender. The presented techniques identify the users' profiles with high accuracy [71-80].

2.4. Social Media for Events

This section presents a definition of the event, then describes different events, including Scientific Events, Personal Events, Trade Events, Crisis Events, Disaster Events, and Protest Events. The four stages of incident detection on Twitter are described in this section. There are three approaches: online event detection, future event detection, and retrospective event in PEA. In this, these three approaches were explained with examples.

2.4.1. Event

A controversial problem in information retrieval is identifying the event in a stream of documents such as tweets [81, 82]. In the past, a stream of documents was created by collecting news articles from traditional media. But today, with the advent of social media, posts, free texts such as blogs, and conversational transcripts are used to create streams of documents. The use of time in the data retrieval process has created a new field called temporal data retrieval. This domain aims to retrieve the most relevant documents with the user query, which is also temporarily related to the user query [83, 84]. For example, the user query might be such that past documents must be retrieved.

Event detection is a branch of Topic Detection and Tracking (TDT) [82, 85, 86]. The field of research for TDT can be divided into five parts:

1. Story segmentation: Dividing a news transcript into separate stories
2. Subject recognition (cluster detection): Grouping documents (stories) based on topics
3. Identify a new event (identify the first story): Identify the start of a new topic in a stream of documents (identify an unseen story)
4. Story link detection (identifying related stories): Answering the question of whether two documents are similar or not (whether two random stories speak about the same subject)
5. Event Tracking: (Trend) Examining the evolution of an event (or topic) and describing how it occurs, identifying the topic, identifying the first story, and trend are some of the things that are related to event tracking.

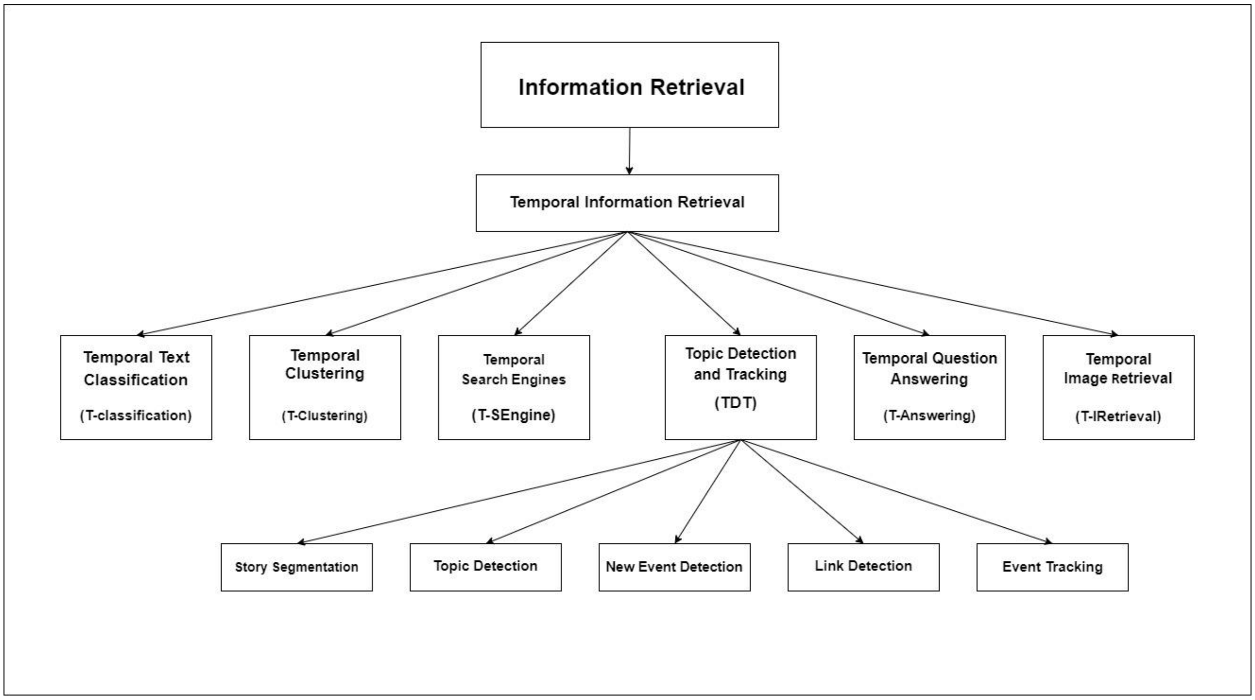


Figure 2.2 The location of event detection in information retrieval.

Figure 2.2 shows a view of the event tracking location on Twitter. An event is defined as an incident that occurs at a specific time [87] and place [82, 86, 88] and is associated with several factors. Due to social media's online nature, the connection between the event and the physical location can be omitted [89]. In social media, an event can be defined as an incident that changes textual data volume. This textual data is on a topic related to the event, and this change is at a specific time [90]. Mahmoud Hassan et al. [91] have defined the event on social media as follows: The occurrence of a public trend in the world (a topic of interest to people) causes discussion related to the event on social media by users - this volume of messages is also used for prediction right after it occurs. According to these definitions, the following can be mentioned as event attributes:

- Events are on a large scale because many users experience them.
- Events affect people's lives, which is the most important reason for users to mention the event.
- Events have a time and place range.

Identifying the event on Twitter has been the focus of many researchers and has its problems. Twitter data is noisy. Extracting information and identifying the event from Twitter is difficult due to the shortness of the messages, the use of slang, abbreviations, acronyms, and misspellings. Twitter also has meaningless messages [92], infected content [93], and rumors [94]. Therefore, the methods used to identify Twitter's event are different from the traditional methods for traditional media and even the methods used on other social media. Streaming documents in traditional media are well written and edited and may even be structured [95, 96].

2.4.2. Types of Event

2.4.2.1. Scientific Event

Seminar and conference are two scientific events for which the user creates content. The scientific events besides trade events and crisis events have attracted the attention of Twitter users. Scientific event details are shared on social networks to inform researchers about these events and engage the general public with science [97].

2.4.2.2. Personal Event

Marriage, employment, admission to the university, child's birth, graduation, and death of the dearest ones are among the personal events. Lee has identified 42 personal events [98]. The difference between personal and public events is the number of shared massages.

2.4.2.3. Trade Event

Product launch, Trade Shows, and meetings are three trade events for which the Twitter users create content. Trade events seek to achieve commercial goals by gathering experts and advertisement in social media such as Twitter.

2.4.2.4. Crisis Event

This Section relates to two scientific events that are discussed on Twitter. Disasters and protests are two crisis events for which Twitter users create content. Crisis events attract the highest attention in social media such as Twitter [99, 100].

2.4.2.4.1. Disaster Event

Disaster events include events such as earthquakes, volcanoes, landslides, famines and droughts, hurricanes, tornados, and cyclones. Sufi [101] has represented new software for analyzing Tornadoes using artificial intelligence (AI) based algorithms and disaster event dataset.

2.4.2.4.2. Protest Event

Protests are a form of social unrest that takes place in opposition to an idea or an action [102]. Social unrest can take many forms, such as protests, occupations, or strikes [103]. Protest happens worldwide, and people use it to voice their protest to the authorities [104]. Predicting any information about the place, time, and causes of protests is of great importance to rulers, policymakers, analysts, and scientists [105].

Not every gathering of people is necessarily a protest, and people can gather for other purposes, such as gathering fans of a sports team or celebrations, festivals, and parades. Therefore, there must be a precise definition of protest. Since not all possible protests can be categorized, it is better to consider non-protest gathering as criteria for exclusion. Gatherings such as concerts, festivals, sporting events, security events, or natural disasters are among these exclusion criteria [105].

In the proposed methods for protest detection, the following information is extracted from that event:

- Protest location
- Protest date/time
- Violence/No violence
- Reason for protest
- People who protest
- Size of the protest

Protesters usually fall into the following categories: religious, political, occupational, educational, legal, moral, media, trade union, or refugee. If the protesting population does not fall into any of the above categories, they will fall into the general population category. The protest event, like other events, has a time, date, and place where the protest takes place. The ultimate goal of event detection is to detect the protest event at a specific time, date, and place. One of the most critical features of protests is whether a protest is peaceful or violent. There is a vital difference between peaceful and violent protests. There are various reasons for the protests, including political, economic, and social reasons. The size of the protest is also an essential feature of protests. There is a profound difference between a protest of ten people and tens of thousands of people [105].

2.4.3. Event detection from Twitter

Figure 2.3 shows the general process of identifying an event from Twitter. The preprocessing component prepares messages for the next step by receiving tweets from the Twitter stream interface and processing them. Named-entity recognition, Part-of-speech (POS) tagging, etc., are some of the things that happen in the preprocessing stage.

The event detection component uses event detection techniques (based on document, features, and topic) to cluster the processed tweets, and each cluster represents an event (Figure 2.3).

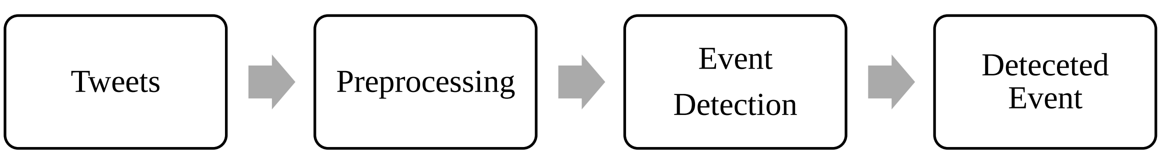


Figure 2.3 The overall process of event detection

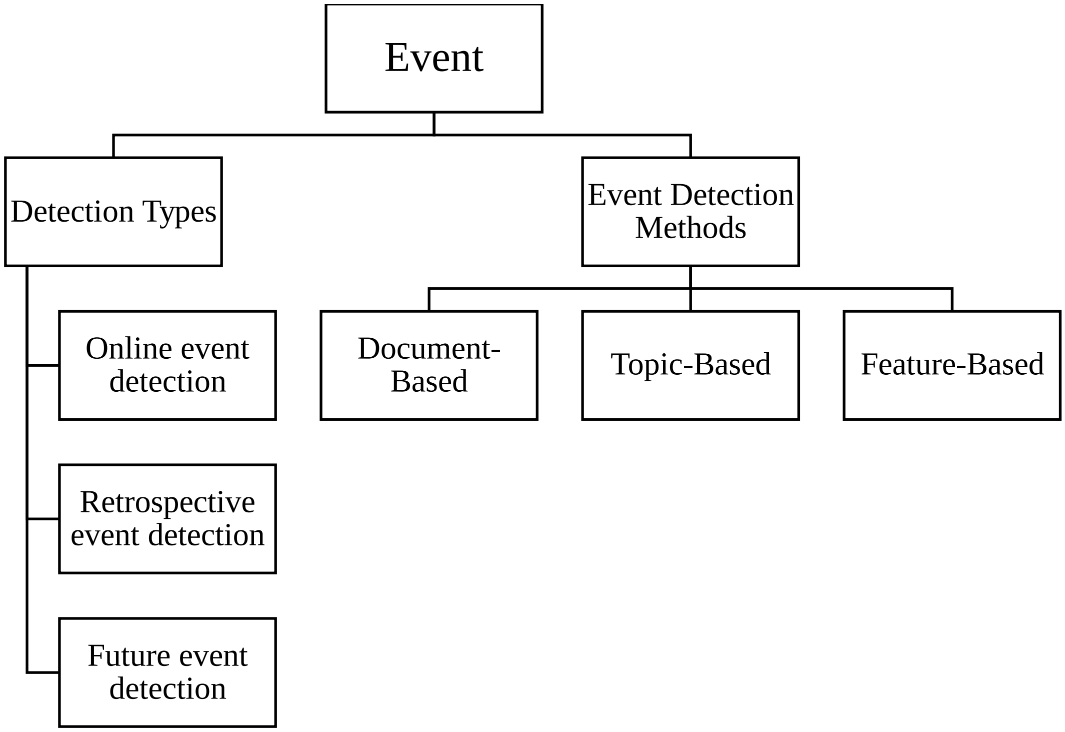


Figure 2.4 The subgroups of event

Event detection from social media can be examined from different aspects of detection type (retrospective, online, and future) and identification method (based on a document, feature, and topic). Figure 2.4 shows the subgroups of event.

2.4.3.1. Detection type in Event Detection

By constantly monitoring a stream of documents, one can discover the first story about a topic of interest, which identifies a new event. Depending on how social media messages are monitored and processed, event recognition systems can be divided into online new event detection, retrospective event detection, and future event detection.

Online event detection: The state in which a new event is detected by constantly monitoring the live stream of documents in the near future is called online event detection.

Event Detection at Onset (EDO) system [106] can detect the event within 3 to 8 minutes after being posted on Twitter. In EDO, data goes through three stages from the moment of arrival to event detection: the graph generation stage, the graph pruning stage, and the event detection. In the first step, to create an event graph, all tweets of a period are broken down into words. Words form graph nodes, and the edge between two nodes is equal to the number of tweets that contain two words. In the pruning stage, the important and emerged (exploded) words are identified by the Kullback-Leibler divergence score, which measures the amount of change in the occurrence of a word during the recent period compared to previous periods. Emerged words are words that have not been seen in the past but have been used suddenly lately. Words that score higher than one threshold are important. Different sources produce different types of information, the combination of which improves event detection systems [107].

Retrospective event detection: A set of pre-collected documents may be used to detect previously unknown events, which is called retrospective event detection. In this case, the basis of all methods is the retrieval of event-related documents by performing a TF-IDF query or analysis on the document set.

Future event detection: If it is possible to detect a future event's signs by monitoring the live streams of documents and using them to predict the event before the occurrence, it is called future event detection.

2.4.3.2. Detection method in Event Detection

Event detection methods from Twitter can be divided into two categories: document-based and feature-based [12, 108]. In some studies, another category called topic-based has been proposed [91]. Liu has presented in another category called clustering, statistics, and probability [109], which are equivalent to document-based, feature-based, and topic-based, respectively.

Document-based: In document-based methods, documents are identified by clustering documents based on the textual similarity of events, assuming that all documents are about the intended events. Since many unrelated messages (noises) are sent on social media along with the messages about the event, it is necessary to determine clusters related to the event. In document-based methods, the use of the TF-IDF technique to display tweets is common [110, 111].

Feature-based: Modeling an event in text streams as an explosion (a rapid frequency increase of a feature) is a common technique in feature-based methods. Zhang et al. [112], based on Markov's latent model, determined the exploded terms and then used them to create a directional graph whose nodes, exploded terms, and edges are the co-occurrence connection. Then, by applying graph-based clustering, the strongly connected components representing the event are extracted from the graph. The system also predicts event popularity based on user influence, user interest in the event, and volume of event tweets.

Topic-based: Topic-based methods use probabilistic models to extract hidden topics from Twitter streaming data to identify an event. In other words, each tweet is mapped to a probabilistic distribution of different topics to discover the hidden semantic structure of the collection of tweets and thus identify the event [113, 114].

3. SOCIAL MEDIA FOR EVENT PROTEST

3.1. Protests

Social protests refer to a collective action that aims at social change in some of the community. It usually occurs when power distribution method, laws governing society, or the policymakers' decisions and the authorities fail to solve the current major problems and meet the needs of the majority members of society [115].

New social protests have an accepted theory and are like a string of beliefs that allow understanding and defining the critical conditions and permitting continuing it or transferring it from a special situation to another situation. On the contrary, they are beyond the gathering of people to protest. They need some forms of organization and communication to allow continuing and stabilizing movement goals.

In contemporary social protests, the focus is on what is known as the life-world. Some factors such as identity, personal life, neighborhood, gender, and quality of life and lifestyle, and such protests seek to fundamentally change the social lifestyle, particularly democracy and civil society.

The following social factors effectively develop modern social movements: expansion of government, development of cognition-awareness industry, and new social media [116].

3.2. Social Media Role in Protests

The role of social media, such as Twitter, in social protests, is crucial. Many world political leaders have blamed social media, such as Twitter, for the protests. In their opinion, other governments and the government's opposition are trying to encourage people to protest and riot by spreading fake news, blackmail, and propaganda against them. In fact, these leaders believe that the source of the protests is not the demands of the people, but Twitter and similar social media.

India: Curfew, internet shutdown to control protests

3 killed, many injured, as protests widen against new citizenship law, seen discriminatory against Muslims

Iran shuts down country's internet in the wake of fuel protests

Netizen Report: Iraq and Ecuador face network shutdowns amid public protests

Technology and human rights news from around the world.

Posted 11 October 2019 18:16 GMT

Iraq shuts down internet again as protests intensify

POSTED ON NOVEMBER 4, 2019

Hong Kong fears internet shutdown after emergency powers are used to ban face masks

October 4, 2019

Figure 3.1 News related to the shutdown of Internet in protests

Figure 3.1 shows that in many countries, social media, and in some cases, the entire Internet is shut down to control protests. Some believe that the role of social media in the protests has been magnified. But what is certain is that social media, such as Twitter, plays a very important role in the occurrence, organization and continuation of protests, and comprehensive research should be done on this role and the behavior of users in the event of protests.

3.3. General Architectural View of Protest Event Detection

Event detection methods have influenced protest event detection methods. Event detection research is also rooted in an older field called Topic Detection and Tracking (TDT) because anomaly detection and topic modeling are the primary activities of event detection [117].

Event detection methods used data from traditional media like blogs, emails, web forums, and news websites. After the advent of social networks, social data streams have replaced these resources, and most protest event detection (PED) methods use social network data [118].

A protest detection method includes five steps: data collection, data preprocessing, location inference, feature extraction, modeling, and evaluation. Figure 3.1 shows general architectural view of PED. Text, image, sound, and video data sets are used as sources of PED methods. Data preprocessing is a common step in all data mining methods because data must be cleaned before being used by algorithms. Data in PED is collected from social networks, so the preprocessing stage is more critical due to non-standardized data in social media. The information published by a user who participated in the demonstration is more important than a user who is thousands of miles away. Identifying the user's location is one of the complicated steps in PED. Extracting the informative feature and choosing a better approach to improve accuracy, precision, recall, and F1-measure is the following step of the PED.

3.3.1. Data preprocessing

Preprocessing is done to achieve a clean text [119]. The primary purpose of this step is to clean the input data and convert it into a more useful form to be used in the creation of a protest detection model. For example, a protest detection model may need filtering noise components such as hyphens and pause words.

Typical tools in NLP are used in text preprocessing and various techniques are usually used to perform the preprocessing step. Some of these techniques include:

- Stemming
- Named Entity Recognition (NER)
- POS tagging
- Extracting time expression
- Converting slang words

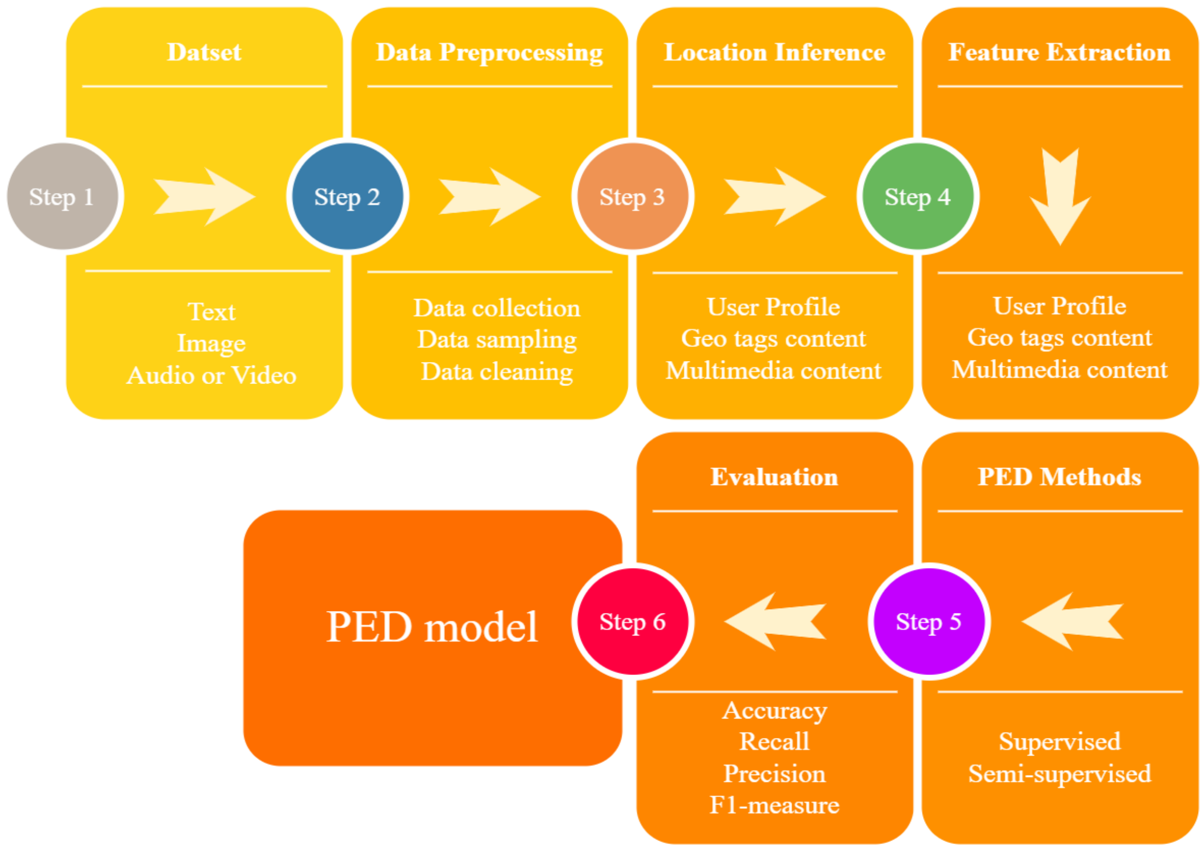


Figure 3.2 General architectural view of PED

The tools used in each of these methods are listed in Table 3.2 input text data is preprocessed by using different methods, such as, stemming, part-of-speech tagging (POS tagging), NER, user-name mentions, slang word conversion and removing stop words, URLs, time expression extraction.

Table 3.1 Preprocessing methods & tools

Preprocessing methods	Tools
Stemming	Lovins stemmer [120] Paice/Husk stemmer [121] Porter stemmer [122]
POS tagging	Gate Twitter POS model [123] Twitter-trained POS tagger [124]
NER	T-NER [125]
Slang-word conversion	Porter Algorithm [126]
Temporal phrase extraction	TempEx [127, 128], SUTime [129]

The preprocessing methods mentioned are for textual data. Video data has different preprocessing methods. Studies that use video or video data sources are beyond the scope of this survey.

3.3.2. Feature Extraction

For real-time analysis of events taking place during the protests, cross-sectional data cannot be beneficial and data streams are widely used in these cases. Textual data streams on social networks can provide the analyst with real-time information about a geographic location or a temporal period. From a textual data stream, features can be extracted that can reveal structural, temporal, spatial and semantic properties of the event in question. These data streams are usually categorized into two categories: specific or unspecified. Specific data streams usually already have pre-determined parameters such as time, place, and specifications. However, most of the existing approaches in the field of event detection fall into the unspecific category. After analyzing the content of the text, they deal with temporal and spatial issues separately.

3.3.2.1. Structural

In structure-based modeling, the role of words in tweets is examined grammatically with Part-of-Speech (POS). This feature is suitable when users follow the grammatical structures. POS tagging is a Natural Language Processing(NLP) process that categorizes words in a text, relying on the word's definition and context. Part-of-speech tags(Such as, noun, verb, adjective, and adverb) represent the characteristic structure of lexical terms within a text, tweet, or sentence[119].

3.3.2.2. Temporal

The temporal and historical information published in the content produced on social networks is essential to detect the protests accurately[119]. The exact time of a protest is one of the most crucial features of a protest. Because data streams on social networks are constantly updated, they can be good data sources for real-time analytics. In social unrest and protests, newer information is more important, and a tweet from three hours ago is far less valuable than a tweet from three minutes ago. PED methods usually use local protest time to analyze tweets, not a coordinated universal time, because this way, some information can be analyzed more accurately.

3.3.2.3. Spatial

The geographical information published in the content produced on social networks is also an essential element for detecting protests. Many protests are organized using social media. In addition, in these organizations, the location information of the protest is published so that the people tending to attend the protest can know the exact location of the protest. There are many ways to detect protests using these spatial features.

3.3.2.4. Semantic

Semantic features are another features examined in modeling the protest detection method. The techniques used to extract these semantic features include semantic analysis, semantic role labeling, and named-entity recognition. Thematic dimension-based identification techniques with supervised classification mechanisms, using semantic features, can classify tweets into different thematic categories, which are usually also considered the social content of users.

3.3.3. Detection models

This section examines the methods and models used in identifying social networking events in more detail. In order to distinguish between predefined sets of event classes, supervised or semi-supervised approaches are used in most event detection techniques. Detection models are also divided into three categories: unsupervised, semi-supervised, and supervised [130].

Most scholars in studies have looked at the PED problem as a hard classification problem. In previous studies, the primary purpose is to predict or identify protests. Will there be a protest or not? Answering this question is the main goal of scholars. The presented methods have used classification and regression algorithms. Machine learning algorithms such as Support Vector Machines(SVM), Naive Bayes, decision trees, and other algorithms were used to predict protests. Different types of regression algorithms were also used in many studies. In some studies, clustering is used in the initial stages. In the third part, these methods are discussed in detail.

Evaluation of PED methods is challenging due to the lack of a common dataset in the field. Evaluation analysis and specific benchmarks are needed to evaluate the performance of various event detection techniques. In an evaluation analysis, the model is usually trained and tested using the studied methods. Performance evaluations are typically measured with four standard criteria: precision, recall, F-measure, and accuracy [131]. Recall means that out of all the protest events that are happening, how many of them are correctly identified. Precision is the proportion of the correct protest events in the events that are classified as protest event. The high precision of a method means that the method has a low false-positive rate and the high recall means that the method has a low false-negative rate. The weighted average of recall and precision is the F-measure. Accuracy is the proportion of the correctly classified events to all events.

Social events are usually complex issues and include a variety of characteristics and components, each of which alone can affect the method's overall performance. Evaluating event detection methods can help us select the best method and understand the improvements needed for existing methods. A database with the same conditions is needed to evaluate these different techniques properly. Most of the work done in the event detection field focuses on a specific topic and evaluates the performance of their model based on a specific time and place.

Table 3.2 shows the comparison of PED methods in details. The methods that use features are in the majority. Only two studies have used the document approach approaches[132, 133]. One of the main goals of PED is to find the location and time information of protests; that is why spatial and temporal features are evident in PED methods. The content published on social networks aims to publish public information about protests, usually containing location and time information.

Table 3.2 Comparison of PED methods in details

Reference	Event Methods		Features				Detection Categories		Evaluation Metrics			
	Document-Based	Feature-Based	Spatial	Temporal	Semantic	Structural	NED	RED	Precision	Recall	F-measure	Others
[132]	▪				▪			▪				▪
[133]	▪					▪		▪	▪	▪		
[134]		▪	▪	▪	▪			▪	▪	▪		▪
[135]		▪	▪	▪				▪	▪	▪	▪	
[136]		▪	▪	▪	▪			▪	▪	▪	▪	
[137]		▪			▪			▪				▪
[138]		▪	▪	▪				▪			▪	▪
[139]		▪	▪		▪			▪				▪
[140]		▪				▪		▪				▪
[7]		▪			▪			▪	▪	▪	▪	
[18]		▪			▪			▪	▪	▪	▪	
[19]		▪	▪	▪	▪			▪	▪	▪	▪	
[141]		▪			▪			▪	▪	▪	▪	
[142]		▪			▪			▪				▪
[14]		▪			▪			▪				▪
[21]		▪			▪		▪		▪	▪		▪
[143]		▪	▪	▪	▪			▪	▪	▪	▪	▪

There are two main categories of protest event detection: new event detection (NED) and retrospective event detection (RED). As the name suggests, NED focuses on detecting new, live, near-real-time events, while RED focuses on discovering events from retrospective and historical data sources[12]. In the absolute majority of PED methods, the goal is retrospective event detection. In text mining studies such as classification of texts, there are common evaluation criteria such as accuracy, which makes the comparison clearer, but in PED methods, various evaluation criteria have been used.

3.4. Related Works on Social Media in Protests

In this section, essential contributions on PED are presented. All the PED works done cannot be presented in this section, so only new and significant contributions are reviewed. The protest detection models are discussed in the following.

Kallus [132] examines protests in Egypt and provides a model for predicting them. This study's main purpose is to predict the occurrence, time-frame, and locations of protests based on the data collected from social media and blogs. By testing data from March 6, 2013, to July 10, 2013, the results were true positive rate(TPR)=75.51% and true negative rate(TNR)=69.31% using the Random Forest Classifier(RFC). The balanced accuracy (BAC) with the average of these two is 72.41%, including a 44.8% reduction in balanced error from no data. The AUC value of the proposed method was 91.3%.

Bahrami et al. [134] predict protests using four machine-learning algorithms. They use Wrapper feature selection method with four classifiers namely: Naïve Bayes, C4.5, SVM , and Logistic Regression(LR). Daily tweet count, date, time, place mention count for each day, number of tweets calling for protest, percentage of negative and violent sentiment tweets, and Mrs. Clinton and president Trump's vote in each state were used as features. The data size is 0.47 million tweets about Trump's executive order banning citizens of seven Muslim countries. The precision was between 80%-88% to detect protest events.

Tuke et al. [135] developed a method called "Pachinko Prediction" to predict protests using Twitter data. In this study, the protests in Australia between 2017 and 2018 have been examined. Machine learning methods are used in this method to calculate posterior event probabilities as an empirical Bayesian approach and classify individual posts on social media as relevant. 226 events in 1663 days are analyzed, and a model for predicting the protests in those days is presented. The empirical Bayesian method has been used to achieve the best model using various features (tweets only, tweets + Month, tweets + Location, and tweets + Month / Location). The best feature combination includes tweets + Month/location with 76% accuracy. The main contribution of this research is the prediction of protests with a very small number of tweets. According to the authors, the proposed method can predict protests with 25 tweets.

Agrawal et al. [136] present the solution approach by using three components; semantic enrichment of events-related tweets (planning, commentary and mobilization, and crowd-buzz), named entity recognition (people expressions extraction, temporal, and spatial location), and location-time-topic correlation miner. It provides a method for early detection and prediction of protests using Twitter data (Two Million tweets from October 1, 2013, to February 28, 2014) and an ensemble learning approach. The dataset includes two events, “Fast for families protest” and “Christmas Island hunger strike”. Experimental outcomes demonstrate the high accuracy of classifiers in terms of F1-score (0.75 and 0.85 for “Fast for families protest” and “Christmas Island hunger strike” events, respectively).

To analyze location occurrences in extracted results, Muthiah et al. [137, 144] developed a probabilistic soft logic technique. This method combines key phrase learning for the identification of relevant information and time normalization for future mentions resolution. The steps of this research include Linguistic Processing, Phrase Filtering, Geocoding, and Warning Generation. The events of ten countries (Brazil, Argentina, El Salvador, Mexico, Chile, Paraguay, Colombia, Uruguay, Ecuador, and Venezuela) were examined. The dataset includes Twitter, Twitter URL, Facebook, mailing list, and blogs. Accuracy was reported to be between 81% and 93%. The authors believe that 75% of protests in countries considered illegal are planned and organized, and researchers can predict these protests by observing social media data like Twitter and Facebook.

Ertugrul et al. [138] provide a predictive modeling framework by using a spatiotemporal learning approach called ActAttn and long short-term memory neural network. The aim was to predict follow-up protests to the Charlottesville rally (Charlottesville, Ferguson I, and Ferguson II) in 2017. The AUC value of the proposed method was 85%.

In a study, Compton et al. [139] develop a framework to identify civil protests in Latin America by using LR. In this study, the input variables of tweets and the output variables are event type, population, location, date, and probability of unrest. The model has successfully predicted civil unrest events in Latin America on Twitter in terms of geographical and textual filters. This system manually examined two hundred eighty-three forecasts generated to reveal 157 forecasts about upcoming civil unrest events in Latin America and 126 forecasts for sports events, simple chatter, or other public functions.

Steinert-Threlkeld et al. [140] associated the coordinated messages (Hashtags, retweets, links , and mentions) on Twitter with the protests in the coming days. The focus of this research is on the protests during the Arab Spring. The dataset contains 14 million tweets from 16 countries in the Middle East and North Africa. The negative binomial regression model and polynomial regression model were used for prediction. There is a strong significant relationship between today's hashtag coordination on Twitter and tomorrow's protests. The negative binomial regression model shows that increasing one standard deviation in the measure of coordination today enhances the number of protests by 25.4% tomorrow.

Krolov et al. [7] provided a predictive model for Baltimore protests in 2015 in order to answer the following two questions.

- How social unrest is predicted using social media communications?
- Can mobilization in social media communication be identified?

LR, Clustering, and SVM algorithms are used to answer these two questions. The results have shown empirical evidence favoring the possibility of using social media communications to predict social unrest. The approach extracts the social science findings to direct and announce how to extract the information and perform data mining. However, there were not enough observations to claim that mobilization can be directly measured with it or other similar constructs based on social media.

Ramakrishnan et al. [18] adopt a multi-model approach to provide a comprehensive model for predicting protests. Features defined by three methods Volume-based, DQE (Dynamic query expansion), and Cascades model. Social media such as Twitter, news, web search, Wikipedia, blog feeds, and other sources are used. The database included Twitter's public API, Bloomberg financial news, RSS news and blog feeds, Talkwalker alerts, Healthmap's alerts and reports, TOR usage data, Google Flu Trends, NASA satellite meteorological data, OpenTable's restaurant cancellation data, web-pages referenced Tweets, and the PAHO health survey. Research main question is “Are Big Data Analytics Systems like EMBERS capable of forecasting civil unrest events?”. Input variables are streaming data from Twitter, Facebook, and RSS feeds, and output variables are date, location, type of events, population, and probability of event occurrence. The EMBERS model obtained a quality score of 3.11 out of 4.0 with an average lead time of 8.8 days and high precision and recall (0.69 and 0.82, respectively) during 15 months among 10 Latin American countries.

Korkmaz et al. [19] have provided a predictive model for Latin American protests. Based on today's characteristics, this study predicts the next day's protests and LR with Lasso Model is used for prediction. The predictive performance for protest forecasting in ten Latin American countries from November 1, 2012, to August 31, 2014 using regularized LR ranges from 55% to 91% in terms of accuracy. The average recall is high in all countries (0.94 to 1.0), and the F1-Score is about 0.68 to 0.95. One of the few kinds of research that uses different sources and languages to create a protest detection model. Twitter, news, political event databases, Tor (The Onion Router), blogs, and exchange rates are data sources in this research. According to the findings, Twitter, blogs, and news are essential sources for forecasting.

An event forecasting model was introduced by Cadena et al. [133] that employs the notion of activity cascades (follower and combined mention plus retweet) on Twitter to forecast the possibility of protests in Brazil, Mexico, and Venezuela. The dataset contains 353 million tweets collected in 1.5 years. The Lasso based LR model was used to predict the probability of a protest on a given day. The accuracy of the model varies from country to country. The proposed model predicts Mexico's events with 95% accuracy and the events of Venezuela and Brazil with 76% accuracy.

Bakerman et al. [141] developed a dynamic LR model to predict future protests one day ahead and five days ahead. The Bayesian variable selection was used to set the parameters of the dynamic LR model. It provides a method for predicting future protests using Twitter data (500 million tweets). The F1-score of the proposed method was 97% to 100%.

Timonda et al. [142] develop a variance-in-time method that smooths google trends scores over time and contains variance in interest in attempts to recognize protests during the BLM movement in the Baltimore protests of April 2015. The research findings highlight the success of using Google Trends in forecasting protests. The AUC value of the proposed method was 90%.

Zhao et al. [14] predict civil unrest using a novel feature learning model based on fused-overlapping group Lasso and an Nth-order strong hierarchy from multiple data sources with different geographical levels (country, state and city level). The AUC value of the proposed method was 91%.

Table 3 shows the protest detection key characteristics, limitations, and performances. Different types of regression algorithms(Logistic regression[7, 134, 139], Regression model[140], The Lasso-based logistic regression[19], and Lasso model[14]) are widely used in the PED methods. Algorithm SVM [7, 134] is next. Different evaluation methods have been used to measure the presented methods' performance, making it difficult to compare the methods. The best performance is related to Bakerman's study [141] with F1-score=97, and the lowest performance is related to Compton [139] with 55% accuracy in predicting protests. Different types of features have been used to predict protests. Features in text mining, such as keywords to economic indicators, have been used in PED methods.

Xu et al.[21] accurately predicted social protests using Tumblr posts and keyword filters. The Tumblr data was collected between 2013-04-01 and 2013-11-04. The study aimed to predict upcoming unrest events in the future based on cascading textual filters. Three filters (Keyword, location, and future data) were applied to determine the nature of the post. The results show that the proposed method can predict upcoming unrest events in the future based on cascading textual filters. Detection rates per month are 97.65% in June, 92.06% in July, and 93.70% in August by order of precision.

Alsaedi et al. [143] use temporal, spatial, and textual features to predict a riot in a small and large protest event. They evaluated the framework on Twitter's large-scale, real-world dataset (2011 riots in England). The presented method is a hybrid method including classification (Naive Bayes, SVM, and LR), online clustering and summarization.

In addition to the papers in Table 3, other studies have been published at conferences represented. Renaud et al. [145] predicted civil unrest events following the 2016 U.S. Presidential Election using a Random Forest classification, highlighting social network structure's importance in studying protests. Grill [146] has examined the main concepts and future risks of civil unrest in a study. This study has reviewed papers and products related to civil unrest before 2019. Hua et al. [147] use transfer learning and label propagation to estimate the locations of events(such as civil unrest, crimes, and disease outbreaks). The proposed Semi-Supervised Targeted-Interest Event Detection (STED) using Twitter data from Latin America. Williams et al. [148] compare offline action and online expression in the 2011 riots in England by the tension engine component of the Cardiff Online Social. They believe that by monitoring tension on social networks, the occurrence of violence is predictable. Many methods are only provided for large-scale protests. Some methods have been evaluated on many limited protest events. The proposed solutions for the limitations of the methods presented in the PED are detailed in the suggestions section.

Hossny et al. [15] use most associated keywords extracted from past protests to create a prediction model for future protests using Twitter data. After testing the keywords as informative features with algorithms, Naive Bayes, SVM, LR, KNN and decision trees, the F1 score up to 0.79 was reported. This study was successful on the protests that occurred in Australia and Indonesia. Rule et al. [149] developed the pressure detection and analysis algorithm to detect phrases of interest and afterward analyzes the pressure surrounding the related events. Gallagher et al. [150] represent a comprehensive content and topics analysis of tweets of #BlackLivesMatter and #AllLivesMatter hashtags during the Fergusen, Missouri protests in 2015. This study is essential in terms of studying divergent discourses in protests. Zhao et al. [17] used an unsupervised approach (Dynamic Query Expansion) to forecast PED. This study generates a homogeneous tweet graph and expands domain-related terms. Goode et al. [16] use tweet volume as the data source to present a differential game theoretic approach to characterize the cost of participation. Qiao and Chen [151] use Hidden Markov Models to predict the opposition protests in Southeast Asia. Table 3.3 shows protest detection features, methods and performances in details.

Table 3.3 Protest detection features, methods and performances

Collection of studies in the field of protest detection			
Ref	Key characteristics	Limitations	Performance
[132]	Using RFC with violent language by the fraction of n-grams	The performance of the RFC was not compared with other algorithms.	0.91†
[133]	The Lasso-based logistic regression based on activity cascades and tweet volume	The proposed model is incapable of predicting protests in the coming days.	0.76-0.95*
[134]	Deploying machine learning algorithms with event specific features	Because the dependence of method features on a specific event, it is complicated to use in other events.	0.80-0.88*
[135]	A pachinko prediction approach based on the empirical Bayesian approach and day, month, location and informative tweets	The implementation of the method is ambiguous and average performance is obtained.	0.74*
[136]	A combination of ensemble learning and Spatio-temporal and named entity	The model is tested on a minimal number of events.	0.75-0.85§
[137]	A combination of key phrase learning and Probabilistic Soft Logic	The method is unable to identify the event on the current day.	0.81-0.93*
[138]	Deploying neural network algorithms The spatiotemporal structure	The model is tested on three BLM protest events.	0.85†
[139]	Using LR with keywords, temporal, and date features	The accuracy of the model is low.	0.55*
[140]	Design of Regression model based on hashtags, retweets, links, and mentions	The method can identify only large-scale protest events.	--
[7]	Deploying LR and SVM algorithms with related tweets about mobilization	The model is tested on a minimal number of events.	0.82-0.85*
[18]	Sequential probabilistic based on volume-based, DQE, and cascades model	The model is only capable of large-scale protest events.	0.69£
[19]	Multi-source models based on the Lasso model	The method can identify only large-scale protest events.	0.68-0.95*
[141]	A combination of Dynamic logistic regression Twitter terms	It is necessary to check the reported performance on other protest events datasets.	0.97 to 100§
[142]	LR based on the mean and variance of a search term	The method relies only on Google Trends data.	0.90†
[14]	Lasso model based on incremental multi-source feature learning	The model is only capable of large-scale protest events.	0.91†
[21]	Filtering approach with keyword/hashtag counting	The algorithm used for implementation is ambiguous.	0.93-0.97£
[143]	Hybrid approach with temporal, spatial, and textual features	It focuses on a specific type of protest.	0.86*

*Accuracy †AUC §F1-score £Precision

3.5. Protest Event Detection Applications

PED is indeed an important tool that has a range of applications for different groups. One of the key areas where PED can be useful is in protest management, as it helps authorities to predict and prepare for potential protests. By analyzing various factors such as social media activity, historical data, and current events, PEA (Protest Event Analysis) can provide valuable insights into the likelihood and location of protests. Another application of PEA is the identification of protesters' identities. With the help of advanced technologies such as facial recognition and license plate readers, PEA can enable authorities to quickly identify individuals who are participating in protests. This can be particularly useful in situations where there may be concerns about public safety or criminal activity. In addition to protest management and identification, PEA can also help to shed light on the reasons behind protests. By analyzing social media conversations and news coverage related to protests, researchers can gain a better understanding of the underlying issues that are driving people to take to the streets. This information can then be used to inform policy decisions and to address the root causes of social unrest.

Social protests are inevitable, especially in difficult economic and social conditions. Public dissatisfaction eventually leads to protests [152]. In recent months, people protested the high energy price in the most prosperous European countries [153]. The reasons for protests can vary widely, ranging from economic grievances to political dissatisfaction or even human rights violations. However, regardless of the specific cause, the social and economic impact of protests can be significant. For example, civil unrest can disrupt supply chains, damage infrastructure, and lead to a decline in tourism. Moreover, protest movements can be particularly costly if they turn violent, leading to property damage, injuries, and even loss of life.

Considering the inevitability of protests, managing protests is the only way to deal with their high economic and social costs. One of the most basic applications of protest identification methods is its management. Anticipation of protests and violence can be used for more effective management. To manage protests effectively, it is crucial to understand the underlying causes and motivations of the demonstrators. Policymakers should engage with protesters and address their concerns proactively. By doing so, they can help prevent peaceful demonstrations from escalating into violent conflicts, which can have severe economic and social consequences.

Many studies have been done about the economic costs of protests in recent years [154]. Unexpectedly, the protests' cost after George Floyd's death has been estimated at two billion dollars [155]. Violence has high human and financial costs [156]. Providing an effective method for indicating violence can prevent these costs. Protest-related violence can lead to significant physical injuries and property destruction, which can result in high medical costs and lost productivity. Furthermore, such events can cause long-lasting psychological trauma for both victims and witnesses. Therefore, effective prevention strategies are needed to reduce the likelihood of violence during protests. Law enforcement agencies can use various techniques, including crowd control tactics, de-escalation strategies, and communication methods, to prevent violent clashes.

The costs of protests are not only limited to economic costs, but society's mental health is also negatively affected by protests and violence [157]. A high prevalence of mental problems after protests has been reported among community members, even when peaceful [158]. The psychological impacts of protests can be severe, particularly for vulnerable groups such as children, the elderly, and people with pre-existing mental health conditions. Such individuals may experience anxiety, depression, or post-traumatic stress disorder (PTSD) due to exposure to violence or witnessing traumatic events. Therefore, it is crucial to provide mental health support services to those affected by protests, including counseling, therapy, and other supportive interventions. Moreover, policymakers should take steps to address the underlying societal issues that lead to protests, including poverty, inequality, and social injustice. By addressing these root causes, they can help reduce the likelihood of future protests and their associated social and economic costs.

The identity of those participating in protests is not only important for researchers and governments, but also for the safety and security of all involved [159]. For instance, during the widespread protests that occurred across the United States in 2020 following the killing of George Floyd by police officers, there were concerns raised about the presence of outside agitators. Protest identification methods could help verify whether such concerns were warranted or not, and provide a clearer picture of who was actually participating in the protests [160].

In addition to identifying participants, protest detection methods can also shed light on other aspects of social movements. For example, researchers can use these methods to analyze the language and rhetoric used by protesters to better understand their motivations, grievances, and goals. This can be particularly useful for policymakers who are looking to engage with protesters and address their concerns in a meaningful way.

The use of masks in social protests has become increasingly prevalent in recent years, as protesters seek to conceal their identities from law enforcement and other authorities. While this can make it challenging to identify individual protesters, protest detection methods can still provide valuable insights at a broader level. For instance, researchers can use these methods to analyze patterns of activity and communication among groups of protesters, even if they cannot determine the identity of each individual participant.

Examining the contents produced by protesters in social networks can reveal not only their political views and affiliations, but also their social and cultural backgrounds. This information can be useful for understanding the diversity of perspectives within a social movement, as well as potential areas of overlap or disagreement between different groups of protesters. Moreover, it can help researchers and policymakers understand the broader social and economic contexts that give rise to mass mobilization, and how these contexts interact with political factors to shape the direction and impact of social movements.

Protests are a common form of expression in contemporary societies, reflecting the voices of groups who feel marginalized or disempowered. Protests can arise for a variety of reasons, including economic, social, political, and cultural factors. For example, economic inequality, corruption, discrimination, and lack of access to basic services can all be catalysts for protests.

However, not all protests turn violent. The use of violence during protests may be perceived as a means to achieve desired outcomes when other avenues have been exhausted or appear ineffective. It can also be a result of a lack of trust in the authorities or a perception that the government is unresponsive to the needs of the people. Moreover, the behavior of individuals within a protest can be influenced by external factors such as provocateurs, police brutality, or the presence of weapons.

In recent years, social media has become an important tool for organizing and mobilizing protests. Through monitoring user behaviors on these platforms, researchers can gain insights into the underlying reasons why protests become violent. Social media provides a platform for sharing information, coordinating actions, and amplifying messages, but it can also contribute to the escalation of tension and violence.

Understanding why some protests turn violent is critical for policymakers, law enforcement agencies, and civil society organizations. By identifying the underlying causes of violent protests, stakeholders can take measures to address them, prevent future occurrences, and preserve the right to peaceful assembly.

4. PROTEST EVENT DETECTION BASED ON TWITTER’S USER BEHAVIORS

4.1. Overview

This chapter explains the presented method for identifying protests and violence in detail. One of the most challenging steps in protest event detection (PED) methods is finding and collecting datasets to implement the method. Usually, each method has a unique dataset, which is impossible to share Twitter users' data due to legal restrictions. This issue makes it difficult to compare methods in PEA. In this chapter, the study datasets are described in detail. Then the relationship between users' behavior and objections is analyzed. The dataset Occupy Wall Street was used to find the relationship between users' behavior and the occurrence of protests. Then, based on the findings of this implementation, the proposed method was explained in the next section.

4.2. Datasets

This section explains the datasets employed in this research. The Occupy Wall Street Dataset as the primary dataset, which includes 1500 tweets about Occupy Wall Street, is first described. This dataset consists of 500 tweets related to one day before the protest, 500 tweets on the day of the protest, and 500 tweets on the day after. The Emotion Analysis dataset is related to emotion analysis, including 47289 tweets, and is used to produce the emotion analysis model. The third dataset includes tweets related to the Event Protest. This dataset has millions of tweets about the BLM protests in the United States. Table 4.1 describes datasets of study and their function.

Table 4.1 Datasets of study and their role

Datasets	The role in thesis
Occupy Wall Street Dataset	This dataset was used to identify the behavior of Twitter users. The proposed method was first tested on this labeled dataset. After the success of this method was specified, it was implemented on the BLM dataset. Dr. Wang has collected and labeled this dataset with 5 coders [9].
Emotion Analysis Dataset	This dataset was used to identify the Opinion Share behavior. By using this dataset, the Opinion Share behavior was identified in the BLM dataset [161].
BLM Dataset	The main research dataset to identify the occurrence of protest and violence in BLM protests.

4.2.1. Occupy Wall Street Dataset

In recent US history, the Occupy Wall Street movement is arguably one of the most recognizable protest movements [162]. A part of the New York Times website has even been allocated to the movement, which collects all related details about the movement [9].

Organizational efforts focusing on the protest initiation were captured by pulling tweet samples from September 16, 17, and 18. Pre-event, day-of, and day-after tweets of the event were chosen for analysis. Large amounts of raw data are not permitted by Twitter to be released. Therefore, researchers pulled tweets by date and search term using a software tool. The related tweets were identified using the hashtag #occupywallstreet. The hashtag selection was due to event organizers' and participants' use of it as the official hashtag. Each selected date had the last 500 tweets, which were chosen for analysis, giving a total number of 1500 tweets [9] .

This primary dataset has been used to investigate the significant relationship between the protest day and different features. This dataset has been used to find whether there is a relationship between the protest day and users' behaviors and tweets' emotions. In this dataset, 1500 tweets were investigated carefully, and five coders labeled the users' behavior. Such a dataset is important because it can reach the important cases in the tweets to approach the protest day. The careful investigation of this dataset will achieve the social networks users' behavior analysis model more simply and more carefully because these behaviors have been investigated manually by the experts. The labeled behaviors in this dataset are discussed in Section 2.2.3.

4.2.2. Emotion Analysis Dataset

In emotion analysis, the supervised approach has been used to classify tweets. The labeled dataset creates the classification model by deep learning algorithms [161]. Specifications of the labeled dataset are found in Table 4.2.

Table 4.2 Emotion dataset specifications

Emotion	Neutral	Happy	Sad	Hate and Angry
Number of documents	96643	16297	15937	5410

In the dataset, Neutral emotion is marked as 0, Happy 1, Sad 2, and Hate and Angry emotions as 3. The emotions dataset has been used to create the emotion analysis classification model. In this dataset, the number of neutral tweets is 9643, the number of happy tweets is 16297, the number of sad tweets is 15937, and the number of tweets containing hate and angry messages is 5410. This dataset has been published in the article written by Buazizi [163].

4.2.3. BLM Dataset

Twitter datasets covering the entire Black Lives Matter movement are up to date. There are 41.8 million posts from 10 million users in the dataset used to identify spatial and temporal patterns using the timestamps and longitude/latitude coordinates, the linguistic patterns related to social networks through follower/friend user data, inter and intra- movement dialog, the social movements and their counter-protests, and the publication of misinformation and news through news articles linking retweets and tweets. From January 2013 to June 2020, Tweets containing the keywords AllLivesMatter, BlackLivesMatter, and BlueLivesMatter were collected from the Twitter API [164].

Figure 4.1 shows how to create the final dataset for the assessment of the method. This figure shows stages of converting tweet IDs into CSV file with the given specifications to dataset preparation.

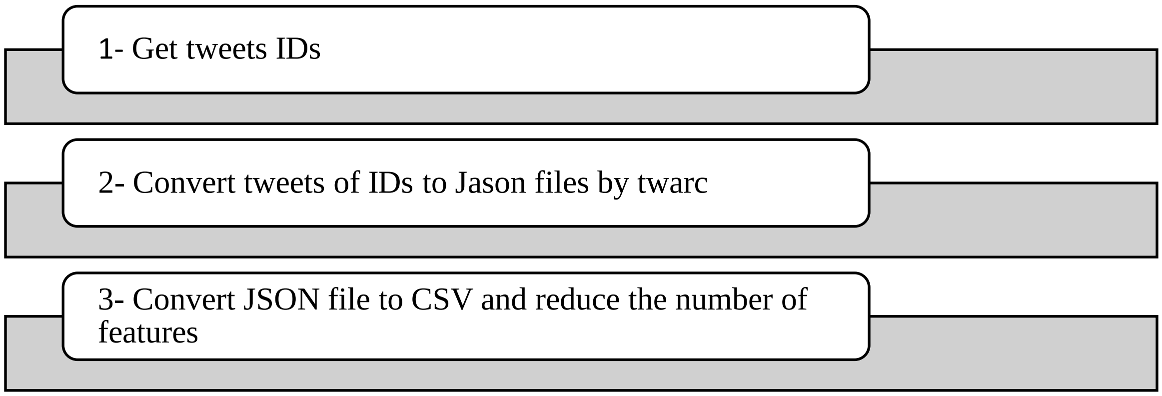


Figure 4.1 Stages of converting tweet IDs into CSV file with the given specifications to dataset preparation

JSON is the abbreviation of JavaScript Object Notation, meaning the notation of objects on JavaScript. JSON is an open standard file format that has allowed exchanging data on the web using attribute-value pairs. JSON is the only method for representing data objects when they are sent on the internet. This standard is a substitute for XML and is native JavaScript. For this reason, it is so popular and can be used widely. The most common use of JSON is to fetch data from web servers upon request. JSON files are large since they have all details of tweets. For example, the dataset #BlackLivesMatter includes more than 200 Gigabytes of the JSON content. For this reason, the additional contents are deleted for simplification and converted into CSV format.

The tweet IDs are converted into JSON files with twarc library before conversion. Based on the rules of Twitter, no one is entitled to share the contents of the tweets and can only share the tweet IDs. Archiving Twitter JSON data can be facilitated through the use of Twarc, which is a Python library and command-line tool. A JSON object represents each tweet, which corresponds exactly to the output obtained from the Twitter API [165].

Two datasets about the Black Lives Matter (BLM) movement are combined in order to create the dataset that is used in this dissertation. Twitter corpus of the Black Lives Matter movement in Giorgi's paper [166] and the ground-truth data obtained from the website of acleddata.com[1] on the occurrence of protest events during the movement. This dataset contains 41.8 million tweets which are collected from ten million users and it includes Black Lives Matter (BLM) movement tweets from 2013 to 2020. Tweets in this dataset are published as tweet ID, and then the contents of the

[1] https://acleddata.com/data-export-tool/

tweets are downloaded with Twitter APIs. 19,388,309 tweets are used in this paper because the information published on the acleddata.com website is about the movement after the death of George Floyd in 2020 and the website data contains events from May 28 to June 30. The following two filters are applied to the downloaded tweets to obtain the final dataset:

- The location field of the tweet's bio information contains the name of cities.
- The presence of hashtags *BlackLivesMatter*, *AllLivesMatter*, and *BlueLivesMatter* in a tweet

After the filtering operation, the initial 19,388,309 tweets are reduced to 2,908,634 tweets in the final dataset. The final dataset contains tweets in 34 days between May 28 and June 30 and each tweet has the city information in its location field. Figure 4.2 shows the number of tweets per day before and after applying the filters.

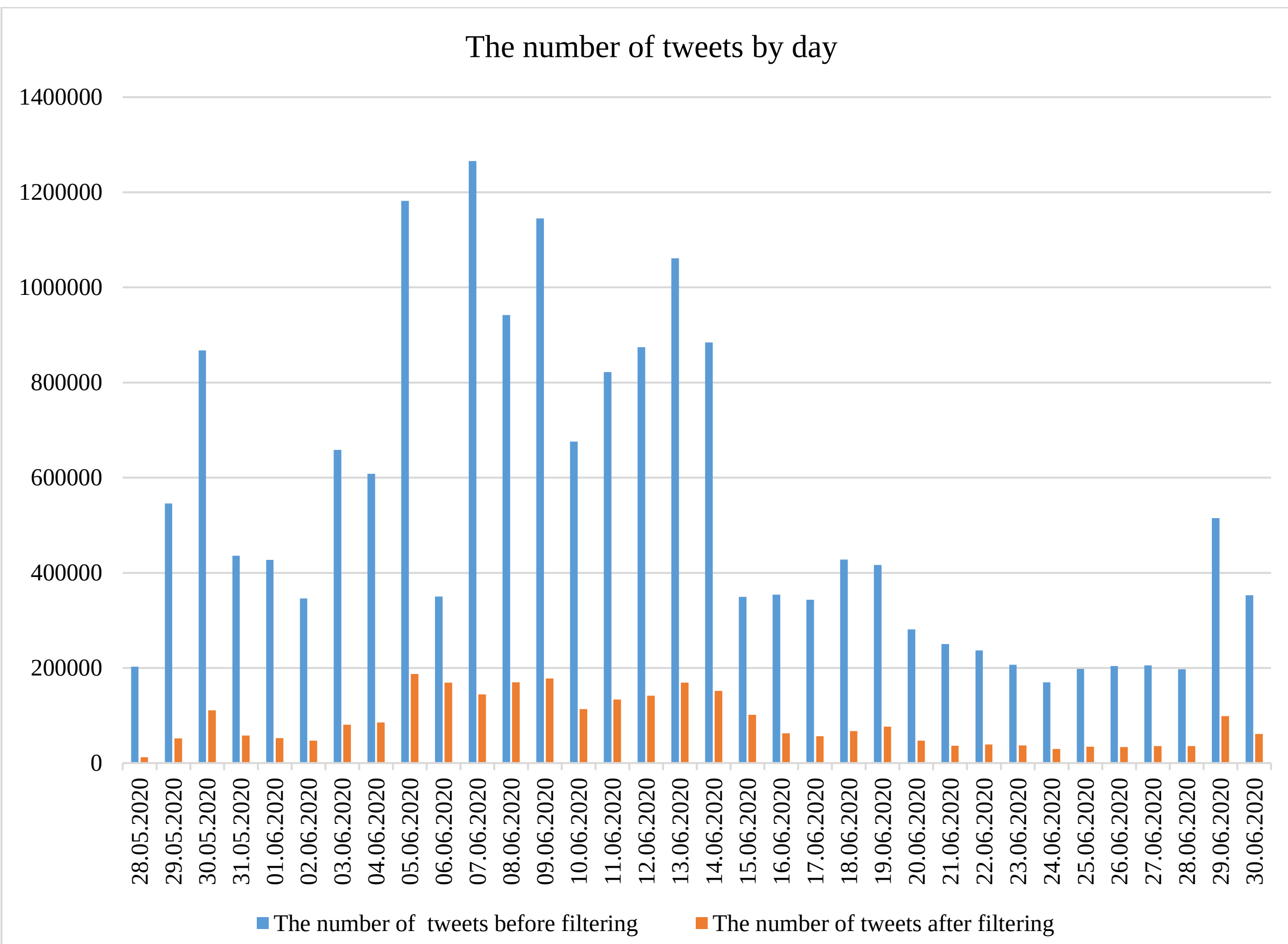


Figure 4.2 The number of tweets by day

From the figure, we can observe the following:

- The highest number of tweets before filtering occurred on June 5, 2020, with 1,182,180 tweets.
- The highest number of tweets after filtering occurred on June 9, 2020, with 177,587 tweets.
- The lowest number of tweets before filtering occurred on May 28, 2020, with 202,849 tweets.
- The lowest number of tweets after filtering occurred on May 28, 2020, with 12,624 tweets.

This figure provides an overview of the daily tweet activity related to the Black Lives Matter movement during the specified period. It shows the significant reduction in the number of tweets after applying the filters, which helps focus on the relevant tweets.

One of the most challenging steps to obtaining the dataset used in this paper was to get information about the event and protest types. Table 4.3 provides a record of the data expected on the acleddata.com website.

Table 4.3. The information is available in the acleddata.com dataset

Event ID	Event Date	Sub event type	City	Notes
5129	30.05.2020	Peaceful protest	Charlottesville	On May 30, 2020, about 1,000 people marched in a protest in Charlottesville (Virginia) in support of the Black Lives Matter movement and against police brutality and the death of George Floyd. [size=about 1,000]

Table 4.3 shows that the website data provides the event ID, date, type, location, and description. In addition to the kind of event, information on the size of the protest and the number of participants is also provided. Then the information about the occurrence of protests is added to the final dataset tweets with spatial information about cities using the website data, if there is a protest

on the specified day and place, the protest column will be equal to one, and if no protest occurs, it will be equal to zero. If this type of protest is violent, the type of protest will be equal to one, and if it is peaceful, the column will be equal to zero.

The last step to obtain the dataset is extracting user behavior features. For all the tweets on a specific date and location, user behavior features are received by the methods presented in Section 3 and added to the dataset. The structure of the dataset after filtering and feature selection is shown in Table 4.4.

Table 4.4 The final dataset

Features	Sub features	Sample	Description
Event ID	ID Number	5129	Unique ID for each protest
Date Information	Date	30.05.2020	Date of protest
	Month	May	Month of protest
	Day	Saturday	Day of protest
Event Location	City	Charlottesville	City of protest
	State	North Carolina	State of protest
Number of tweets in the date and location	Number Of Tweets	1978	Total number of tweets in Charlottesville on May 30
The ratio of Tweets containing EventLogShare Behavior	Date Info	0.03	Date tweets to total tweets ratio
	Time Info	0.1	Time tweets to total tweets ratio
	Place Info	0.129	Spatial tweets to total tweets ratio
The ratio of Tweets containing Opinion Share Behavior	Sad	0.291	Sad tweets to total tweets ratio
	Happy	0.198	Happy tweets to total tweets ratio
	Hate-anger	0.185	Hate and anger tweets to total tweets ratio
The ratio of Tweets containing General Information Behavior	Neutral	0.326	Neutral tweets to total tweets ratio
Event Information	Event Type	1	0: Non-protest events; 1: Protest
Protest Information	Protest Type	0	0: Peaceful; 1: Violent

The final cleaned dataset contains 1414 protest events and 3078 non-protest events from 460 cities in 37 states. A non-protest event means no protest occurred on that day and location. Protest events include 1414 protests in the BLM movement between May 28 and June 30, among which 285 were violent, and 1129 were peaceful (Figure 4.3).

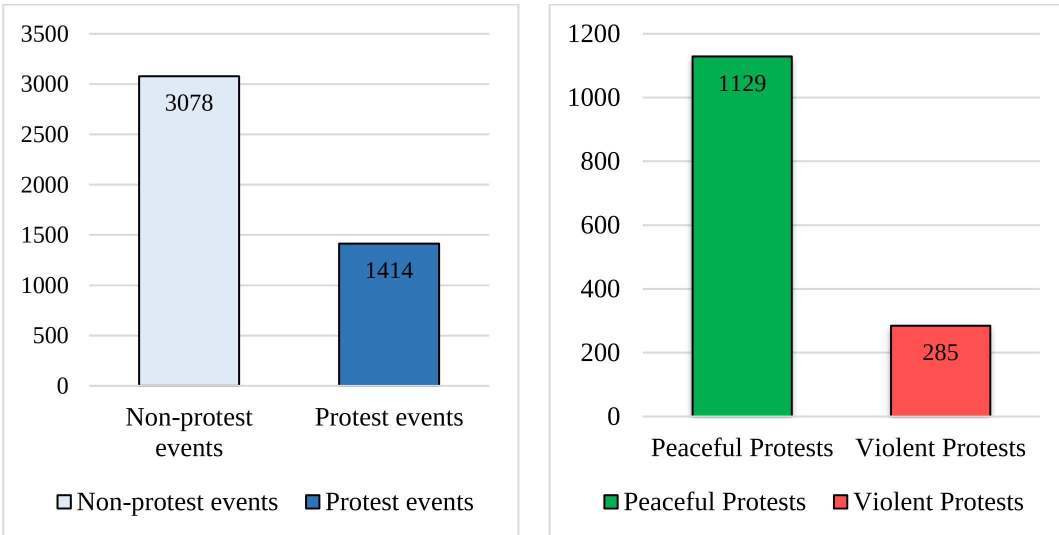


Figure 4.3 The number of protests and protest type in the dataset

Tables 4.5 and 4.6 give the observed number of protest events and non-protest events for the state and city. Table 4.5 shows the states in which the most frequent protests occurred. The state of North Carolina is ranked first with 116 protests. One hundred sixteen protests occurred in eleven cities of North Carolina in 34 days.

Table 4.5 The states with the most frequent protests

States	Protest events	Non-protest events
North Carolina	116	216
Ohio	93	168
Illinois	90	171
Florida	80	209
Pennsylvania	72	146
Tennessee	61	70
California	66	190
Oregon	61	135
Texas	55	170

Table 4.5 shows that North Carolina, Ohio, and Illinois are the top three states with the highest number of protest events. By examining social media posts, sentiment, and engagement in these states, researchers can potentially predict future protests and understand the underlying causes and motivations. Similarly, Table 4.6 shows the cities with the most frequent protests, with Portland, Richmond, and Seattle leading the list. Analyzing social media data from these cities can provide valuable insights into the dynamics of protests, the issues being raised, and the factors contributing to the frequency of these events.

As shown in Table 4.5, the states in which the most frequent protests occurred. The state of North Carolina is ranked first with 116 protests. One hundred sixteen protests occurred in eleven cities of North Carolina in 34 days. Based on the data provided, it is evident that North Carolina has the highest number of protest events (116) compared to other states, followed by Ohio (93) and Illinois (90). It is also interesting to note that the ratio of protest events to non-protest events is highest in Tennessee, with 61 protest events and only 70 non-protest events. This suggests that the intensity of protests in Tennessee is relatively higher than in other states.

Furthermore, despite a larger population, California has fewer protest events (66) compared to states like North Carolina, Ohio, and Illinois. This could indicate that the factors driving protests in these states may be more pronounced or that the population is more politically active. Additionally, Oregon has a relatively high number of protest events (61) compared to its non-protest events (135), suggesting a higher level of political engagement or specific local issues driving protests in the state. The data in Table 1 highlights the varying frequency of protests across different states, which could be influenced by factors such as population size, political climate, and local issues.

Table 4.6 The cities with the most frequent protests

Cities	Protest events	Non-protest events
Portland	33	1
Richmond	28	6
Seattle	27	7
Detroit	25	9
Louisville	25	9
Oakland	25	9
Phoenix	25	9
Chicago	24	10
Columbus	24	10
Memphis	23	11

Table 4.6 shows the cities where the most frequent protests occurred, with Portland ranking first with 33 protests. The data reveals that Portland has a significantly higher number of protest events (33) than non-protest events (1), indicating a strong focus on protest activity in the city. This could be attributed to specific local issues, political climate, or a highly engaged population.

Other cities, such as Richmond, Seattle, and Detroit, also have more protest events than non-protest events. This suggests that these cities may be experiencing similar factors driving protests, such as political or social issues that resonate with the local population. In contrast, cities like Chicago, Columbus, and Memphis have a more balanced ratio of protest events to non-protest events, which could indicate a more diverse range of activities and concerns within these cities.

It is also worth noting that the cities listed in Table 4.6 are geographically diverse, representing various regions across the United States. This highlights the widespread nature of protests and the potential for different factors to drive protest activity in different cities. Further investigation into each city's specific issues and contexts could help better understand the reasons behind the high frequency of protests and inform potential strategies to address these concerns.

Social media platforms, particularly Twitter and Facebook, can be valuable tools for predicting and understanding protests in specific states and cities. By analyzing user engagement, sentiment, and content in the areas with the highest number of protest events, researchers can potentially forecast future protests and gain a deeper understanding of the factors driving these events.

4.3. User Behaviors and Protests

4.3.1. Relationship between Twitter user behaviors and protest day

Social network users show different behaviors toward social unrest. Some users do not react to this social unrest; others react to the protests and share content. Many questions are raised in this field. In this section, three main questions are discussed in this field. These questions are as follows:

Q1: How do users behave in the face of protests?

Q2: How do users use Twitter to communicate about protests?

Q3: What content do users share regarding Protests via Twitter?

The proposed method is explained below to get an accurate answer to these questions. In this regard, these questions will be answered using the initial research data set. In the data set prepared by Dr. Wong, 1,500 tweets about the OWS protests have been scrutinized and labelled by five coders. In this tagging, each tweet is labelled to the relevant behavior. For example, a tweet about public information sharing behavior has been tagged to this category. They were able to specify coding categories and continued this process until a reliability of 0.91 was obtained.

This data set has been used to answer three questions. We believe that users' behavior patterns in Occupy Wall Street protests can be obtained using this labeled dataset, which can be used to analyze other events, such as the BLM movement.

Using the Chi-square, the existence of a relationship between user behavior and the day of protest in the Wall Street movement has been evaluated. The Chi-square test is one of the most important non-parametric tests, which is indicated by X^2. Fisher introduced the chi-square test. This test can show the significance of the difference between the observed and expected frequencies obtained from a community [167].

The evaluation of independence tests through crosstabulation, also known as a bivariate table, is predominantly conducted using Chi-Square statistics. Simultaneous distribution of two classified variables can be achieved through crosstabulation, where the intersection of variable classification appears in the table cells. By comparing the observed pattern of responses in the cells with the expected pattern, the independence test examines whether there is a relationship between variables or if they are independent of each other. The researcher can assess whether the number of observed cells differs from the number of expected cells by statistically computing the chi-square test and comparing it to a critical value of the chi-square distribution.

$$\mathbf{X^2 = \sum (Fo - Fe)^2 / Fe}$$

In the formula and test of Chi-Square:

X²: Chi-square index

Fo: Frequency observed

Fe: Frequency expected

As shown in the Chi-square formula, the Chi-Square statistic predicts the difference between what is observed in the data and what is expected in the absence of a relationship between behaviors.

According to calculations, there is a significant relationship between EventLogShare behavior and the day of the protest (p-value<0.05). There is also a significant relationship between GeneralInfoShare behavior and the day of the protest (p-value<0.05). However, the relationship between EventLogShare behavior and the day of the protest is stronger. The results of using the chi-square test can be seen in Table 4.7.

Table 4.7 The relationship between the day of the protest and the behavior of Twitter users

Behaviors	χ^2	p-value
EventLogShare	123.820	4.544×10^{-28}
GeneralInfoShare	47.820	4.172×10^{-11}
OpinionShare	2.666	0.263
PublicOutreach	11.066	0.004

There is no significant relationship between OpinionShare behavior and the day of the protest (p-value>0.05). There is significant relationship between PublicOutreach behavior and the day of the protest (p-value<0.05).

According to the data collection and tagging in source and the statistical studies presented in this study, it can be concluded that identifying EventLogShare behavior among users is most important to predict the day of the protest.

4.3.2. Recognition of EventLogShare

The definitions of Twitter user behavior were discussed in Section 2.2. Sharing geographical and temporal information is one of the four behaviors of users in the protest event. Based on modeling, this behavior has a significant relationship with the initial data set and the day of the protest. In this section, temporal and geographical information will be identified. This section is done with the Python programming language.

In this thesis, a new method is proposed to identify the behavior of EventLogShare in the protest event. This method uses the Rollbase, SUTime library [168], and regular expressions to identify tweets containing EventLogShare.

Information including date, time, and location must be extracted from the tweets to identify this behavior. For this purpose, the rule-based method, the regular expression, and the SUTime library have been used (Figure 4.4).

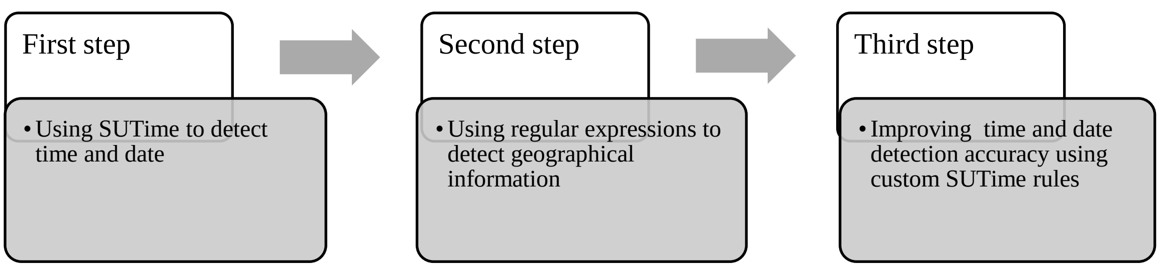


Figure 4.4 Behavior detection steps of EventLogShare

SUTime libraries are used to identify and normalize time expressions. This is how this library works: for example, next saturday at 3 pm is received as input, and 2016-02-17T15: 00 is resulted as the output according to the current reference date. As a deterministic rule-based system in the Stanford CoreNLP pipeline, the library facilitates working with temporal information and document annotations. The current ruleset only supports English, but some researchers have developed some other rulesets for other languages such as Swedish. Angel Chang is the creator of SUTime. In the Stanford JavaNLP project, other parts and classes were added to the library. TokensRegex is a generic framework for mapping semantics to objects and defining in-text patterns, which is known as the basis of SUTime. This research utilized the SUTime library in Python.

Regular expressions provide a tool for programmers to find the words, phrases, and patterns in a text string. Regular expressions have a special set of operators that can be combined to analyze and search text. The comprehensiveness of these operators allows us to search all our requests on the string.

Rule-based text classification systems are based on the linguistics rules. In this context, the linguistics rules are the rules of correspondence (or dependence) produced by humans to create a correspondence between a particular linguistic pattern and a label (or class). In other words, as soon as the system detects a particular language pattern in the data, it automatically uses the label corresponding to that pattern to label or categorize it. Once the rules for doing so have been developed in text mining methods, the system will automatically identify the various linguistic structures in the text data and label them accordingly.

The SUTime library was used to identify the time and date. The SUTime library identified the EventLogShare behavior, the results of which are presented below. Using regular expressions and data related to time information, we tried to identify the geographical information published in the tweets to identify places' names.

Table 4.8 The accuracy of behavior detection of EventLogShare

EventLogShare	Evaluation
Precision	0.63
Recall	0.82
Accuracy	0.71

According to the results obtained in Table 4.8, the accuracy of the EventLogShare detection method is 63, recall is 82, and precision is 71. Then, adding custom expressions and rules (custom SUTime rules), detecting the initial dataset's behavior improved the precision by 9%, recall by 2%, and accuracy by 7%. According to the results obtained in Table 4.9, the detection method's accuracy is 78, recall is 84, and precision is 72.

Table 4.9 The detection accuracy of EventLogShare behavior after adding custom SUTime rules

EventLogShare	Evaluation
Precision	0.72
Recall	0.84
Accuracy	0.78

4.3.3. Emotion Analysis on Tweets

In the previous Section, studying the relationship between users' behavior and protest day shows that there is a significant relationship between the protest day and two behaviors of users in social networks. Then, the presence or absence of the significant relationship between the tweets' emotions and the protest day is studied.

Using the emotions analysis, which is explained in the next paragraphs, the primary dataset was labeled. In this regard, 1500 tweets were collected and labeled in the primary dataset into four sad, happy, hate, and neutral emotions, and then, the relationship between these emotions and the protest day was studied using the chi-square test.

Table 4.10 shows the results of this analysis. There was a significant relationship between hate emotion and protest day (p-value<0.05). In this regard, the protest will likely occur if the hate emotion increases in the tweets. There is also a significant relationship between neutral tweets and protest day (p-value<0.05). Still, no meaningful relationship between sad and happy tweets and protest day was reported.

Table 4.10 The relationship between the day of the protest and the behavior of Twitter users

Variable	χ^2	p-value
Neutral	22.91	0.000011
Hate	15.23	0.000491
Happy	4.054	0.132
Sad	1.17	0.557

Emotion Analysis has been used in this thesis to predict the protest day and the probability of violence in the protests. In this Section, Emotion Analysis is discussed briefly. The subject of Emotion Analysis was raised 20 years ago, during which research has been conducted on it.

Emotion Analysis is a subset of Sentiment Analysis or opinion mining, and the definitions that have been given for them can be used for Emotion Analysis [169]. In Sentiment Analysis, classification is done based on the positive or negative emotion, but the Emotion Analysis is classified in more detail into feelings of anger, happiness, sadness, etc. [170]. Table 4.11 gives these emotions in detail.

Table 4.11 Sentiments and emotions

Sentiments	Emotions
Positive	Surprised, Angry, Disgusted, Fearful, Sad, and Love
Negative	Hate, Shy, Praiseful, Regretful, and Nervous

Sentiment Analysis (SA) is a study field that tries to express the emotions, behavior, opinions, and analysis of different persons regarding the entities and their features. This entity can include products, services, organizations, individuals, events, and subjects [171]. SA has attracted much attention since 20 years ago. One of the reasons for such attention is that SA is highly interesting to audiences. Such a broad scope of audiences led several researchers and scholars to pay attention to this subject. Researchers have worked on Sentiment Analysis in different fields, thereby studying this subject from their own perspective. They have also chosen a name for their research subject based on their perspective, causing us to confront different names for the Sentiment Analysis in the literature. These names include Emotion Analysis, Opinion Mining, Opinion Analysis, Opinion Extraction, Sentiment Mining, Subjective Mining, and many other names [172]. Such a broad scope of names made finding the related works problematic because very successful research may be done in this field, but it may not be evident to the experts due to its different names.

The first research on SA dates back to 2002. Bo pang presented a method for classification of film opinions using the supervised technique [173]. Then, Turney presented the SA method with the unsupervised technique in the same year [174].

Following them, two thousand articles and studies have been published in this field, and different methods have been presented to identify the sentiments. In recent years, deep learning algorithms have been used for this purpose. In this thesis, these algorithms have been used for Sentiment Analysis.

SA is currently used for predicting the film, results of the election, depression, stock prices, and several other cases. Sentiment Analysis can be done at word, feature, sentence, and document levels. It is mostly done at the word level. Sentiment Analysis means finding and identifying the sentiment of a user about a special subject or event. SA is done in text, video, and voice, but since this thesis is about the Sentiment Analysis in the text, SA on the text has been considered by default.

Users can express his/her emotion through text and express his/her feelings of happiness, sadness, hatred, etc., by writing an opinion about an event. This is not the only way for a user to express his/her emotions, but these emotions may be expressed through voice or video, but it can be claimed that the number of opinions that are presented through text is much higher than that of the opinions expressed by voice or video. The main focus of the Sentiment Analysis is on the texts which are produced by web users by different means. Users produce text every day through social networks or messaging devices and many other ways. Considering the place of social networks such as Twitter and messaging devices in routine life, a major part of the text is produced every day, and this is the unique chance of the Sentiment Analysis on the texts. Since the main focus of the Sentiment Analysis is on the text, the Sentiment Analysis can be regarded as one of the text mining fields. For this reason, the methods which have been applied in text mining have been widely used by researchers in Sentiment Analysis [172].

The importance of the Sentiment Analysis can be understood through a high number of studies in this field. Today, several studies are conducted in this field. The reason for such attention to this field is the increasing importance of the users' opinions. We all are curious to know others' opinions, and we tend to know what opinions and beliefs they follow in different parts of the world. From the commercial viewpoint, before shopping, we tend to know the emotion of those who have purchased the goods before us. Sentiment Analysis has become more important after the emergence of social networks because the emergence of social networks allowed users to express their opinions in all fields as easily as possible. For example, on Twitter, you can comment on it using the hashtag, and this comment can be seen by those who follow this hashtag. The Sentiment Analysis in social networks also caused growth and increasing attention to the Sentiment Analysis.

One of the ambiguities which are raised in the Sentiment Analysis is the difference of opinions or sentiment. Is there such a difference? If yes, what does this difference mean?

In this research, the relationship between the protest day and opinion sharing behavior was not significant (Table 4.7). A significant relationship was found in the emotion on the tweets (Table 4.10). This difference can be related to the difference in sentiment and opinion.

In the dictionary, sentiment means a behavior, thought, or judgment created through sense, while opinion means an attitude, judgment, and analysis made in mind. Based on the definition, it seems that the concept of "opinion" is related to a person's specified attitude toward an entity originating out of his/her mind, while sentiment is more related to a sense and shows a sense of an individual about an entity. Therefore, it can be concluded that the concept of "opinion" is more based on the mind than the concept of "sentiment," and the concept of "sentiment" is more based on sense than the concept of "opinion." For more clarification, note the following example. In the following example, there are two sentences containing opinion and sentiment.

"Today's protests in Taksim Square are worrisome. I fear that a bad event happens! "

"I think that if the new minister comes to power, the protests in Hong Kong will be exacerbated."

The first sentence shows a sentiment about an event that the speaker considers, and the second sentence expresses the opinion which results from the speaker's thought. The audience who hears these two sentences recognizes the difference between these two sentences because he/she responds to the first sentence: "I am worried too." He/she responds to the second sentence: "I also agree with you." These two responses express the difference between opinion and sentiment. Besides these differences, there are deep similarities between opinion and sentiment. Opinion and sentiment can be regarded as an entity and interrelated ,considering a partial difference.

Different opinions in SA are classified into twofold, threefold, and seldom fourfold groups. In the SA, opinions are mostly classified into positive and negative groups, but in the Emotion Analysis, there are 13 cases. Most of the works performed on the SA have focused on two positive and negative groups.

4.4. The Proposed Prediction Method Based on User Behaviors

The protest event detection based on twitter's user behaviors consisted of five steps: (1) data preparation, (2) classification of raw data, (3) identify user behaviors, (4) setup day-location pairs and features, and (5) evaluation of Bayesian Logistic Regression (BLR) algorithm and report prediction results and indicators (Figure 4.5). An overview of the proposed method is shown in Figure 4.5 This figure shows how to predict protest event and protest type based on user's behavior.

In the first step, the required data are extracted from the Twitter social network. In the second step, the extracted data are saved in JSON format. JSON's files are heavy and intertwined and contained much information. In this step, the extra information is deleted, and then the data are pre-processed and cleaned. Information on the day and place of the protest is extracted and stored.

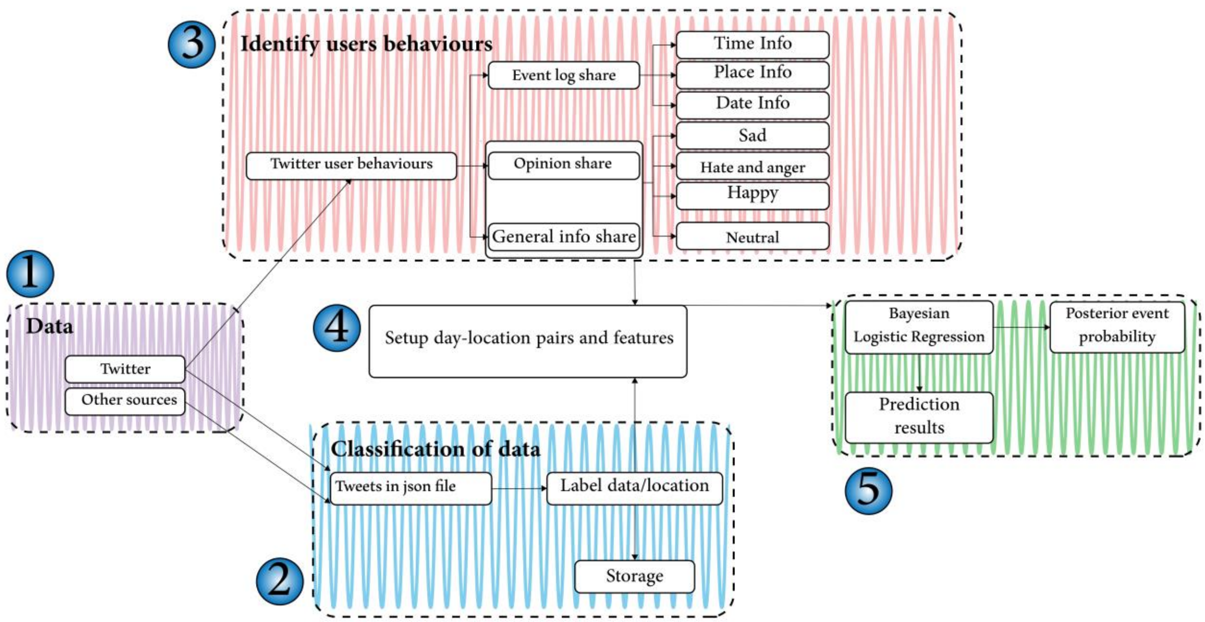


Figure 4.5. Overview of the presented method

Raw tweets obtained from Twitter are processed in the third step, and the Twitter users' behaviors are identified. User behavior can be classified into three groups. Wang et al. in their paper ranked the users' behaviors on social networks in events such as Occupy Wall Street [9]. The paper classified users' behaviors in social networks during protests, claiming that the users' behaviors in social networks during the Wall Street protests are as follows: Event log sharing (Event Log Share), sharing general information about the event (General Info Share), sharing opinion about the event (Opinion Share), and call for action. Wang et al. believe that when an event such as a social protest occurs, users try to show one of the above-mentioned social behaviors on social networks.

Event log sharing (ELS): The event log sharing behavior of the users is widely observed in a social protest. One of the users' main behaviors on Twitter in street protests is that users try to share spatial, temporal, and logistic event logs. Through this content creation; users try to share the event log information. Three groups of tweets describe this behavior:

- Time information: Tweets containing time information such as protest hour.[2]
- Date information: Tweets containing date information.[3]
- Spatial information: Tweets containing information such as the name of the street or square where the protest occurs.[4]

We use Stanford NLP's SUTime to identify tweets with time and date information[175]. A list of places collected from the Internet is used to identify tweets with spatial information.

Opinion Share (OS): For each event, users share opinions about that event. These opinions can be positive[5] and negative(sadness and hatred[6]) about the event. The third case is the sharing of neutral views about the event.

General Information Share (GIS): General information sharing is another behavior expressed by users of social networks, particularly Twitter, in protest events. The main aim of these tweets is to share general information about the events. These tweets post the news of the event and have a neutral tone[7].

In order to determine whether tweets contain OS and GIS behaviors, a deep learning classification algorithm is used. A combined deep learning algorithm is used to identify OS behaviors and GIS behavior. Three groups of emotions that are considered for OS behavior are happiness, sadness and anger-hatred. Neutral emotions are also considered for GIS behavior. We use "multi-channel"

---

[2] Example: RT @angelxxelyse: A rally/protest is scheduled Friday (tomorrow) 5 p.m. at Peter's Park in the South End Boston. Please consider attending…

[3] Example: RT @HeadOverFeels: On June 6th join the #DoctorWhoBlackout! We're supporting #BlackLivesMatter with this livetweet of THE GHOST MONUMENT a…,https://t.co/6zqPO3gIve

[4] Example: RT @MrAndyNgo: Antifa groups in Portland, Ore. have announced a 6 p.m. gathering at Laurelhurst Park. This is a middle class residential ne…

[5] Example: RT @HealthworksFit: #BlackLivesMatter. We stand in solidarity with the Black community near and far against systemic racism and injustice.…,https://t.co/cdedP2V4eP

[6] Example: Don't let these mothafuckas fool you https://t.co/DoUW6SV9Tb

[7] Example: RT @JoshuaPotash: This is the Brooklyn Bridge right now!! #BlackLivesMatter https://t.co/8qgC2afTkh,https://t.co/Ct7uHeRic8

combinations of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) units to classify tweets into one of the four emotional classes (sad, happy, hate-anger, and neutral) [161]. Other categories such as linguistic features and public outreach ignored in this paper. These are rare tweets.

In the fourth step, the day and place pairs that are obtained in the second step the extracted features that are obtained in the third step are combined to create the dataset whose structure is given in Table 4.4.

In the fifth step, Bayesian Logistic Regression (BLR) algorithm has been used to calculate the accuracy of the prediction and the indicator identification of events. It is vital to identify the features that should be monitored to identify the protest and the probability of violence. The prediction results show the model's accuracy, and the posterior probability determines the best indicator for monitoring events.

We define a Bernoulli random number for the probability of protest occurrence (Y_{ij}) [176]. Y_{ij} is the probability of a protest event occurring in specific day(i) and city(j). If Y_{ij} is equal to 1 the protest event will occur ,and it will not happen if Y_{ij} is equal to 0. This is also applicable for the protest type prediction.

$$\Pr(Y_{ij} = 1) = p = 1 - \Pr(Y_{ij} = 0) = 1 - q \qquad (1)$$

The probability mass function f of Bernoulli distribution over possible n outcomes can be expressed as:

$$\mathrm{f}(n; p) = p * n + (1 - p)(1 - n) \quad for\ \ n \in 0{,}1 \qquad (2)$$

In the Bernoulli distribution, which is a discrete distribution, there are two possible outcomes: n=0 and n=1. The occurrence of n=1 ("protest event") has a probability of p, while n=0 ("no event") has a probability, $q = 1 - p$ where $0 < p < 1$.

By leveraging prior knowledge, the Bayesian approach can achieve robustness with high efficacy [177]. The algorithm can achieve a better fit of the posterior distribution of the objective function by increasing the number of samples [178].

An event X is described by a set of features. Based on Bayes theorem, the conditional probability whether a given event is a protest event or not can be computed as follows:

$$P\left(Y_{ij}\middle|X\right) = \frac{P\left(X\middle|Y_{ij}\right)P(Y_{ij})}{P(X)} \tag{3}$$

In what follows, we use the logistic link function

$$P\left(Y_{ij}\middle|X\right) = \frac{e^{(\alpha+\beta_i x_i)}}{1+e^{(\alpha+\beta_i x_i)}} \tag{4}$$

Which has the probability mass function

$$P\left(Y_{ij}\middle|X\right) = \frac{1}{1+e^{-(\alpha+\beta_i x_i)}} = \frac{1}{1+e^{-(logit)}} \tag{5}$$

$$P\left(Y_{ij}\middle|X\right) = \frac{1}{1+e^{-(\alpha+\beta_0 x_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\beta_4 x_4+\beta_5 x_5+\beta_6 x_6+\beta_7 x_7+\beta_8 x_8)}} \tag{6}$$

The interpretation formula is as follows:

$$Logit = \alpha + \beta_0 * (Number\ Of\ Tweet) + \beta_1 * (Day) + \beta_2 * (Date\ Info) + \beta_3 * (Time\ Info) + \beta_4 * (Place\ Info) + \beta_5 * (Sad) + \beta_6 * (Happy) + \beta_7 * (Hate\ and\ anger) + \beta_8 * (Neutral) \tag{7}$$

In the Bernoulli distribution, the theta parameter is important and it is calculated using alpha and beta parameters as follows.

$$\theta = \ logistic(\alpha + \beta_i x_i) \tag{8}$$

Alpha and beta parameters with normal distribution used to estimate the theta parameter to generate the BLR algorithm. Based on Bayesian logic, the numbers in the data have been randomly selected, starting with the stated parameters. Bayes uses a distribution to express these numbers.

The BLR algorithm is employed to identify protest indicators and violence indicators. Also the prediction model is created by this algorithm.

5. EXPERIMENTAL RESULTS

This chapter presents the results obtained from implementing the proposed method for predicting protests and violence using user behavior on Twitter. Two events were identified and predicted: protest events and violent events. Results for both events are reported in the 5.1 and 5.2 sections. To find the best combination of user behavior, various combinations were evaluated to achieve the best results. Therefore, at the beginning of the chapter, different combinations of user behavior were defined, and based on them, five groups were formed on which the method was implemented. The results obtained from different combinations of user behavior are presented in tables and charts. The following sections of this chapter explain evaluation parameters, user behavior combinations, and the obtained results in detail.

Evaluation of machine learning performance and data mining methods involves using various metrics, making it challenging to compare different approaches presented in various articles. While some papers only use one evaluation criterion, such as accuracy, others may utilize the F1-score metric. In predictive induction based on classification, classification accuracy is typically adopted as a standard criterion, though other measures such as F1-score, Precision, Recall, Sensitivity, and Specificity are also used. In fields related to knowledge discovery and data mining, such as association rules, descriptive induction methods, confidence, support, comprehensibility, and interestingness are utilized as alternative evaluation criteria. Table 5.1 shows variables of confusion matrix.

Table 5.1 Variables of Confusion Matrix

Predicted Label	Real Label		
		Positive	Negative
	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

To explain the evaluation criteria of the proposed method, four variables need to be defined. A true positive (TP) refers to an outcome when the model accurately predicts the occurrence of a protest event. A true negative (TN) represents an outcome when a correct prediction of the absence of a protest event is made by the model. On the other hand, a false positive (FP) occurs when the model incorrectly predicts a protest event, while a false negative (FN) arises when the model inaccurately predicts the absence of a protest event.

$$\text{Precision} = \frac{\sum \text{TP}}{\sum \text{TP+FP}} \tag{9}$$

$$\text{Recall} = \frac{\sum \text{TP}}{\sum \text{TP+FN}} \tag{10}$$

$$\text{Accuracy} = \frac{\sum \text{TP+TN}}{\sum \text{TP+FP+FN+TN}} \tag{11}$$

In cases where the cost of false positives and false negatives varies, the F1-score criterion can be utilized as it provides a balanced combination of accuracy and precision measures.

$$\text{F1} - \text{score} = \frac{2\ (\text{Precision} * \text{Recall})}{\text{Precision+Recall}} \tag{12}$$

The presented evaluation criteria have been applied to assess the accuracy of the protest event detection based on twitter’s user behaviors. The subsequent sections involve evaluating classification models using metrics derived from the four outcomes mentioned earlier.

A combination of features is considered to achieve the best protest prediction model. We considered five features combination to protest and violent event detection based on twitter's user behaviors, as follows:

EventLogShare only: This combination only uses the EventLogShare (Time Info, Place Info, and Date Info) features (**Group I**).

OpinionShare only: This combination only uses the OpinionShare (sad, happy, hate-anger) features (**Group II**).

GeneralInforShare only: This combination only uses the GeneralInforShare (neutral) feature (**Group III**).

EventLogShare + OpinionShare + GeneralInforShare: This combination uses all user behavior features of tweets in Twitter (**Group IV**).

EventLogShare + OpinionShare + GeneralInforShare + Number of Tweets + Day: This combination uses all user behaviors on Twitter plus the number of indicative tweets and the protest day information (**Group V**).

Each features combination is used to predict the protest and the possibility of violence in the protest. The 5.1 and 5.2 sections reports the obtained accuracy precision, recall, and F1-score results.

A Spearman's rank-order correlation method is used to examine the strength of the relationship between features and outcomes (event type and protest type). Besides user behavior features, the month and day of the protests have also been collected. Because there is information about only two months in our data, only the day parameter has been considered. In addition, the number of tweets per day has been used as a feature in combinations. Correlation results for features are also presented in the next section.

In 5.1 section, users' behavior was used to predict the protest day. In the 5.2 section, the prediction of violent protests was considered. The results of both predictions were presented in detail in different tables.

5.1. Protest events

We first tested the strength of the relationship between features (the protest day, the number of tweets, Date Info, Time Info, Place Info, sad, happy, hate-anger, and neutral) and outcomes (event type and protest type). A Spearman's rank-order correlation was run to determine the relationship between features and outcomes. There was a moderate, positive correlation between Time Info and protest events, which is statistically significant (r = 0.52, p-value = 6.87×10^{-261}). There is a moderate, positive correlation between Date Info and protest events, which is statistically significant (r = 0.44, p-value = 2.77×10^{-195}). There is a weak, positive correlation between the protest day features and protest events. There is a weak, positive correlation between the number of tweets and protest events.

Table 5.2 presents the performance results of features combination for protest event classification using different feature groups. The table shows the accuracy, precision, recall, and F1-score for each feature group. The table contains five feature groups (Group I to Group V) and their respective performance metrics. The table demonstrates how different feature groups affect the performance of the proposed method in terms of accuracy, precision, recall, and F1-score.

According to Table 5.2, the best combination of features belongs to user behavior besides the number of informative tweets and the protest day. In this combination, the accuracy is 92%, the precision is 85%, the recall is 87%, and the F1 is 86%. The performance of the triple users' behavior combination is very close to that best features combination. EventLogShare behavior has been more successful in predicting the protest day among the users' behavior.

The effectiveness of the features of the Group I to predict protest events in the proposed method is demonstrated below: accuracy= 0.88; precision = 0.84; recall = 0.86; f1-score = 0.85. The performance of combined features to predict protest events in Group I is higher than Group II and III, but it is lower than Group IV and V.

The performance of the features of the Group II to predict protest events in the proposed method is as follows: accuracy= 0.55; precision = 0.68; recall = 0.52; F1-score = 0.60. The performance of OpinionShare feature to predict protest events is lowest. Using OpinionShare feature without combining it with other features provides poor performance for predicting protest events.

The performance of the features of the Group III to predict protest events in the proposed method is as follows: accuracy= 0.69; precision = 0.50; recall = 0.35; F1-score = 0.41. Using GeneralInforShare behavior alone also provides low performance in detecting protest events. This feature outperformed only behavior OpinionShare and provided lower performance than other features combinations.

The performance of the features of the Group IV to predict protest events in the proposed method is demonstrated below: accuracy= 0.91; precision = 0.83; recall = 0.89; F1-score = 0.86. The combination of the three behaviors (EventLogShare, OpinionShare, and GeneralInforShare) provides acceptable performance for predicting protests. This combination of behaviors after Group V is the best combination for predicting protest events.

The effectiveness of the features of the Group V to predict protest events in the proposed method is as follows: accuracy= 0.92; precision = 0.85; recall = 0.87; F1-score = 0.86. The best performance for predicting protests is a combination of five features (EventLogShare, OpinionShare, GeneralInforShare, Number of Tweets, and Day). In this combination, the highest performance is obtained.

Table 5.2 Protest event classification accuracy, precision, recall, and F1 score

Features Groups	Accuracy	Precision	Recall	F1-score
Group I	0.88	0.84	0.86	0.85
Group II	0.55	0.68	0.52	0.60
Group III	0.69	0.50	0.35	0.41
Group IV	0.91	0.83	0.89	0.86
Group V	**0.92**	**0.85**	**0.87**	**0.86**

In this table, Group V has the highest Accuracy, Precision, Recall, and F1-score, making it the best-performing group in terms of protest event classification. Group IV also performs well, with only slightly lower values than Group V. Group II and Group III have lower performance metrics, indicating that they may not be as effective in classifying protest events. Figure 5.1 shows Protest events classification accuracy, precision, recall, and F1 score.

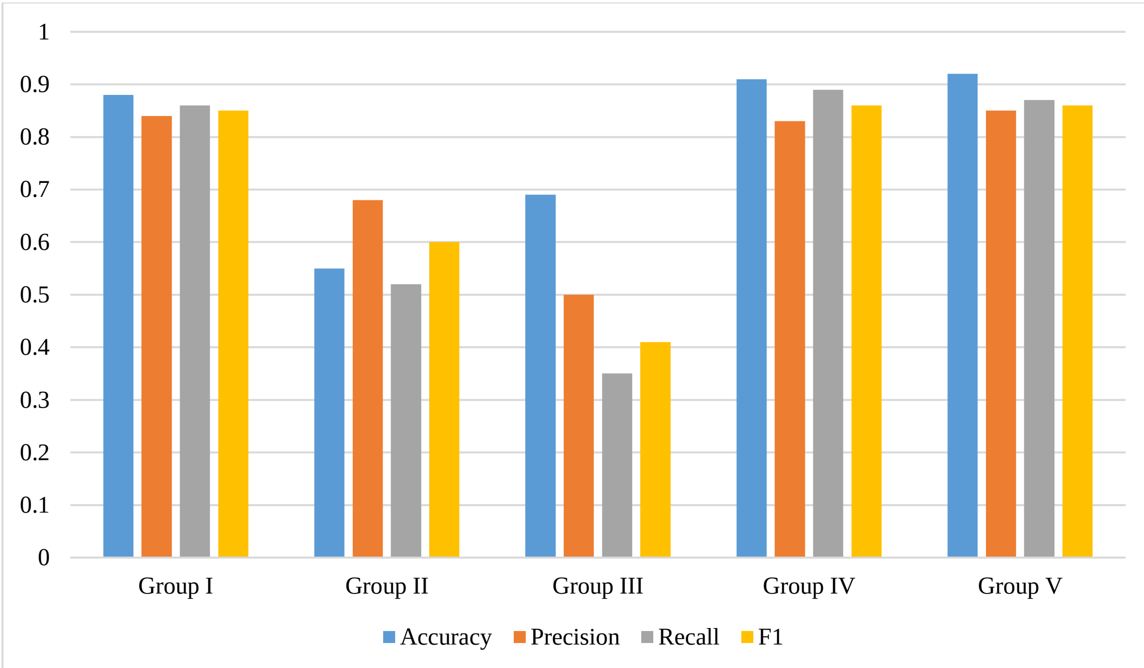


Figure 5.1. Protest events classification accuracy, precision, recall, and F1 score

Table 5.3 shows the classification accuracy, precision, recall, and F1 score of the best features combination to predict protest events.

Table 5.3 provides a more detailed view of the best features combination (Group V) for protest event detection. It shows the precision, recall, F1-score, and accuracy for both event types (No Events and Protest). The features combination achieved an accuracy of 92% for both "No Events" and "Protest" event types. For "No Events," the features combination had a precision of 93%, recall of 90%, and F1-score of 91%. For "Protest" events, the features combination had a precision of 77%, recall of 83%, and F1-score of 80%. The table contains the performance metrics for the best features combination, which uses EventLogShare, OpinionShare, GeneralInforShare, Number of Tweets, and Day as features.

Table 5.3 A more detailed table of the best features combination results in a protest event

The best features combination	Event Type	Precision	Recall	F1	Accuracy
EventLogShare + OpinionShare + GeneralInforShare + Number Of Tweet + Day	No Events	93	90	91	92
	Protest	77	83	80	

The best features combination achieves an accuracy of 92% in predicting protest events. For No Events, the features combination has a precision of 93%, recall of 90%, and F1-score of 91%. For Protest events, the features combination has a precision of 77%, recall of 83%, and F1-score of 80%.

Figure 5.2 shows no protest vs. protest (y = 0, y = 1). The S-shaped line is the mean value of theta parameter (θ). This line can be interpreted as the probability of a protest, given that we know the Time Info (a) and Date Info (b) tweets ratio. The boundary decision is represented as a vertical line. According to the boundary decision, the values of the TimeInfo tweets ratio to the left correspond to y = 0 (no protest), and the values to the right to y = 1 (protest). X vector has theta values that were described in equation 8. The values of α and β are respectively equal to 0.65 and 0.87. The ratio of Time Info tweets and Date Info tweets on protest and no protest day shows with orange and blue color, respectively (Figure 5.2 and Figure 5.3).

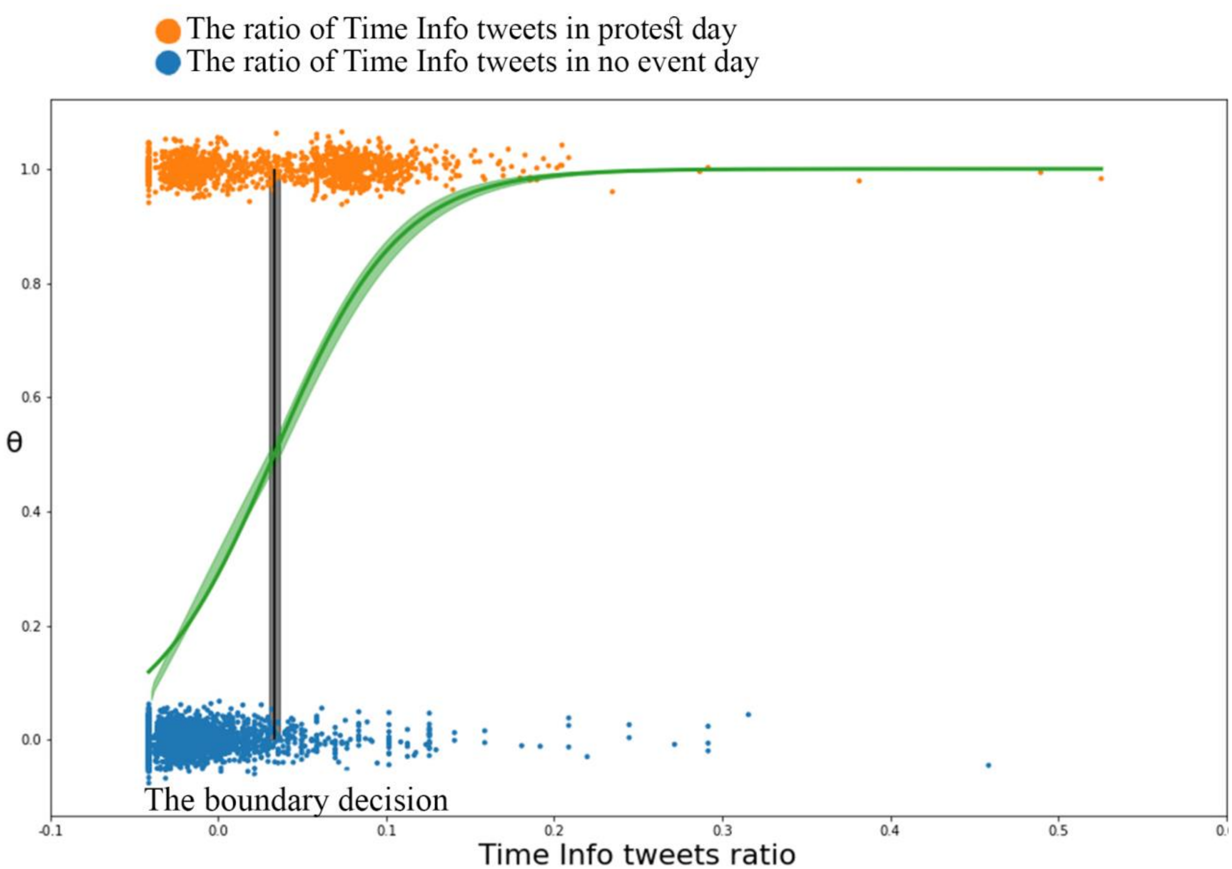
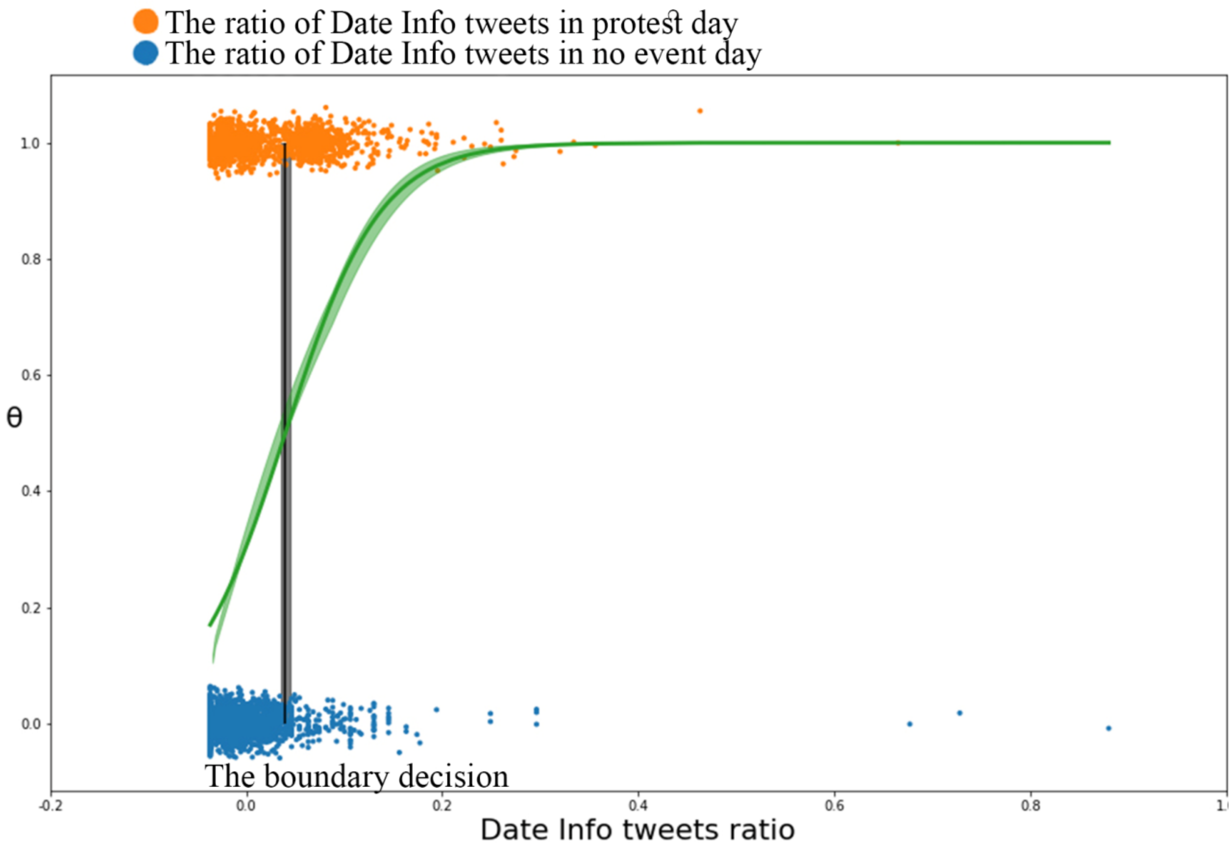


Figure 5.2. The fitted sigmoid curve and the decision boundary for Time Info



(b)

Figure 5.3. The fitted sigmoid curve and the decision boundary for Date Info

The best combination of features is presented in Table 5.3. The best combination is used to identify the cut-off point for the ratio of Time Info tweets and Date Info tweets. The time and date tweets' cut-off points are 0.04 and 0.06, respectively. The rate of Date Info tweets greater than 6% of all tweets and Time Info tweets greater than 4% are crucial indicators for forecasting protests. Each point in Figure 5.2, Figure 5.3 and Figure 5.4 shows the event in the city and the day. The total number of events is 4492, which includes 3078 no protest events and 1414 protest events, which are displayed in Figure Figure 5.2 and Figure 5.3 with 4492 points (1414 orange points and 3078 blue points).

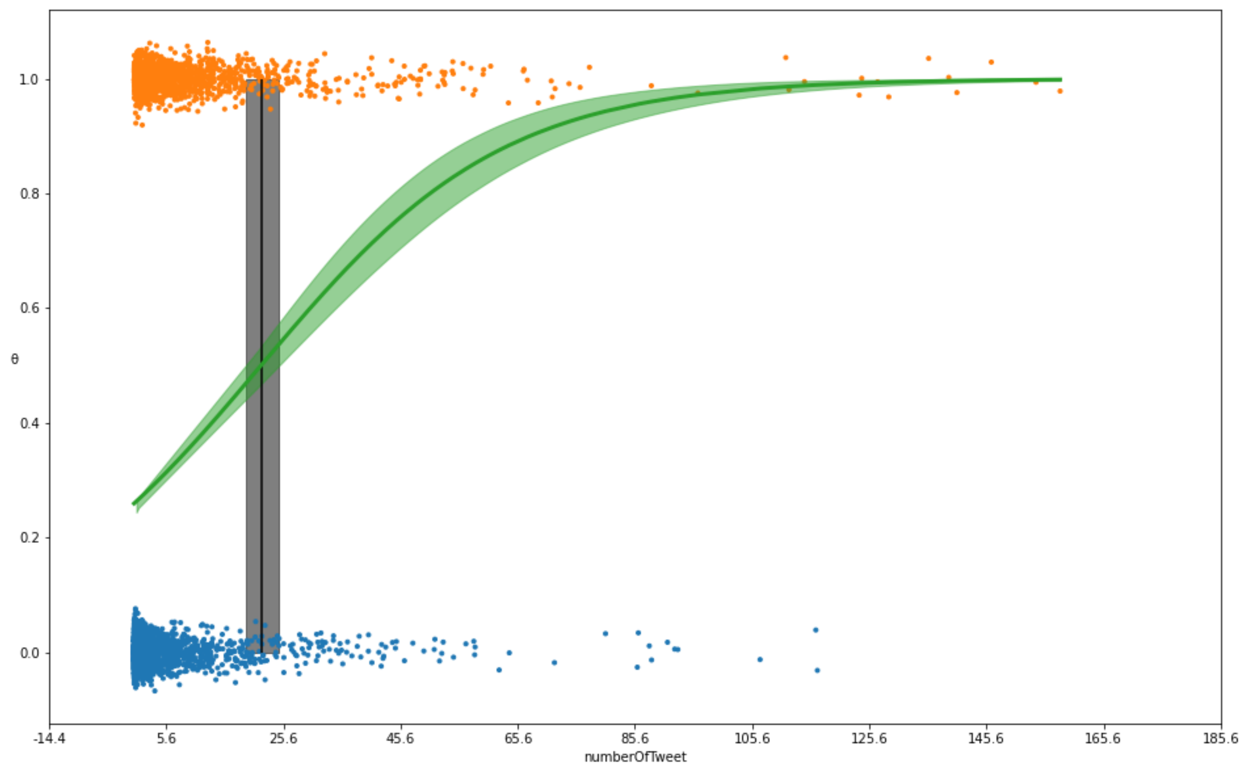


Figure 5.4. The fitted sigmoid curve and the decision boundary for number of tweet

Figure 5.4 shows the fitted sigmoid curve and the decision boundary for number of tweet. The number of tweets on protest and no protest day shows with orange and blue color.

5.2. Protest types

Like the protest event, the violent event has also been studied similarly. A strong, positive correlation was found between the hate-anger tweets ratio and violent events, which was statistically significant (r = 0.62, p-value = 1.37×10^{-42}). A negative correlation of moderate strength was detected between the sad tweets ratio and violent events, which was statistically significant (r =-0.56, p-value = 2.5×10^{-36}). Also, there was a moderate, negative correlation between happy tweets ratio and violent events.

Table 5.4 presents the performance results of the features combinations for protest type classification using different feature groups. The table shows the accuracy, precision, recall, and F1-score for each feature group. The table contains five feature groups (Group I to Group V) and their respective performance metrics. The table demonstrates how different feature groups affect the performance of the protest event detection based on twitter's user behaviors in terms of accuracy, precision, recall, and F1-score.

Based on the results obtained from Table 5.4, the best combination of features belongs to user behavior besides the number of informative tweets and the protest day. In this combination, the accuracy is 92%, the precision is 85%, the recall is 87%, and the F1 is 86%. The performance of the triple users' behavior combination is very close to best combination. Among users' behaviors, OpinionShare behavior has been more successful in predicting violence.

The performance of the features of the Group I to predict violent events in the proposed method is demonstrated below: accuracy= 0.79; precision = 0.50; recall = 0.50; F1-score = 0.45. The performance of combined features to predict violent events in Group I is higher than Group III, but it is lower than Group II, IV and V.

The performance of the features of the Group II to predict violent events in the proposed method is demonstrated below: accuracy= 0.90; precision = 0.82; recall = 0.87; F1-score = 0.84. OpinionShare's behavior alone provides favorable performance in detecting violent events. Contrary to the identification of protests, by identifying the behavior of OpinionShare users on Twitter, it is possible to identify the emotions of users, and by monitoring hate-anger tweets using sentiment analysis, violent events can be predicted with high accuracy. It is obvious that monitoring the OpinionShare behavior of Twitter users is very crucial in detecting violent events.

The performance of the features of the Group III to predict violent events in the proposed method is demonstrated below: accuracy= 0.71; precision = 0.50; recall = 0.35; F1-score = 0.41. GeneralInforShare's behavior alone provides the worst performance in detecting violent events. Spatial and time information does not perform as well as they were in the protest detection in detecting violent events.

The performance of the features of the Group IV to predict violent events in the proposed method is demonstrated below: accuracy= 0.91; precision = 0.83; recall = 0.89; F1-score = 0.86. The combination of the three behaviors (EventLogShare, OpinionShare, and GeneralInforShare) provides sufficient performance for forecasting violent events. This combination of behaviors after Group V is the most acceptable combination for predicting protest events.

The performance of the features of the Group V to predict violent events in the proposed method is as follows: accuracy= 0.92; precision = 0.85; recall = 0.91; F1-score = 0.87. The best performance for predicting violent events is a combination of five features (EventLogShare, OpinionShare, GeneralInforShare, Number of Tweets, and Day). In this combination, the most elevated performance is obtained.

Table 5.4 Protest types classification accuracy, precision, recall, and F1 score

Features Groups	Accuracy	Precision	Recall	F1-score
Group I	0.79	0.50	0.50	0.45
Group II	0.90	0.82	0.87	0.84
Group III	0.71	0.50	0.50	0.41
Group IV	0.91	0.83	0.89	0.86
Group V	**0.92**	**0.85**	**0.91**	**0.87**

Table 5.5 shows the classification accuracy, precision, recall, and F1 score of the best features combination for predicting violence in protests.

In this table, Group V again has the highest Accuracy, Precision, Recall, and F1-score, making it the best-performing group in terms of protest types classification. Group IV is also a strong performer, with only slightly lower values than Group V. Group I and Group III have lower performance metrics, indicating that they may not be as effective in classifying protest types.

Table 5.5 provides a more detailed view of the best features combination (Group V) for protest type classification. It shows the precision, recall, F1-score, and accuracy for both protest types (Peaceful and Violent). The features combination achieved an accuracy of 92% for both "Peaceful" and "Violent" protest types. The table contains the performance metrics for the best features combination, which uses EventLogShare, OpinionShare, GeneralInforShare, Number of Tweets, and Day as features. For "Peaceful" protests, the features combination had a precision of 0.97, recall of 0.94, and F1-score of 0.95%. For "Violent" protests, the features combination had a precision of 0.73, recall of 0.87, and F1-score of 0.79. The values of α and β are respectively equal to 0.71 and 0.79.

Table 5.5 A more detailed table of the best features combination results in protests type

The best features combination	Protest Type	Precision	Recall	F1	Accuracy
EventLogShare + OpinionShare + GeneralInforShare + Number Of Tweet + Day	Peaceful	0.97	0.94	0.95	92
	Violent	0.73	0.87	0.79	

Figure 5.5 shows violent events classification accuracy, precision, recall, and F1 score. Table 5.4 information was presented in this figure.

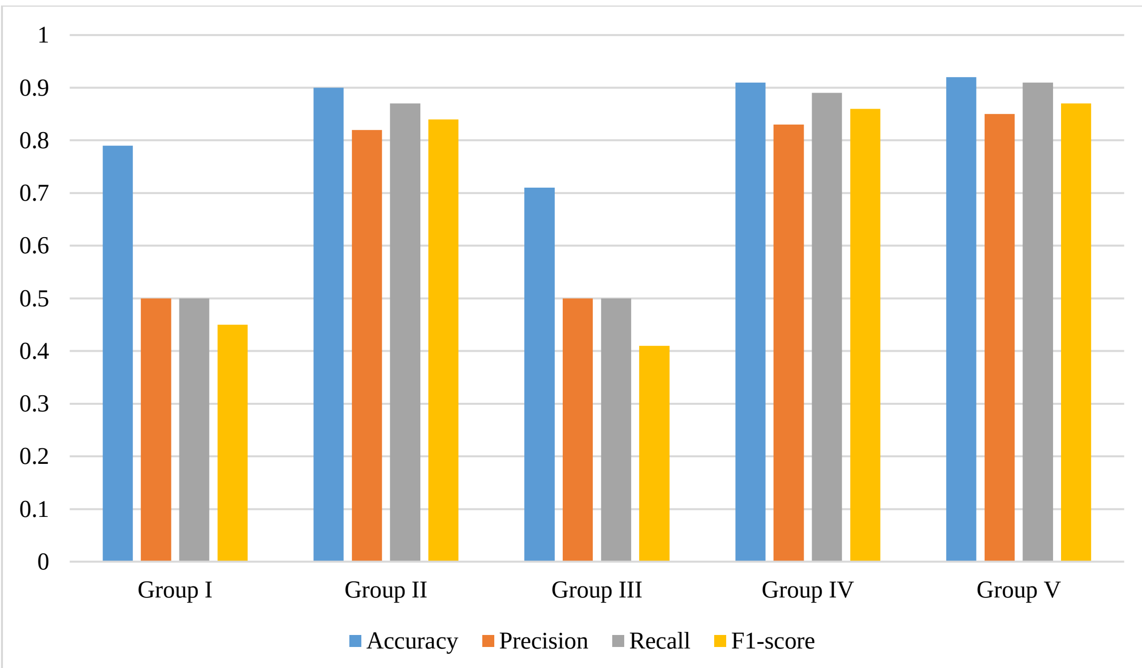


Figure 5.5. Violent events classification accuracy, precision, recall, and F1 score

Based on the boundary decision (vertical line), the values of hate-anger tweets ratio to the left correspond to y = 0 (no violence), and the values to the right to y = 1 (violence). Based on the extracted features, the proposed algorithm is implemented on five features combinations, and the best combination for predicting protests and violence in protests is obtained. The triple Twitter users' behaviors besides the day and the number of informative tweets to predict the day of protest and violence provide the best features combination.

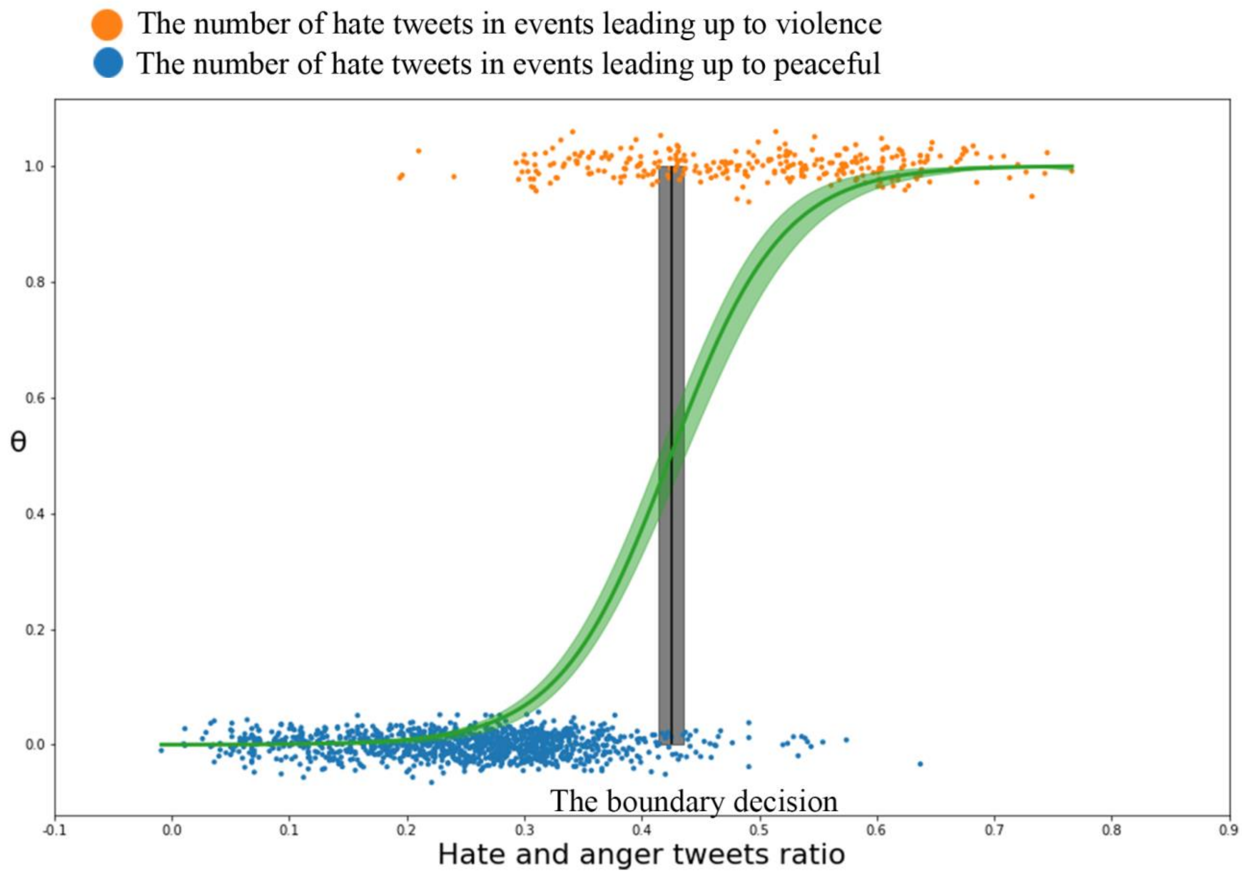


Figure 5.6. The fitted sigmoid curve and the decision boundary for hate-anger tweets ratio

The best combination is used to identify the cut-off point for the ratio of hate-anger tweets. The hate-anger tweets' cut-off points are 0.42. A rate of hate-anger tweets greater than 42% is an essential indicator for predicting the violence in protests. The ratio of hate-anger tweets on violence and no violence shows with orange and blue colors, respectively (Figure 5.6).

In conclusion, Group V consistently performs the best in both protest event classification and protest types classification, with the highest values for Accuracy, Precision, Recall, and F1-score, which uses all user behavior features on Twitter, including the number of indicative tweets and the protest day information. This features combination achieved the highest accuracy, precision, recall, and F1-score for both tasks. Group IV is a close second in terms of performance. These results suggest that the features combinations in Group V and Group IV are more effective in classifying protests and protest types compared to the other groups. The results suggest that incorporating a combination of user behavior features, such as EventLogShare, OpinionShare, GeneralInforShare, the number of tweets, and the day of the protest, can lead to more accurate and reliable predictions of protest events and their potential for violence.

6. DISCUSSION, CONCLUSION AND THOUGHTS FOR THE FUTURE

Today, with the free circulation of information and the increasing access of citizens to political information, political actions in various forms have risen, such as activities in citizen action groups, protests, and boycotts [12]. Each of these political actions can be accompanied by opinions in the online space, and the analysis of these opinions expressed by citizens in social networks and other sources can be of great value to researchers and decision-makers [20]. Content analysis is a field of study that automatically develops categories for the content people produce on social networks and assigns texts to categories. We believe that in PEA, the identification of early indicators is more important than the accuracy of classification to categories. Based on the finding of the BLR algorithm, ELS behavior is very important for predicting protest day. The rate of Time Info and Date Info tweets are a powerful indicator for predicting the protest day. Regarding violent events, OS behavior is also very useful. It is possible to predict the violence in protests by monitoring the rate of hate-anger tweets.

While previous research has reported high levels of accuracy in identifying protest events, the performance of this dissertation stands out for several reasons. First, our study achieved an impressive accuracy rate of 92%, which is at the higher end of the range reported by other papers in this field. This suggests that our approach was particularly effective in identifying protest events.

Moreover, our study also achieved a precision rate of 85%, which further indicates the reliability and effectiveness of our methodology. Precision is an important metric in protest identification studies because it measures the proportion of correctly identified protest events among all events identified as protests. A high precision rate suggests that our algorithm was able to accurately distinguish between protest and non-protest events.

Another factor that sets our study apart is the high number of protest events included in our dataset. By analyzing a large number of protests, our study was able to identify patterns and trends that may not have been apparent in smaller datasets. This adds to the overall reliability and validity of our findings.

Finally, one of the most significant contributions of our study is the identification of early indicators of protest activity. By identifying factors that are associated with the likelihood of future protests and violent in them, our study can help policymakers and law enforcement officials take proactive steps to prevent or manage potential unrest. This is a crucial contribution to the field of protest and violent identification research and has important implications for public safety and social stability.

Overall, the combination of high accuracy and precision rates, a large dataset, and the identification of early indicators make our dissertation an important contribution to our understanding of protest activity and how it can be effectively identified and managed.

6.1. Online Protest Event Detection

The content of tweets published by those who want to participate in protests is essential in real-time and has a different essence than the data of text classification, opinion mining, recommender systems, etc. For this reason, we believe that identifying early indicators in the studies related to the prediction of the day of protest and violence in them is more important than the prediction accuracy as the success criteria of "hard" classification problems.

There are three approaches: online event detection, future event detection, retrospective event, in PEA. In protest events, online analysis of the event is fundamental. Identifying the protest after it happens is not worth much. What is essential is to identify the protest before it happens. In previous works, the retrospective approach for PEA has been used. This research is important for examining the content of social networks and investigating the process of protest formation and the role of social networks in it. However, analyzing a protest event online and predicting its occurrence before it happens is much more critical and valuable. Identifying indicators for the occurrence of a protest is a way to determine the occurrence before it occurs.

Interpretable research results are early indicators for system monitoring experts such as police staff. The value of the boundary decision is specified in the results (Time Info tweets rate = 0.04, Date Info tweets rate = 0.06, and hate-anger tweets rate = 0.42). The system monitoring experts can warn of upcoming protests and violence by monitoring this value. By monitoring hate-anger tweets, the police can learn about the increased likelihood of violence in protests. Preventing violence in protests by using the interpretable results extracted from the content of tweets published on Twitter is one of the most important contributions of this study. Predicting violence in protests can prevent many financial and human costs.

Korkmaz et al. [24] have mentioned economic reasons in Latin America as an early indicator of protests. Tuke et al. [39] have determined the role of weekdays and months, besides the number of informative tweets in a tropical country such as Australia, as an early indicator. Bakerman et al. [5] have identified specific keywords as the most important prediction indicator of protest day. Other works with machine learning algorithms and text mining approaches have been presented for predicting protests and have not presented an early indicator [1, 2, 4, 14, 18, 22, 33, 42].

6.2. The Open Dataset Problem

The most crucial problem in protest identification studies is the lack of an open dataset. Because social network policies do not allow user content publication, there is no comprehensive data set. Many studies have unique datasets and have been collected by researchers. Thus, it is not easy to make comparisons between these studies. Most other papers reported an accuracy between 75% and 95%. The performance of this dissertation includes accuracy (92%) and precision (85%) are more reliable and valuable because of the high number of protest events in the dataset and the identification of early indicators.

The lack of a common dataset has made the comparison difficult, however, by examining previous works, it is clear that the method presented in this thesis provided acceptable accuracy. Besides, one of the most important contributions of this thesis is to provide warning indicators about the occurrence of protest and violence, which has been less addressed in previous works.

6.3. Small and large-scale protests

In previous studies, small and large-scale protests have not been examined simultaneously. Some studies studied large-scale protests [22, 28, 34] and some small-scale protests [38]. In the present study, small and large-scale protests were investigated, and the presented method was successful in both. One of the essential advantages of this research is providing a method for examining small and large-scale protests.

Unlike previous works focused on small-scale or large-scale protests, small-scale and large-scale protests were analyzed in this study simultaneously. This research detects both small and large scales protests successfully. Interpretable results are invaluable to the system expert. The indicators can be critical in predicting protests and violence.

The study conducted a comprehensive analysis of both small-scale and large-scale protests, which is a departure from previous works that focused solely on either one. This approach allowed for a more nuanced understanding of the dynamics and factors that contribute to protests of different magnitudes. By detecting both small and large-scale protests successfully, the study offers valuable insights into how protests can escalate in size and intensity, and how they can be effectively managed.

6.4. Limitations of study

In our study, we encountered limitations because there was no data available on the exact hour of protests during the BLM movement. The BLM movement has been a significant social movement in recent years, drawing attention to issues of systemic racism and police brutality. However, due to the decentralized nature of the movement, it has been challenging to obtain accurate information about the exact hours of protests. However, The data published on site acleddata.com greatly helped this study in the field of creating a dataset. While the JSON file did contain information on the hours, minutes, and seconds of each tweet, this information could not be used to implement the model based on the precise time of protests.

Another limitation of our study was the lack of similar studies with common datasets for comparison purposes. This made it difficult to compare our findings with other studies and evaluate the accuracy of our predictions. Additional investigation is needed in this area to establish a standard dataset for forecasting protest events accurately.

Regarding the Place Info variable, another limitation we encountered was the lack of a complete list of streets, squares, and places in the United States of America. With such a list, we could have made better use of this indicator in our study and provided more accurate predictions. However, without this information, this variable's impact was limited, and we had to rely on other factors to predict the likelihood of protests.

In conclusion, our study highlights some of the challenges and limitations associated with predicting protest events using social media data. While Twitter provides an important information source to and analyze protest events, insufficient details available on the timing and location of these events can make it difficult to develop accurate models. Additionally, the lack of a common dataset for comparison purposes and incomplete information on geographical locations limits the effectiveness of some variables in our model. Future research in this area should focus on addressing these limitations and developing more accurate models for predicting protest events.

6.5. Conclusion and Thoughts for the Future

In this dissertation, we propose a detailed analysis of Twitter's open data to forecast future protests and violent events. The findings of this study showed a high correlation between the occurrence of protests and tweets, which indicates the reliability of Twitter as an indicator for predicting protest events and violence in them. Therefore, the behavioral patterns of Twitter users on social media have become a significant source for capturing, comprehending, and analyzing protest events. We introduce a novel method for forecasting the day of protest and the possibility of violence during the protest using Twitter user behavior and the Bayesian Logistic Regression algorithm. The study dataset was obtained from the combination of the two open data and then based on the triple Twitter users' behaviors, the desired features were extracted from it. Based on the extracted features, the proposed algorithm was implemented on five models, and the best combination for predicting protests and violence in protests was obtained. The triple Twitter users' behaviors besides the day and the number of informative tweets to predict protests and violence in protests provide the best model. The results indicate the frequency of tweets that contain date and time information is the best indicator for identifying protests. Hate-anger tweet rates are also the best indicator of violence in protests.

The size of a protest is a crucial factor in understanding its impact and effectiveness. Our subsequent investigation will involve developing a framework for estimating the size of protests based on our dataset. This will involve integrating multi-source data to more accurately estimate The number of people who took part in the protest. Debates over the size of protests have become common, with protesters seeking to emphasize widespread commitment and governments seeking to downplay the extent of the protest by understating the number of protesters. By using multi-source data, we aim to provide a more accurate estimate of the size of protests that is not subject to manipulation. Our proposed framework will incorporate a range of features to estimate the size of protests. These include the month, day, location, black rate per capita income, and unemployment rate of states. We will also use user tweets as another source of data to provide additional information on the size of protests. To estimate the size of protests accurately, we will use a combination of variables extracted from Twitter user behavior, such as the number of tweets, spatial tweets, and the rate of sad, happy, neutral, and hate tweets. These features will be used alongside other relevant variables to create a robust model that can accurately estimate the size of protests. Overall, developing a framework for estimating the size of protests is a critical next step in our research. By combining multi-source data and incorporating additional features into our models, we hope to provide more accurate estimates of protest size that can aid policymakers, activists, and the general public. With this framework, we can better understand the impact and effectiveness of protests and contribute to ongoing debates about their significance in contemporary society.

REFERENCE

[1] Cao, L., et al., *Behavior informatics: A new perspective.* IEEE Intelligent Systems, 29(4): pages 62-80. **2014**.

[2] Shibutani, T., *Human nature and collective behavior: Papers in honor of Herbert Blumer*. 2023: Taylor & Francis. **2023**.

[3] Wei, L., et al., *Do social media literacy skills help in combating fake news spread? Modelling the moderating role of social media literacy skills in the relationship between rational choice factors and fake news sharing behaviour.* Telematics and Informatics, 76: p. 101910. **2023**.

[4] Lorenz-Spreen, P., et al., *A systematic review of worldwide causal and correlational evidence on digital media and democracy.* Nature human behaviour, 2023. **7**(1): p. 74-101.

[5] Mirbabaie, M., et al., *Social media in times of crisis: Learning from Hurricane Harvey for the coronavirus disease 2019 pandemic response.* Journal of Information Technology, 35(3): p. 195-213. **2020**.

[6] Zafarani, R., M.A. Abbasi, and H. Liu, *Social media mining: an introduction*. Cambridge University Press. **2014**.

[7] Korolov, R., et al. *On predicting social unrest using social media*. in *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. IEEE. **2016**.

[8] Chen, L., J. Chen, and C. Xia, *Social network behavior and public opinion manipulation.* Journal of Information Security and Applications, 64: p. 103060. **2022**.

[9] Wang, Z. and K. Caskey, *# Occupywallstreet: an analysis of twitter usage during a protest movement.* Social Networking, 5(04): p. 101. **2016**.

[10] Ortiz-Ospina, E. and M. Roser, *The rise of social media.* Our world in data, **2023**.

[11] Aichner, T. and F. Jacob, *Measuring the degree of corporate social media use.* International Journal of market research, 57(2): pages 257-276. **2015**.

[12] Atefeh, F. and W. Khreich, *A survey of techniques for event detection in twitter.* Computational Intelligence, 31(1): pages 132-164. **2015**.

[13] Fernando, S., et al., *Towards a large-scale twitter observatory for political events.* Future Generation Computer Systems, 110: pages 976-983. **2020**.

[14] Zhao, L., et al., *Spatio-Temporal Event Forecasting Using Incremental Multi-Source Feature Learning.* ACM Transactions on Knowledge Discovery from Data (TKDD), 16(2): pages 1-28. **2021**.

[15] Hany Hossny, A. and L. Mitchell, *Event detection in Twitter: A keyword volume approach.* arXiv e-prints, p. arXiv: 1901.00570. **2019**.

[16] Goode, B.J., et al., *Pricing a protest: Forecasting the dynamics of civil unrest activity in social media.* PloS one, 10(10): p. e0139911. **2015**.

[17] Zhao, L., et al., *Unsupervised spatial event detection in targeted domains with applications to civil unrest modeling.* PloS one, 9(10): p. e110206. **2014**.

[18] Ramakrishnan, N., et al. *'Beating the news' with EMBERS: forecasting civil unrest using open source indicators*. in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. **2014**.

[19] Korkmaz, G., et al., *Multi-source models for civil unrest forecasting.* Social network analysis and mining, 6(1): pages 1-25. **2016**.

[20] Doyle, A., et al., *Forecasting significant societal events using the embers streaming predictive analytics system.* Big data, 2(4): pages185-195. **2014**.

[21] Xu, J., et al. *Civil unrest prediction: A tumblr-based exploration*. in *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer. **2014**.

[22] Haile, I.M. and Y. Qu, *Mitigating Risk in Financial Industry by Analyzing Social-Media with Machine Learning Technology*. European Journal of Electrical Engineering and Computer Science, 6(2): pages 33-37. **2022**.

[23] Talbot, C.V., et al., *How people with dementia use twitter: A qualitative analysis.* Computers in Human Behavior, 102: pages 112-119. **2020**.

[24] Weerkamp, W. and M. De Rijke. *Activity prediction: A twitter-based exploration*. in *SIGIR workshop on time-aware information access*. **2012**.

[25] Dickinson, T., et al. *Identifying prominent life events on twitter*. in *Proceedings of the 8th International Conference on Knowledge Capture*. **2015**.

[26] Kwak, H., et al. *What is Twitter, a social network or a news media?* in *Proceedings of the 19th international conference on World wide web*. **2010**.

[27] Phuvipadawat, S. and T. Murata. *Breaking news detection and tracking in Twitter*. in *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. IEEE. **2010**.

[28] Sakaki, T., M. Okazaki, and Y. Matsuo. *Earthquake shakes Twitter users: real-time event detection by social sensors*. in *Proceedings of the 19th international conference on World wide web*. **2010**.

[29] Brzozowski, M.J. and D.M. Romero. *Who Should I Follow? Recommending People in Directed Social Networks*. in *ICWSM*. **2011**.

[30] Farzindar, A. and D. Inkpen. *Proceedings of the Workshop on Semantic Analysis in Social Media*. in *Proceedings of the Workshop on Semantic Analysis in Social Media*. **2012**.

[31] Bollen, J., H. Mao, and X. Zeng, *Twitter mood predicts the stock market.* Journal of computational science, 2(1): pages 1-8. **2011**.

[32] Tsakalidis, A., et al., *Predicting elections for multiple countries using Twitter and polls.* IEEE Intelligent Systems, 30(2): pages 10-17. 2015.

[33] Gerber, M.S., *Predicting crime using Twitter and kernel density estimation.* Decision Support Systems, 61: pages 115-125. **2014**.

[34] Small, T.A., *What the hashtag? A content analysis of Canadian politics on Twitter.* Information, communication & society, 14(6): p. 872-895. **2011**.

[35] Wang, W. *Data Science for Social Good-2014 KDD Highlights*. in *AAAI*. **2015**.

[36] Sobel, M., D. Riffe, and J.B. Hester, *Twitter diplomacy? A content analysis of eight US embassies' Twitter feeds.* The Journal of Social Media in Society, 5(2): pages 75-107. **2016**.

[37] Tillmann, P., *Trump, Twitter, And Treasuries.* Contemporary Economic Policy, **2020**.

[38] Sadineni, P.K., *Mining in Big Data: Challenges, Solutions.* Solutions (November 20, 2020), **2020**.

[39] Kumar, A., S.R. Sangwan, and A. Nayyar, *Multimedia social big data: Mining*, in *Multimedia Big Data Computing for IoT Applications*. Springer. pages 289-321. **2020**.

[40] Li, J. and C. Cardie, *Early stage influenza detection from twitter.* arXiv preprint arXiv:1309.7340, **2013**.

[41] Culotta, A. *Towards detecting influenza epidemics by analyzing Twitter messages*. in *Proceedings of the first workshop on social media analytics*. **2010**.

[42] Eichstaedt, J.C., et al., *Psychological language on Twitter predicts county-level heart disease mortality.* Psychological science, 26(2): pages 159-169. **2015**.

[43] Serfass, D.G. and R.A. Sherman, *Situations in 140 characters: assessing real-world situations on Twitter.* PloS one, 10(11): p. e0143051. **2015**.

[44] Wald, R., et al. *Using Twitter content to predict psychopathy*. in *2012 11th International Conference on Machine Learning and Applications*. IEEE. **2012**.

[45] Golbeck, J., et al. *Predicting personality from twitter*. in *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE. **2011**.

[46] Sumner, C., et al. *Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets*. in *2012 11th international conference on machine learning and applications*. IEEE. **2012**.

[47] Guntuku, S.C., et al. *Studying personality through the content of posted and liked images on Twitter*. in *Proceedings of the 2017 ACM on web science conference*. **2017**.

[48] De Choudhury, M., S. Counts, and E. Horvitz. *Predicting postpartum changes in emotion and behavior via social media*. in *Proceedings of the SIGCHI conference on human factors in computing systems*. **2013**.

[49] De Choudhury, M., et al., *Predicting depression via social media.* Icwsm, 13: pages 1-10. **2013**.

[50] Hechtlinger, Y., N. Klein, and H.A. Song, *Predicting Latent User Attributes on Twitter*. **2015**

[51] Aldahawi, H., *Mining and analysing social network in the oil business: Twitter sentiment analysis and prediction approaches*. Cardiff University. **2015**.

[52] Dijkman, R., et al., *Using twitter to predict sales: A case study*. arXiv preprint arXiv:1503.04599, **2015**.

[53] Asur, S. and B.A. Huberman. *Predicting the future with social media*. in *2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology*. IEEE. **2010**.

[54] Lassen, N.B., R. Madsen, and R. Vatrapu. *Predicting iphone sales from iphone tweets*. in *2014 IEEE 18th International Enterprise Distributed Object Computing Conference*. IEEE. **2014**.

[55] Jansen, B.J., et al., *Twitter power: Tweets as electronic word of mouth.* Journal of the American society for information science and technology, 60(11): pages 2169-2188. **2009**.

[56] Bollen, J., A. Pepe, and H. Mao, *Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena.* arXiv preprint arXiv:0911.1583, **2009**.

[57] Mao, H., S. Counts, and J. Bollen, *Quantifying the effects of online bullishness on international financial markets*. ECB Statistics Paper. **2015**.

[58] Eysenbach, G., *Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact.* Journal of medical Internet research, 13(4): p. e123. **2011**.

[59] Shen, W., et al. *Linking named entities in tweets with knowledge base via user interest modeling*. in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. **2013**.

[60] Sang, E.T.K. and J. Bos. *Predicting the 2011 dutch senate election results with twitter*. in *Proceedings of the workshop on semantic analysis in social media*. **2012**.

[61] Skoric, M., et al. *Tweets and votes: A study of the 2011 singapore general election*. in *2012 45th hawaii international conference on system sciences*. IEEE. **2012**.

[62] Preoţiuc-Pietro, D., et al. *Beyond binary labels: political ideology prediction of twitter users*. in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. **2017**.

[63] Conover, M.D., et al. *Predicting the political alignment of twitter users*. in *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE. **2011**.

[64] Makazhanov, A., D. Rafiei, and M. Waqar, *Predicting political preference of Twitter users.* Social Network Analysis and Mining, 4(1): p. 193. **2014**.

[65] Pennacchiotti, M. and A.-M. Popescu. *A machine learning approach to twitter user classification*. in *Fifth international AAAI conference on weblogs and social media*. **2011**.

[66] Tumasjan, A., et al. *Predicting elections with twitter: What 140 characters reveal about political sentiment*. in *Fourth international AAAI conference on weblogs and social media*. Citeseer. **2010**.

[67] Chung, J.E. and E. Mustafaraj. *Can collective sentiment expressed on twitter predict political elections?* in *AAAI*. **2011**.

[68] Schumaker, R.P., A.T. Jarmoszko, and C.S. Labedz Jr, *Predicting wins and spread in the Premier League using a sentiment analysis of twitter.* Decision Support Systems, 88: p. 76-84. **2016**.

[69] Godin, F., et al. *Beating the bookmakers: leveraging statistics and Twitter microposts for predicting soccer results*. in *KDD Workshop on Large-Scale Sports Analytics*. **2014**.

[70] Sinha, S., et al., *Predicting the NFL using Twitter.* arXiv preprint arXiv:1310.6998, **2013**.

[71] Burger, J.D., et al. *Discriminating gender on Twitter*. in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. **2011**.

[72] Culotta, A., N.R. Kumar, and J. Cutler. *Predicting the Demographics of Twitter Users from Website Traffic Data*. in *AAAI*. Austin, TX. **2015**.

[73] Sap, M., et al. *Developing age and gender predictive lexica over social media*. in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. **2014**.

[74] Ugheoke, T.O. and R. Saskatchewan, *Detecting the gender of a tweet sender*. M. Sc. Project Report. Department of Computer Science, University of Regina, **2014**.

[75] Deitrick, W., et al., *Gender identification on twitter using the modified balanced winnow*. **2012**.

[76] Nguyen, D., et al. *Why gender and age prediction from tweets is hard: Lessons from a crowdsourcing experiment*. in *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. **2014**.

[77] Gadiya, M. and S. Jain, *Gender Prediction using Images Posted on Online Social Networks.* **2016**.

[78] Arroju, M., A. Hassan, and G. Farnadi. *Age, gender and personality recognition using tweets in a multilingual setting*. in *6th Conference and Labs of the Evaluation Forum (CLEF 2015): Experimental IR meets multilinguality, multimodality, and interaction*. **2015**.

[79] Flekova, L., D. Preoţiuc-Pietro, and L. Ungar. *Exploring stylistic variation with age and income on twitter*. in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. **2016**.

[80] Vicente, M., F. Batista, and J.P. Carvalho. *Improving Twitter Gender Classification using Multiple Classifiers*. in *Proceedings of the 8th European Symposium on Computational Intelligence and Mathematics (ESCIM-2016), Sofia, Bulgaria*. **2016**.

[81] Allan, J., *Introduction to topic detection and tracking*, in *Topic detection and tracking*. Springer. pages 1-16. **2002**.

[82] Allan, J., *Topic detection and tracking: event-based information organization*. Vol. 12.: Springer Science & Business Media. **2012**.

[83] Campos, R., et al., *Survey of temporal information retrieval and related applications.* ACM Computing Surveys (CSUR), 47(2): pages 1-41. **2014**.

[84] Alonso, O.R., *Temporal information retrieval*. Citeseer. **2008**.

[85] Allan, J., et al., *Topic detection and tracking pilot study final report.* **1998**.

[86] Yang, Y., T. Pierce, and J. Carbonell. *A study of retrospective and on-line event detection*. in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. **1998**.

[87] Boettcher, A. and D. Lee. *Eventradar: A real-time local event detection scheme using twitter stream*. in *2012 IEEE International Conference on Green Computing and Communications*. IEEE. **2012**.

[88] Xie, L., H. Sundaram, and M. Campbell, *Event mining in multimedia streams.* Proceedings of the IEEE, 96(4): pages 623-647. **2008**.

[89] Imran, M., et al., *Processing social media messages in mass emergency: A survey.* ACM Computing Surveys (CSUR), 47(4): pages 1-38. **2015**.

[90] Dou, W., et al. *Event detection in social media data*. in *IEEE VisWeek Workshop on Interactive Visual Text Analytics-Task Driven Analytics of Social Media Content*. **2012**.

[91] Hasan, M., M.A. Orgun, and R. Schwitter, *A survey on real-time event detection from the twitter data stream.* Journal of Information Science, 44(4): pages 443-463. **2018**.

[92] Hurlock, J. and M.L. Wilson. *Searching twitter: Separating the tweet from the chaff*. in *Fifth International AAAI Conference on Weblogs and Social Media*. **2011**.

[93] Lee, K., B.D. Eoff, and J. Caverlee. *Seven months with the devils: A long-term study of content polluters on twitter*. in *ICWSM*. **2011**.

[94] Castillo, C., M. Mendoza, and B. Poblete. *Information credibility on twitter*. in *Proceedings of the 20th international conference on World wide web*. **2011**.

[95] Khodabakhsh, M., M. Kahani, and E. Bagheri, *Predicting future personal life events on twitter via recurrent neural networks.* Journal of Intelligent Information Systems,: pages 1-27. **2018**.

[96] Khodabakhsh, M., et al., *Detecting life events from twitter based on temporal semantic features.* Knowledge-Based Systems, 148: pages 1-16. **2018**.

[97] Mahmoudpour, B., *Introducing and Evaluating the Scientific Event.* Science and Technology Policy Letters, 11(3): p. 98-112. **2021**.

[98] Li, J., et al. *Major life event extraction from twitter based on congratulations/condolences speech acts*. in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. **2014**.

[99] Adel, H., et al., *Improving crisis events detection using distilbert with hunger games search algorithm.* Mathematics, 10(3): p. 447. **2022**.

[100] Wiriyathammabhum, P., *ClassBases at CASE-2022 Multilingual Protest Event Detection Tasks: Multilingual Protest News Detection and Automatically Replicating Manually Created Event Datasets.* arXiv preprint arXiv:2301.06617, **2023**.

[101] Sufi, F.K., *AI-Tornado: An AI-based Software for analyzing Tornadoes from disaster event dataset.* Software Impacts, 14: p. 100357. 2022.

[102] Gerbaudo, P., *The pandemic crowd.* Journal of International Affairs, 73(2): p. 61-76. **2020**.

[103] Lancaster, L., *Unpacking discontent: Where and why protest happens in South Africa.* South African Crime Quarterly, 64: p. 29-43. **2018**.

[104] Boulianne, S., M. Lalancette, and D. Ilkiw, *"School strike 4 climate": social media and the international youth protest on climate change.* Media and Communication, 8(2): pages 208-218. **2020**.

[105] Hoegh, A., et al., *Bayesian model fusion for forecasting civil unrest.* Technometrics. 57(3): pages 332-340. **2015**.

[106] Katragadda, S., et al. *Detection of event onset using twitter*. in *2016 international joint conference on neural networks (IJCNN)*. IEEE. **2016**.

[107] Katragadda, S., R. Benton, and V. Raghavan, *Framework for real-time event detection using multiple social media sources.* **2017**.

[108] Cordeiro, M. and J. Gama, *Online social networks event detection: a survey*, in *Solving Large Scale Learning Tasks. Challenges and Algorithms*, Springer. pages 1-41. **2016**.

[109] Liu, X., M. Wang, and B. Huet, *Event analysis in social multimedia: a survey*. Frontiers of Computer Science, 10(3): pages 433-446. **2016**.

[110] Hasan, M., M.A. Orgun, and R. Schwitter. *TwitterNews+: a framework for real time event detection from the Twitter data stream*. in *International Conference on Social Informatics*. Springer. **2016**.

[111] Kaleel, S.B. and A. Abhari, *Cluster-discovery of Twitter messages for event detection and trending.* Journal of computational science, 6: pages 47-57. **2015**.

[112] Zhang, X., et al., *Event detection and popularity prediction in microblogging.* Neurocomputing, 149: pages 1469-1480. **2015**.

[113] Zhou, D., L. Chen, and Y. He. *An unsupervised framework of exploring events on twitter: Filtering, extraction and categorization*. in *Twenty-ninth aaai conference on artificial intelligence*. **2015**.

[114] Madani, A., O. Boussaid, and D.E. Zegour, *Real-time trending topics detection and description from twitter content.* Social Network Analysis and Mining, 5(1): p. 59. **2015**.

[115] Pistoni, C., et al., *Collective Action, Protest, and Covid-19 Restrictions: Offline and Online Community Participation in Italy and Australia*, in *Psychosocial Perspectives on Community Responses to Covid-19*. Routledge. pages 66-78. **2023**.

[116] Haferkamp, H. and N.J. Smelser, *Social change and modernity*. 1992: University of California press Berkeley. **1992**.

[117] Allan, J., R. Papka, and V. Lavrenko. *On-line new event detection and tracking*. in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. **1998**.

[118] D'Andrea, E., et al., *Real-time detection of traffic from twitter stream analysis.* IEEE transactions on intelligent transportation systems, 16(4): pages 2269-2283. **2015**.

[119] Duong, H.-T. and T.-A. Nguyen-Thi, *A review: preprocessing techniques and data augmentation for sentiment analysis.* Computational Social Networks, 8(1): p. 1-16. **2021**.

[120] Lovins, J.B., *Development of a stemming algorithm.* Mech. Transl. Comput. Linguistics, 11(1-2): pages 22-31. **1968**.

[121] Paice, C.D. *Another stemmer*. in *ACM Sigir Forum*. ACM New York, NY, USA. **1990**.

[122] Willett, P., *The Porter stemming algorithm: then and now.* Program, **2006**.

[123] Gimpel, K., et al., *Part-of-speech tagging for twitter: Annotation, features, and experiments*. Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science. **2010**.

[124] Derczynski, L., et al. *Twitter part-of-speech tagging for all: Overcoming sparse and noisy data*. in *Proceedings of the international conference recent advances in natural language processing ranlp 2013*. **2013**.

[125] Yadav, V. and S. Bethard, *A survey on recent advances in named entity recognition from deep learning models.* arXiv preprint arXiv:1910.11470, **2019**.

[126] Trimastuti, W., *An analysis of slang words used in social media.* Jurnal Dimensi Pendidikan dan Pembelajaran, 5(2): pages 64-68. **2017**.

[127] Lin, C., et al. *Representations of time expressions for temporal relation extraction with convolutional neural networks*. in *BioNLP 2017*. **2017**.

[128] Lin, C., et al. *A BERT-based universal model for both within-and cross-sentence clinical temporal relation extraction*. in *Proceedings of the 2nd Clinical Natural Language Processing Workshop*. **2019**.

[129] Chang, A.X. and C.D. Manning. *Sutime: A library for recognizing and normalizing time expressions*. in *Lrec*. **2012**.

[130] Ahmed, K.R., et al., *A simple and robust wetland classification approach by using optical indices, unsupervised and supervised machine learning algorithms.* Remote Sensing Applications: Society and Environment, 23: p. 100569. **2021**.

[131] Miao, J. and W. Zhu, *Precision–recall curve (PRC) classification trees.* Evolutionary intelligence, 15(3): pages 1545-1569. **2022**.

[132] Kallus, N. *Predicting crowd behavior with big public data*. in *Proceedings of the 23rd International Conference on World Wide Web*. 2014]. **2014**.

[133] Cadena, J., et al., *Forecasting social unrest using activity cascades.* PloS one, 10(6): p. e0128879. **2015**.

[134] Bahrami, M., et al., *Twitter reveals: using twitter analytics to predict public protests.* arXiv preprint arXiv:1805.00358, **2018**.

[135] Tuke, J., et al., *Pachinko Prediction: A Bayesian method for event prediction from social media data.* Information Processing & Management, 57(2): p. 102147. **2020**.

[136] Agarwal, S. and A. Sureka. *Investigating the potential of aggregated tweets as surrogate data for forecasting civil protests*. in *Proceedings of the 3rd IKDD Conference on Data Science, 2016*. **2016**.

[137] Muthiah, S., et al. *Planned protest modeling in news and social media*. in *Twenty-Seventh IAAI Conference*. **2015**.

[138] Ertugrul, A.M., et al., *Activism via attention: interpretable spatiotemporal learning to forecast protest activities.* EPJ Data Science, 8(1): p. 5. **2019**.

[139] Compton, R., et al. *Detecting future social unrest in unprocessed twitter data:"emerging phenomena and big data"*. in *2013 IEEE International Conference on Intelligence and Security Informatics*. IEEE. **2013**.

[140] Steinert-Threlkeld, Z.C., et al., *Online social networks and offline protest.* EPJ Data Science, 4(1): pages 1-9. **2015**.

[141] Bakerman, J., et al., *Dynamic logistic regression and variable selection: Forecasting and contextualizing civil unrest.* International Journal of Forecasting, **2021**.

[142] Timoneda, J.C. and E. Wibbels, *Spikes and variance: using Google Trends to detect and forecast protests.* Political Analysis, 30(1): pages 1-18. **2022**.

[143] Alsaedi, N., P. Burnap, and O. Rana, *Can we predict a riot? Disruptive event detection using Twitter.* ACM Transactions on Internet Technology (TOIT), **17**(2): pages 1-26. **2017.**

[144] Muthiah, S., et al. *Planned protest modeling in news and social media*. in *Twenty-Seventh IAAI Conference*. Citeseer. **2015**.

[145] Renaud, M., et al. *Social network structure as a predictor of social behavior: The case of protest in the 2016 US presidential election*. in *XVIII International Conference on Data Science and Intelligent Analysis of Information*. Springer. **2018**.

[146] Grill, G., *Future protest made risky: Examining social media based civil unrest prediction research and products.* Computer Supported Cooperative Work (CSCW), 30(5): pages 811-839. **2021**.

[147] Hua, T., et al. *STED: semi-supervised targeted-interest event detectionin in twitter*. in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. **2013**.

[148] Williams, M.L., et al., *Policing cyber-neighbourhoods: tension monitoring and social media networks.* Policing and society, 23(4): pages 461-481. **2013**.

[149] Rule, W., et al. *Social pressure analysis of local events using social media data*. in *2018 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE. **2018**.

[150] Gallagher, R.J., et al., *Divergent discourse between protests and counter-protests:# BlackLivesMatter and# AllLivesMatter.* PloS one, 13(4): p. e0195644. **2018**.

[151] Qiao, F. and K. Chen. *Predicting protest events with Hidden Markov models*. in *2016 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*. IEEE. **2016**.

[152] Ortiz, I., et al., *An Analysis of World Protests 2006–2020*, in *World Protests*. 2022, Springer. p. 13-81.

[153] Steckel, J.C., et al., *Effects of the energy price crisis on European households.* 2022.

[154] Au, A.C.-H., *The economic costs of protests and policy recommendations for preventing them: the case of Hong Kong.* Journal of Chinese Economic and Foreign Trade Studies, **2021**.

[155] Schultz, D., *The $2 billion-plus price of injustice: a methodological map for police reform in the George Floyd era.* Law & Ineq., 39: p. 571. **2021**.

[156] Mourtgos, S.M., I.T. Adams, and J. Nix, *Elevated police turnover following the summer of George Floyd protests: A synthetic control study.* Criminology & Public Policy, 21(1): pages 9-33. **2022**.

[157] Çelebi, E., J. Adam-Troian, and Y. Mahfud, *Positive links between exposure to police violence, PTSD, and depression symptoms among yellow vests protesters in France.* Journal of interpersonal violence, 37(5-6): pages 2267-2288. **2022**.

[158] Ni, M.Y., et al., *Mental health during and after protests, riots and revolutions: a systematic review.* Australian & New Zealand Journal of Psychiatry, 54(3): pages 232-243. **2020**.

[159] Boyd, M.J., et al. *Understanding the Security and Privacy Advice Given to Black Lives Matter Protesters*. in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. **2021**.

[160] Kowalewski, M., *Book Review: Protests in the Information Age: Social Movements, Digital Practices and Surveillance by Lucas Melgaço and Jeffrey Monaghan*. SAGE Publications Sage UK: London, England. **2022**.

[161] Liu, T., *Multi-class emotion classification for short texts.* Published online at https://tlkh. github. io. Accessed, 16: p. 11. **2020**.

[162] Gerbaudo, P., *From Occupy Wall Street to the Gilets Jaunes: On the populist turn in the protest movements of the 2010s.* Capital & Class, 47(1): pages 107-124. **2023**.

[163] Bouazizi, M. and T. Ohtsuki, *A pattern-based approach for multi-class sentiment analysis in Twitter.* IEEE Access, 5: pages 20617-20639. **2017**.

[164] Giorgi, S., et al., *Twitter corpus of the# blacklivesmatter movement and counter protests: 2013 to 2020.* arXiv preprint arXiv:2009.00596, **2020**.

[165] Binkley, P., *twarc-report README. md.* **2015**.

[166] Giorgi, S., et al. *Twitter Corpus of the# BlackLivesMatter Movement and Counter Protests: 2013 to 2021*. in *Proceedings of the International AAAI Conference on Web and Social Media*. **2022**.

[167] Shen, C., S. Panda, and J.T. Vogelstein, *The chi-square test of distance correlation.* Journal of Computational and Graphical Statistics, 31(1): pages 254-262. **2022**.

[168] Chang, A.X. and C.D. Manning. *Sutime: A library for recognizing and normalizing time expressions*. in *Lrec*. **2012**.

[169] Hung, L.P. and S. Alias, *Beyond Sentiment Analysis: A Review of Recent Trends in Text Based Sentiment Analysis and Emotion Detection.* Journal of Advanced Computational Intelligence and Intelligent Informatics, 27(1): pages 84-95. **2023**.

[170] Mann, S., et al., *Twitter sentiment analysis using enhanced bert*, in *Intelligent Systems and Applications: Select Proceedings of ICISA 2022*. Springer. pages 263-271. **2023**.

[171] Liu, B., *Sentiment analysis and subjectivity*. Handbook of natural language processing, 2(2010): p. 627-666. **2010**.

[172] Ravi, K. and V. Ravi, *A survey on opinion mining and sentiment analysis: tasks, approaches and applications.* Knowledge-Based Systems, 89: pages 14-46. **2015**.

[173] Pang, B., L. Lee, and S. Vaithyanathan, *Thumbs up? Sentiment classification using machine learning techniques.* arXiv preprint cs/0205070, **2002**.

[174] Turney, P.D., *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews.* arXiv preprint cs/0212032, **2002**.

[175] Chang, A.X. and C.D. Manning, *Sutime: A library for recognizing and normalizing time expressions.* LREC 2012. 3735: p. 3740. **2012**.

[176] Nahla, B.S., *Generalized convolution product of an infinitely divisible distribution and a Bernoulli distribution.* Communications in Statistics-Theory and Methods, pages 1-8. **2021**.

[177] Garrido-Merchán, E.C. and D. Hernández-Lobato, *Dealing with categorical and integer-valued variables in bayesian optimization with gaussian processes.* Neurocomputing, 380: pages 20-35. **2020**.

[178] Sun, D., et al., *Assessment of landslide susceptibility mapping based on Bayesian hyperparameter optimization: A comparison between logistic regression and random forest.* Engineering Geology, 281: p. 105972. **2021**.

APPENDICES

A - Publication Derived from thesis

Hossein Zadeh, M., Cicekli, I. (2023). Protest Event Analysis: A New Method Based on Twitter's User Behaviors. Information Technology and Control, 52(2), 457-470. https://doi. org/10.5755/j01.itc.52.2.33077