KNOWLEDGE MANAGEMENT OF SOCIAL MEDIA DATA FOR DISASTER MANAGEMENT

Ph.D. THESIS

by

Annie Singla

CENTRE OF EXCELLENCE IN DISASTER MITIGATION AND MANAGEMENT
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE- 247667
JUNE, 2022

KNOWLEDGE MANAGEMENT OF SOCIAL MEDIA DATA FOR DISASTER MANAGEMENT

A THESIS

Submitted in partial fulfilment of the requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

in

Disaster Management

by

Annie Singla

CENTRE OF EXCELLENCE IN DISASTER MITIGATION AND MANAGEMENT
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE- 247667
JUNE, 2022

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work presented in the thesis entitled **KNOWLEDGE MANAGEMENT OF SOCIAL MEDIA DATA FOR DISASTER MANAGEMENT** is my own work carried out during a period from December, 2016 to May, 2022 under the supervision of Dr. Rajat Agrawal, Professor, Department of Management Studies, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India.

The matter presented in the thesis has not been submitted for the award of any other degree of this or any other Institute.

Date: __________ **(ANNIE SINGLA)**

SUPERVISOR'S DECLARATION

This is to certify that the above mentioned work is carried out under my supervision.

Date: __________ **(RAJAT AGRAWAL)**

INTELLECTUAL PROPERTY RIGHTS AND PUBLICATIONS

Copyright

1. Granted copyright invention from Government of India, entitled ”DisDSS - a web-based decision-support system to predict the disaster-related social media messages”, registration no. SW-15177/2022 dated 10.01.2022.

Travel Funds

1. Dean of Resources and Alumni Association Fund for European Geoscience Union 2022, Vienna, Austria.

2. Professional Development Fund for European Geoscience Union 2022, Vienna, Austria.

3. Centre Development Fund for European Geoscience Union 2022, Vienna, Austria.

Research Articles

1. Published research paper entitled “Investigating Challenges and Enablers of using Social Media for Disaster Management” at Global Memory, Knowledge and Communication.

2. Submitted research paper entitled “iStage: An Intelligent Hybrid Deep Learning Based Framework to Determine the Stage of Disaster Management Cycle from a Social Media Message” at Journal of Information Science (Under review).

3. Submitted research paper entitled ”A Social Media-Based Comparative Study on Multi-Hazards Utilizing Deep Learning Techniques to Identify the Relevancy of Message to Disaster” at Ingeniera Solidaria (Under review).

4. Submitted research paper entitled ”DisDSS: A Web-Based Decision-Support System to Predict the Disaster-Related Social Media Messages” at Global Memory, Knowledge and Communication(Under revision).

5. Submitted ”iRelevancy: Developing a Framework to Identify the Relevancy of a Social Media Message to a Disaster” at Global Memory, Knowledge and Communication (Under revision).

List of Conferences

1. ”Management of Social Media Based Data for Covid-19 Disaster using Deep Learning” at Dealing with Disasters International Conference (DwD 2022) at Northumbria University, United Kingdom (June 09-10, 2022)

2. ”DisDSS 2.0: A Multi-Hazard Web-based Disaster Management System to Identify Disaster-Relevancy of a Social Media Message for Decision-Making Using Deep Learning Techniques” at European Geosciences Union (EGU) 2022 at Vienna, Austria (May 23-27, 2022).

3. ”Comparing the effectiveness of Convolutional Neural Network and Long Short-Term Memory Network for Disaster Based Social Media Messages - Using Thunderstorm and Cyclone as Case Studies” presented at the International Conference on Research in Management & Technovation 2021 (ICRMAT 2021). Proceedings is included in Annals of Computer Science and Information Systems, Poland (ISSN- 2300-5963).

4. “Management of Covid-19 using Social Media Data” accepted at 5th World Congress on Disaster Management (WCDM-2020).

5. ”Use of social media for disaster management” accepted at Pacific Risk Management ‘Ohana (PRiMO), held at Honolulu, Hawaii (August 06–09, 2018).

6. Invited to attend the UNISPACE+50 organized by Office for Outer Space Affairs in celebrating the 50th anniversary of the first United Nations Conference on the Exploration and Peaceful Uses of Outer Space at Vienna, Austria (June 2018).

Trainings and Workshops

1. Organized and participated in workshop “Risk and Safety Assessment of Built Environment in IIT Roorkee campus” at Centre of Excellence in Disaster Mitigation & Management, Indian Institute of Technology Roorkee (March 21st , 2018).

2. Participated in the national level training programme on “Knowledge Management for Disaster Risk Reduction & Resilience (KM4DRR&R)” at National Institute of Disaster Management (NIDM), New Delhi (April 23rd – 27th , 2018).

3. Participated in “Basics of Intellectual Property Rights” at Department of Management Studies in Indian Institute of Technology Roorkee (Jan 20th, 2018).

ABSTRACT

Disasters affect the lives, and infrastructure in a negative manner. With the internet fad, social media has become inevitable in our lives, generating tsunami of data. An individual cannot retrieve the same message on social media within a blink of eye. During disastrous times, data is even more critical. People use social media at catastrophic times. Data, being raw, and unstructured, needs to be in a knowledgable format, so that effective decisions can be taken at right time.

The overarching aim of this dissertation is to advance the management of knowledge for disaster management using social media data. Conventional knowledge management systems are not optimal enough to support disaster management processes. The dynamic nature of disasters offer different situations. With the evolving times, the knowledge management systems needs to be evolved to handle complex environments.

In the **first objective** of our research work, we explore the challenges and enablers of social media usage for disaster management by understanding views and perspectives of people working in disaster management domain. This was done with the help of literature review and data collected through focus group discussion.

The participants chosen for focus group discussion are homogeneously working in disaster management domain but are from heterogeneous backgrounds like civil, architecture, mechanical, management and computer science. The number of participants are 10, ranging from 21 to 42 years of age. 8 male participants and 2 female participants are there. Half of the participants are Master students pursuing disaster management. The remaining four participants are doctoral students and one professor of Centre of Excellence in Disaster Management is amongst the 10 participants.

The methodology developed is in concoction of existing literature and the efficacy of qualitative data obtained from focus group discussion, using Atlas.ti software following inductive thematic approach.

The transcripts are transcribed manually by the moderator. After acquiring validation from the participants, raw data is categorized using inductive thematic approach in Atlas.ti software. The results are finalized after expert validation. The themes are developed using the panoply of coding functions - Open coding, Quick coding, List coding and In-vivo coding - available in Atlas.ti 8. The identified challenges are physical, software, cultural, demographic, authenticity, and regulatory.

The identified enablers are rise in mobile penetration, democratic participation, increase in

living standards, two-way real-time communication, global reach, and cheaper source of information. The study results in contribution by explaining ”what” challenges and enables the usage of social media for disaster management. The research objective sheds a new light on the understanding of social media as a vital player in disaster management and contributes to enlarge the scope of advance research on the relationship between social media and disaster management.

For the **second** objective, The researcher wants to develop a framwork for reliable and accurate identification of disaster-related social media messages for effective management of disasters. This goal can be achieved using an efficient deep learning model with rich social media disaster-based data. In this objective, a novel deep learning based framework, iRelevancy, is proposed for identifying the disaster relevancy of a social media message using deep learning algorithms. The proposed system is evaluated with cyclone Fani data and compared with state-of-the-art deep learning models as well as the recent relevant studies. The performance of the experiments are evaluated by the accuracy, precision, recall, f1-score, area under receiver operating curve, and area under precision-recall curve score. The results show that our model is more effective for the identification of disaster-relevancy of a social media message, in comparison to other state-of-the-art methods. The predictive performance of the proposed model is illustrated with experimental results on cyclone fani data along with misclassifications. Further, to analyze the performance of the proposed model, the results on other cyclonic disasters, i.e., cyclone Titli, cyclone Amphan, and cyclone Nisarga are presented. In addition, the framework is implemented on catastrophic events of different nature, i.e., Covid-19. The research study can assist disaster managers in effectively manoeuvring disasters at the time of distress.

In the **third objective** of our research work, we aim to propose a framework to identify the stage of disaster, i.e., pre, during, post, or irrelevant from a social media message. Extracting knowledge from the social media data during different stages of disaster management cycle is a challenging task. Deep learning has shown great potential in automatic identification from a large amount of raw data in various domains. We propose iStage, i.e., an intelligent hybrid deep learning based framework to determine the stage of the disaster to take right decisions at the right time. To demonstrate the effectiveness of iStage, it is applied on cyclonic and Covid-19 disasters. The considered disaster datasets are cyclone Fani, cyclone Titli, cyclone Amphan, cyclone Nisarga and Covid-19. The experimental results demonstrate that the proposed model outperforms Long Short-Term Memory Network and Convolutional Neural Network models. The proposed approach returns best possible solution among existing research studies considering different evaluation metrics- accuracy, precision, recall, f-score, area under receiver operating characteristic curve, and area under precision-recall curve. Hence, iStage can assist disaster personnel in a better way to manage disasters.

The **fourth** objective aims to propose a web-based smart disaster management system for decision-making that will assist disaster professionals to determine the nature of disaster-related social media message. We consider Covid-19 as our case study. We initiate our research by reviewing the literature pertaining to the usage of social media in Covid-19, and the existing social media-based disaster management systems. Hence, we identify that the existing systems lack the web-interface, considering the importance of the social media message, due to which the most significant messages, i.e., help-seeking and help-offering are not explored. Therefore, in this objective, we address this issue with a web-application.

It is worth mentioning that a fusion-based deep learning model is introduced to objectively determine the nature of a social media message. The developed system leads to a better performance in accuracy, precision, recall, F-score, area under receiver operating curve, and area under precision-recall curve, compared to other state-of-the-art methods in the literature. The contribution of this objective is three folds: Firstly, it presents a new covid dataset of social media messages with the label of nature of message. Secondly, it presents a fusion-based deep learning model to classify the social media data. Thirdly, it presents a web-based interface to visualise the structured information. Furthermore, the architecture of DisDSS is analyzed based on practical case study, i.e., covid-19. The proposed deep learning based model is embedded into a web-based interface for decision-support. To the best of our knowledge, this is the first social media-based disaster management system in India.

The objectives of the dissertation deals with knowledge management of social media data for disaster management domain. The objectives transform the raw and unstructured data into a knowledgable format, so that disaster managers are able to handle disasters effectively and efficiently. The dissertation advances the academic knowledge in better understanding the role of social media for disaster management. The study is important as it helps in determining the challenges and enablers of social media usage for disaster management.

The research has provided another dimension to the social media usage understanding for disaster management. As such, the study extends the inadequate knowledge of barriers and enablers. No doubt, the usage of social media during disasters is ever increasing and additional knowledge would assist in the formulation of effective policies in shielding the society from the menance of disasters. Another contribution of this research is that challenges and enablers studies are largely focused on the Western part of the world. Contrary to this, the research considers discussion group from India, which has gained little research attention thus far.

The dissertation sets up some practical implications. The research recommends the proposed deep learning approach, which outperforms the baseline models, evaluating on different paramters. Another practical contribution is the creation of a Twitter dataset of cyclone Fani and Covid-19. Cyclone Fani dataset has relevancy label and stage of disaster label. Covid-19 dataset has label of nature of the social media message. Also, the results demonstrate the prediction results and web-interface.

ACKNOWLEDGEMENTS

Having an opportunity to acknowledge the help I have had, in completing my research work towards the doctoral degree, is really the thing I have been waiting for. The very first name that comes to my mind is of my learned supervisor, **Professor Rajat Agrawal**, Department of Management Studies, Indian Institute of Technology Roorkee. He has been *saarthi* of my PhD rath. I wish to express my deep sense of gratitude, and appreciation to him, for his stimulating supervision with invaluable suggestions, keen interest, constructive criticism, and constant encouragement during the course of the study. I have learnt a lot from him, not only in academics, but in all spheres of life. I wholeheartedly acknowledge his full cooperation that I received from the very beginning of this work, upto the completion in the form of thesis.

During this research, the Centre has had three heads, Professor B K Maheshwari, Professor Mahua Mukherjee, and Professor Sumit Sen, and fortunately, they encouraged me by extending every sort of help as and when sought for.

The author is specially thankful to Mr. Tahir Hussain, Centre of Excellence in Disaster Mitigation and Management, for his invaluable help, and advice from time to time. I would like to make a special note for Professor Ajay Gairola, who has been a father figure to me, and been there for me in the most difficult times. I thank all the faculty members of the Centre for their cooperation. The financial support provided by Ministry of Education (formerly as Ministry of Human Resource Development), New Delhi to complete the research work is highly acknowledged.

When hurdles appeared insurmountable and target unachievable, the encouragement and camaraderie of friends helped keep things in perspective. I thank all my lovely friends, who helped lighten the burden, especially to Iqra, Tahseen, Smita, Vanitha.

The ever enthusiastic help of my family members is supremely acknowledged. I am in dearth of proper words to express my abounding feelings for my loving husband, *Mr. Aman Garg*, who has been there for me in thick, and thin. Words are not enough to thank my *in-laws*, who left me carefree for household worries, and let me concentrate on my research work. My daughter *Hannah Garg*, is my lifeline, who being a toddler, has been my tiny supporter. She never bothered me in my work, rather she has been my constant little motivator. My sister-in-law *Mrs. Rinki Garg*, and brother-in-law *Mr. Lucky Garg*, are my friends-cum-family. They are my go-to people, whenever I have something to worry about.

My mother *Mrs. Parveen Singla*, is next to God. She is someone, who believed in me from the very core. She constantly motivated me, in entire tenure. My father, *Mr. Rakesh Singla*, a gentleman, who has given me wings to fly, and dream big in life. My baby brother, who is not baby anymore, *Mr. Nixon Singla*, is my fun partner. I am highly grateful for having such a lovely family.

Last, but not the least, I thank Almighty, for blessing me with so much in life.

(Annie Singla)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Rationale for Research

In recent years, disasters have occured frequently across the globe, resulting in extensive losses to lives, and property, significantly impacting the social order, productivity, and everyday life. According to Centre for Research on the Epidemiology of Disasters and United Nations Office for Disaster Risk Reduction (2021), there are 389 reported disasters in 2020. In addition, disasters claim lives of 15,080 people, 98.4 million people are affected globally, and US 171.3 billion dollars are spent on economic damage. India, due to its, physiographic and climatic conditions is one of the most disaster-prone countries of the world. More than 40 million hectares (12 per cent of land) is prone to floods and river erosion. Of the nearly 7,500 km long coastline, close to 5,700 km is prone to cyclones and tsunamis. Nearly 68% of the cultivable area is vulnerable to drought. Large tracts in hilly regions are at risk from landslides and some are prone to snow avalanches. Vulnerability of the nearly 7,500 km long coastline, close to 5,700 km is prone to cyclones and tsunamis. (National Disaster Management Authority, 2019).

International agreements such as the Sendai framework for disaster risk reduction encourages the use of Social Media (SM) to strengthen disaster risk communication. SM is a new way of communication these days. These are apps that allow citizens to post text, image, video and links to websites. The Prime Minister of The Republic of India enunciated ten-point agenda, consisting of utilization of opportunities provided by SM, and mobile technologies for disaster risk reduction. SM is transforming disaster response. It is helping response agencies in quickly organizing themselves and enabling citizens to connect more easily with authorities. In disaster, affected people use social media to help each other. Those responsible for DM must recognize the potential of SM and develop applications relevant to various aspects of disaster risk management. (National Disaster Management Authority, 2019).

Modern SM applications are considerably penetrated into lives of people, providing an insight into people's opinions, beliefs, and sentiments. SM connects people of different genders, race, ethnicity, and origin. According to Global Digital Report 2021 (Simon Kemp, 2021), Figure: 1.1, the global population being 7.83 billion at the start of 2021, and 5.22 billion people are using mobile phones, equating the number to 66.6% of the total population. At this moment, 4.66 billion people use internet, i.e., 59.5%. The same report pinpoints the number of SM users, mentioning that 4.20 billion active SM users are there, i.e., 53.6% of the population.

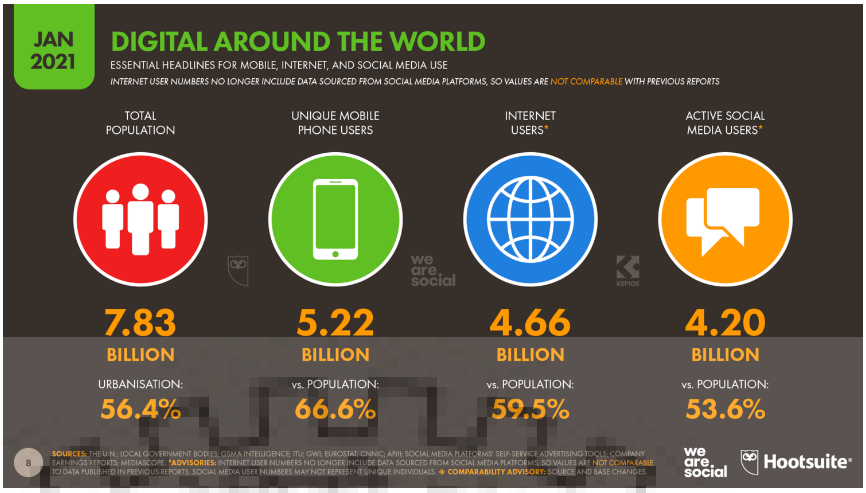


Fig. 1.1: Global Digital Report, 2021 (Simon Kemp, 2021).

Currently, Twitter is a major example of SM usage with millions of users scattered around the globe. As of 2020, Twitter had more than 450 million active users. According to Internet Live Statistics 2021, Twitter disseminates 500 million tweets each day, approximating it to 200 billion tweets in a year [1]. The importance of SM in disasters has been reviewed by several scholars (Abedin et al., 2014; Akter and Wamba, 2017; Gaspar et al., 2019; Goswami et al., 2016; Simon et al., 2015). Feldman et al. (2016) conduct a research study on communicating flood risk at SM outlets. Al-Saggaf and Simmons (2015) explore the use of SM during two natural disasters (i.e., the Jeddah floods of 2009 and 2011) in Saudi Arabia, utilizing YouTube, Facebook, Al-Saha-Al-Siyasa, and Al Arabiya platforms. Rexiline Ragini et al. (2018) explores tweets of Indo-Pakistan heavy floods and landslides of 2015, cyclone HUDHUD 2014, and cyclone Nilofer 2014 to detect the people at risk.

Luckily, the data is relatively available in public domain for research purposes. However, extracting relevant information immersed in noisy, irrelevant, and unstructured SM data is a challenge in SM studies. This has brought into focus the significance of KM in DM. Caballero-Anthony et al. (2021) define KM as a process of creating, capturing, codifying, storing, sharing, distributing, and effectively using knowledge. KM means gaining the correct information, at the right time, and right place. Managing knowledge is paramount in turbulent, fast changing environment of disasters.

Knowledge is information in action as per context, relevance and usefulness. It is classified in two forms: tacit and explicit. Explicit knowledge is documented knowledge that is objective, rational and technical, and it can be transformed into process or strategy. Tacit knowledge is a store of subjective or observed learning experiences. It comprises of insights, expertise, understanding, and skill sets (Meneghello et al., 2020).

KM in DM refers to the process of acquiring, managing, and utilising disaster information

[1] https://www.internetlivestats.com/twitter-statistics/

and knowledge for the support of disaster operations.

1.2 Aim of the Research

Traditional KM systems are not optimally configured to support the DM, as disasters offer completely unique situations. Resource constraints may require new ways to think about existing responsibilities, and functions. Knowledge availability varies more extremely than in normal situations, i.e., manier times, little information is sufficient to make informed decisions, and other times, multiple reportings to the information may unnecessarily increase the information processing computational load. Since, faster decision-making is imperative in a compressed timeline to stabilize the disastrous situation. Hence, DM systems that will be useful in DM must be flexible enough to handle unexpected situations, and robust enough to be reliable in degraded, and complex environments (Yates and Paquette, 2011). Decision-making processes require accurate and reliable knowledge. Figure: 1.2 showcases the crux of the thesis, i.e., KM to augment the effectiveness of DM via SM.

Fig. 1.2: Knowledge Management via Social Media to enhance the effectiveness of Disaster Management

Artificial Intelligence (AI) through Deep Learning(DL) algorithms allow systems to acquire, process, and use knowledge to perform tasks and to unlock knowledge to deliver it to people. We need to keep in mind that when it comes to AI, and a crucial domain DM, we need both responsible uses and design (Rhem, 2021). Timely information in DM is vital, but it still is challenging. DL algorithms are important part of prediction system using disaster data (Caragea et al., 2016a;

Huang et al., 2020a; Nguyen et al., 2019). In terms of the application of SM data for DM, research suggests that lack of tools for managing SM data during a disaster makes it difficult for disaster professionals, to not understand how SM data can be useful for public.

Recent technological advancements in the digital era have the potential to boost the knowledge-driven economy. Although some attention has been paid to specific concepts, and technologies as well as their impact on KM, the KM for DMC utilizing SM is pretty new. No comprehensive studies have addressed the full extent of the problem thus far. The research on KM is still in embryonic stage.

Disaster organizations require specific knowledge needs that enable them to conduct operations. Efforts are made more effective if there is a systematic and well-integrated system of knowledge sharing and exchange in place (Caballero-Anthony et al., 2021).

A failure to share information and knowledge can have adverse impact on collective decision-making during disasters. This results in a lack of coordination and inefficiency of disaster operations. The misallocation of resources, delayed operations, and overlapping of responses are some of the possible failures that could arise. Hence, it is imperative that the capacities are maximized. This can only occur if there is a KM system for information sharing and coordination. (Caballero-Anthony et al., 2021). KM system includes knowledge-based systems, document management systems, semantic networks, object oriented and relational databases, Decision Support System (DSS), expert systems and simulation tools Dorasamy et al. (2013).

1.3 Contributions

Aiming to address this gap, the thesis set sights on the following objectives as contributions:

1. Determining the challenges and enablers in the use of SM for DM. The objective investigates the challenges faced in the usage of SM for DM, by considering views and perceptions of people working in homogenous DM environment. In doing so, an FGD is conducted to understand the role of SM intially and what are the barriers in the usage in DM. The objective further identifies the enablers affecting the SM usage for DM. Finally, the objective addresses the issue in Indian context.

2. Developing a framework to identify the relevancy of a SM message to a disaster. The objective determines the disaster-relevancy of SM message. The considered disasters are cyclone Fani, cyclone Titli, cyclone Amphan, cyclone Nisarga, and Covid-19. The SM platform is Twitter. A hybrid DL model is proposed with six evaluation metrics.

3. Developing a framework to determine at what stage of DMC, the disaster is, from a SM message. The objective contributes by developing a dataset of cyclone Fani with label of stage at which disaster is. The developed framework, iStage, works on DL algorithms, i.e., CNN and LSTM. The iStage is evaluated considering six different performance metrics.

Finally, the testing is perfomed on cyclone Fani testing dataset, cyclone Amphan, cyclone Nisarga, cyclone Titli, and Covid-19.

4. Developing a web-based DM system to determine the nature of SM message for decision-making. In KM, a key enabler is a KM system, designed to identify, share, retrieve, and use knowledge. The significant KM systems coincide with technological advancements that enable, mammoth amounts of data to be processed, stored and disseminated. KM systems enhance the visibility of the knowledge and the culture of collaboration and sharing (Caballero-Anthony et al., 2021). KM is one of the building blocks of DM. KM foster the ability to take decisions and augments the response speed, increase in coordination, helps in capacity development of disaster organizations by supporting the creation and transfer of tacit and explicit knowledge Kusumastuti et al. (2021). This objective advocates the use of DSS that can be deployed as a single access point. The system facilitates to avail the knowledge to the stakeholders/ users in an appropriate form for their decision-making process. A case study of recent Covid-19 disaster is successfully used to demonstrate the efficacy of the developed KM system, utilizing Twitter data Inan et al. (2018).

1.4 Thesis Organization

The rest of the dissertation is organized as follows: Chapter 2 explicates the conceptual background. Chapter 3 reviews the literature. Chapter 4 investigates the challenges, and enablers for using SM in DM. Chapter 5 proposes iRelevancy, i.e., a framework to identify the relevancy of a SM message to a disaster. Chapter 6 proposes iStage: a framework to determine the stage of DMC, in which disaster is, from the SM message. Chapter 7 develops DisDSS: a web-based interface to determine the nature of SM. Chapter 8 concludes the dissertation with the limitations, and future research directions.

2. CONCEPTUAL BACKGROUND

2.1 Knowledge Management Cycle

The three main stages of KM cycle are:

- Knowledge capture and/or creation
- Knowledge sharing and/or dissemination
- Knowledge acquisition and/or application

Knowledge capture refers to the identification and subsequent codification of existing internal knwoledge and know-how within the organization and/or external knowledge from the environment. Knowledge creation is the development of new knowledge and know-how— innovations that did not have a previous existence within the company. In the transition from knowledge capture/creation to knowledge sharing and dissemination, knowledge content is assessed. Knowledge is then contextualized in order to be understood ("acquisition") and used ("application"). This stage then feeds back into the first one in order to update the knowledge content. Figure: 2.1 showcases the different stages of KM cycle.

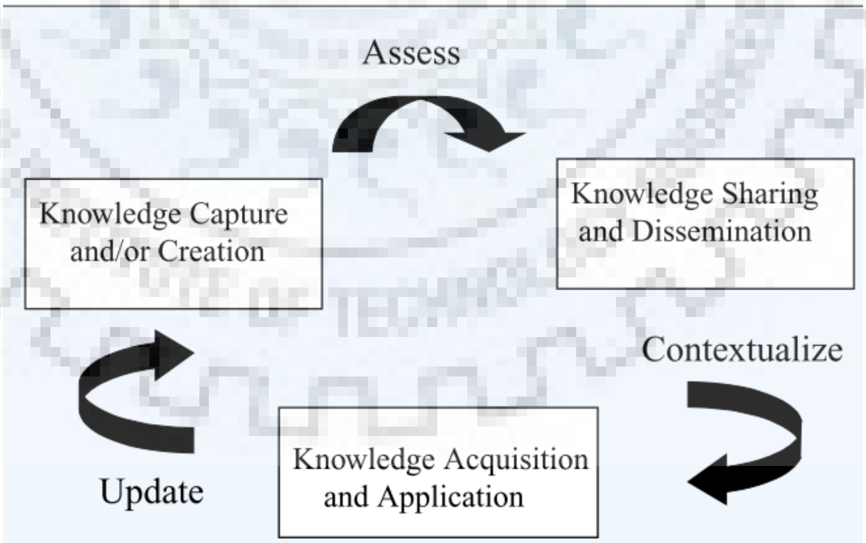


Fig. 2.1: Knowledge Management Cycle

2.2 SECI Model

Nonaka and Takeuchi provide a fundamental model of KM, known as SECI model. It mainly focuses on transforming tacit knowledge into explicit knowledge. The model consists of four stages:

1. socialization (tacit-to-tacit)
2. externalization (tacit-to-explicit)
3. combination (explicit-to-explicit)
4. internalization (explicit-to-tacit)

The four stages posit that knowledge is created by the creative tension between tacit and explicit know-how, leading to a dynamic flow of activities that facilitates the generation, transfer and application of knowledge. These activities repeat at increasing levels of complexity and enable knowledge expansion through a knowledge spiral. Various scholars have argued the limitations of SECI model. Li and Gao (2003) reviewed the SECI in a critical manner and suggested the model may not be applicable across all domains as it demonstrates the validity only on Japanese manufacturing companies. Powell (2007) criticizes the SECI model and state that the base on which model works is flawed, i.e., the fundamental assumption that knowledge can be transformed between tacit and explicit knowledge. Richter (2006) criticizes the lack of empirical basis of SECI model. Figure: 2.2 illustrates the SECI model with all the four stages.

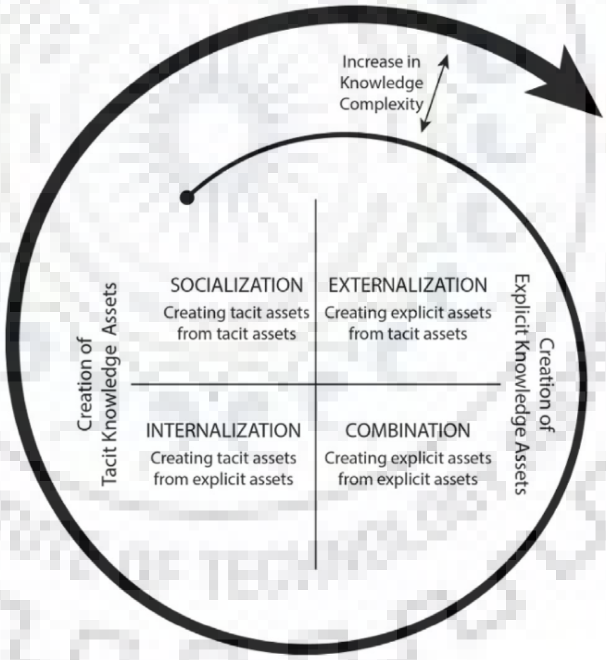


Fig. 2.2: SECI Model (Bandera et al., 2017)

2.3 The Choo Sense-making KM Model

Choo describes a KM model that stresses sense-making, knowledge creation, and decision-making concepts. The Choo KM model focuses on how information elements are selected and subsequently fed into actions. In the sense-making stage, one attempts to make sense of the information streaming in from the external environment. Priorities are identified and used to filter the information. Individual construct common interpretations from the exchange and negotiate information fragments combined with their previous experiences (Koenig, 2011). Figure 2.3 showcases the choo sense-making KM model.

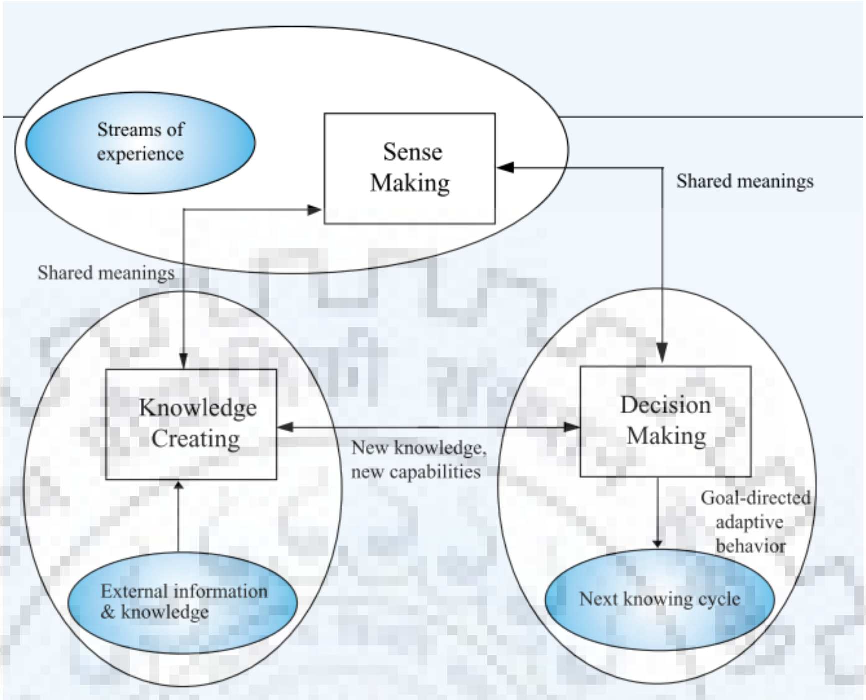


Fig. 2.3: Choo KM Model (Koenig, 2011).

2.4 Wiig Knowledge Management Model

Wiig approaches KM model with the following principle: in order for knowledge to be useful and valuable, it must be organized. Knowledge should be organized differently depending on what use will be made of the knowledge. Some useful dimensions to consider in Wiig's KM model include: (1) completeness, (2) connectedness, (3) congruency, and (4) perspective and purpose. Completeness addresses the question of how much relevant knowledge is available from a given source. Connectedness refers to the well-understood and defined relations between the different knowledge objects. A knowledge base is said to possess congruence when all the facts, concepts, perspectives, values, judgments, and associative and relational links between the knowledge objects are consistent. Perspective and purpose refer to the phenomenon through which one "know something" but often from a particular point of view or for a specific purpose. Wiig's KM model goes on to define different levels of internalization of knowledge.

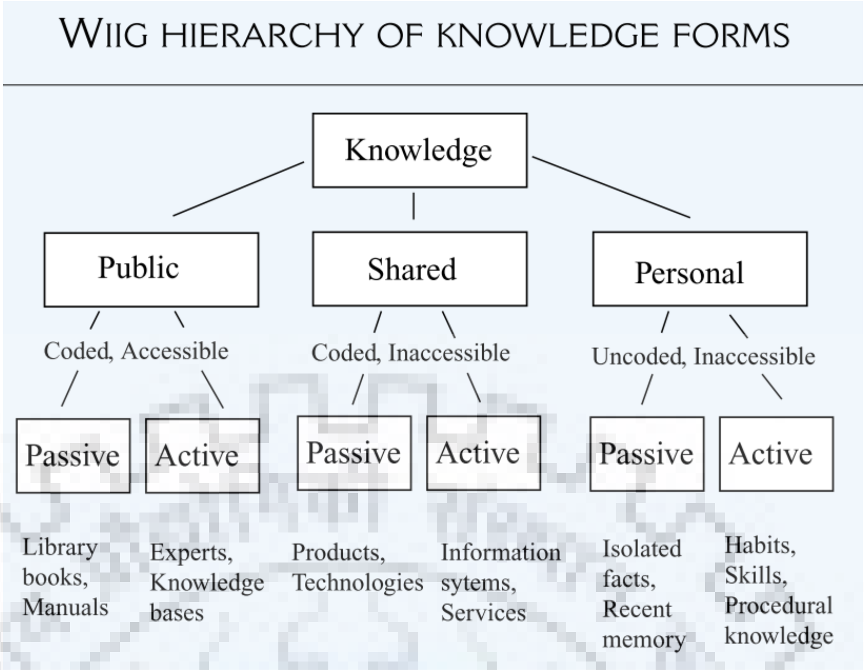


Fig. 2.4: Wiig Knowledge Management Model Koenig (2011)

2.5 The Boisot I-Space KM Model

The Boisot KM model is based on the key concept of an "information good" that differs from a physical asset. Boisot distinguishes information from data by emphasizing that information is what an observer will extract from data as a function of his or her expectations or prior knowledge. Boisot proposes the following two key points:

1. The more easily data can be structured and converted into information, the more diffusible it becomes.

2. The less data that has been so structured requires a shared context for its diffusion, the more diffusible it becomes.

The I-Space model can be visualized as a three-dimensional cube with the following dimensions:

1. codified—uncodified;

2. abstract— concrete;

3. diffused—undiffused (Koenig, 2011)

Figure 2.5 depicts the Boisot I-space KM model.

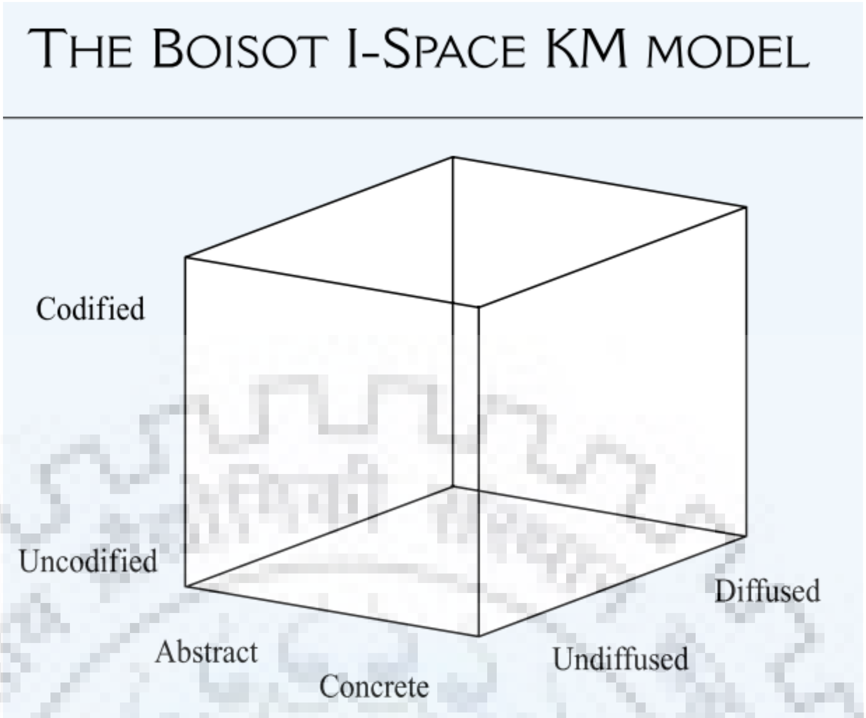


Fig. 2.5: The Boisot I-Space Knowledge Management ModelKoenig (2011)

2.6 DIKW Architecture

DIKW model has roots in KM. It is a hierarchial model with a pyramid shape [Figure: 2.6], and is known as Data-Information-Knowledge-Wisdom hierarchy. It has four key elements[Figure: 2.7 [1]], and are summarized as follows:

- Data: Data are symbols that represent properties of objects, events, and their environments. Data are individual facts, figures, signals, measurements, etc. Data is a collection of facts in a raw, or unorganized form [Figure: 2.7a].
- Information: Information is inferred from data. It is meaningful, relevant, usable, significant, or processed data. Human mind seeking questions in the form of ”what”, ”when”, ”where”, ”who”, or ”how many”, are processed into an answer. When this happens, the data becomes ”information”. Data itself is of no use, until it is in a relevant form. As a result, the difference between data, and information is functional, rather structural [Figure: 2.7b].
- Knowledge: Knowledge answers the ”know-how” question. The pieces of information are connected that help us understand how to apply the information to achieve our goal. The ”know-how” part of the DIKW hierarchy transforms the information to instructions [Figure: 2.7c].
- Wisdom: Wisdom is top of DIKW hierarchy. It is knowledge applied in action, taking into account what is known (knowledge), and what does the most good [Figure: 2.7d].

[1] https://www.ontotext.com/knowledgehub/fundamentals/dikw-pyramid/

In row with these ideas, the following metaphor applies: data: 'know–nothing', information: 'know–what', knowledge: 'know–how', and wisdom: 'know–why'. If we consider data, and information are like a look back into past, knowledge, and wisdom are associated with what we do now, and what we want to achieve in the future (Aven, 2013).

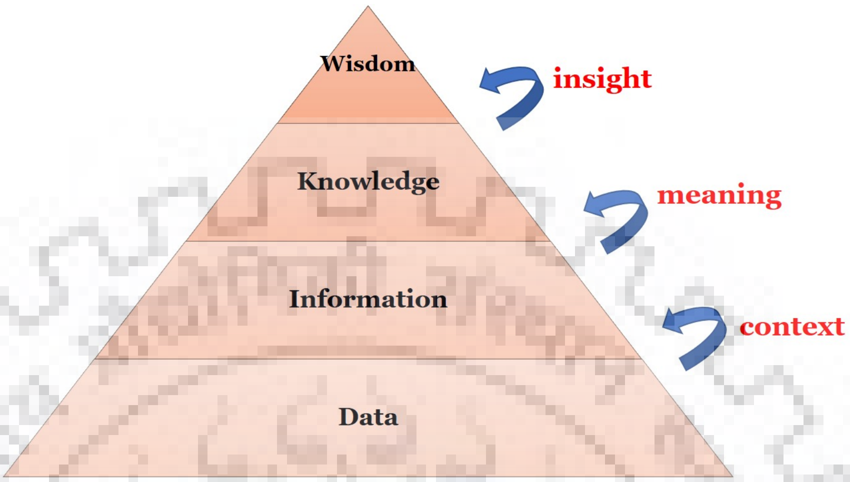


Fig. 2.6: DIKW Architecture

Fig. 2.7: Elements of the DIKW hierarchy.

Most of the literature and the practice of KM focuses on the management of existing knowledge, and less on the notion of innovation through KM. Fast-forwarding to the present, the environments are evolved with seamless data collection, sophisticated search, and mining systems. The KM and DM need to be instrinsically connected, very limited literature focuses on how disaster managers use KM. One important aspect of KM is the need to structure and link the KM effort to contextual imperatives. The DIKW model looks at simple ways to extract insight from all sorts of data to make useful decisions. The DM community can take full benefit from the full maturity of data towards wisdom by enabling management to take informed decisions Pretorius et al. (2019). DIKW is preferred over other KM models because of its logical relationship between elements. Itd escribes the progression from data to knowledge and then to wisdom. The pyramid is a powerful metaphor as it represents the hierarchial strata, structure, stability, integrity, maturity, royalty, and authenticity. It is a simple model to assist in understanding complicated, complex or novel constructs Williams (2014).

2.7 Disaster Management Cycle

According to National Disaster Management Authority (2019), a disaster is defined as " *A serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability and capacity, leading to one or more of the following: human, material, economic, and environmental losses and impacts.*" and DM is "the organization, planning, and application of measures preparing for, responding to and recovering from disasters." DM cycle is a system with several components segragated into three time periods, i.e., *pre, during, and post-disaster*. It is a sum total of all the aspects involved in DM, i.e., disaster risk reduction, mitigation, preparedness, response, recovery, and building back better. Figure: 2.8 illustrates the DM cycle periods with its varying functions.

Fig. 2.8: Disaster Management Cycle.

2.8 Deep Learning Techniques

In the realm of AI, DL techniques are subset of Machine Learning (ML). They are biologically inspired computational networks consisting of single interconnected processing components, i.e., neurons, arranged in layers. CNN and LSTM techniques are chosen in the dissertation. CNN is very effective in reducing the number of parameters without losing on the quality of models and LSTM has benefit of dealing with vanishing gradient which is encountered while dealing with traditional Recurrent Neural Network (RNN).

2.8.1 Convolutional Neural Network

CNNs are a type of DL algorithm highly inspired by the working of the human brain (Petmezas et al., 2021). It is a mathematical structure consisting of multiple stacked layers -convolutional, pooling, and fully connected. The functioning of each layer of CNN is described in Table: 2.1. The output of each layer is fed into the next layer as input, as seen in Figure: 2.9.

Layer	Functionality
Embedding	The very first layer of the model converts the positive integers into fixed-sized dense vectors
Convolution	Extracts high-level features by applying convolution operation
Pooling	Reduces the spatial size of a convoluted feature to reduce the computational load to process the data
Dropout	To overcome the overfitting issue, dropout layers randomly dissociate the connections between neurons of connected layers as two or more neurons start to detect the same feature repeatedly, i.e., co-adaptations.
Fully-connected	Connects every neuron from the previous layer to the next layer

Tab. 2.1: CNN Layer Functionality.

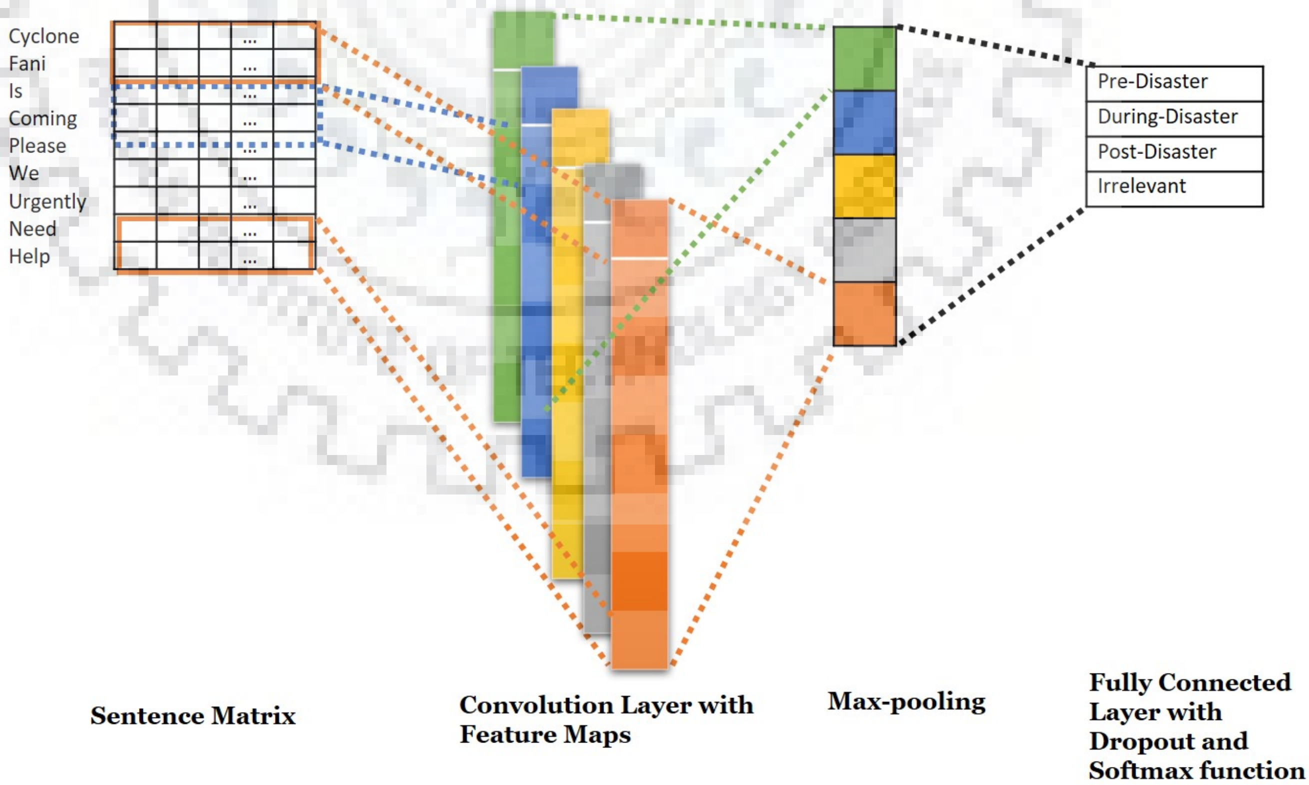


Fig. 2.9: CNN Architecture

Mathematically, each layer is represented as:

Input Layer:

$$x_{1:n} = x_1 \oplus x_2 \oplus \cdots \oplus x_n \tag{2.1}$$

where $\oplus$ = concatenation operator,
$x_i \in R^d$ = d-dimensional word vector corresponding to the i-th word in the sentence.

Convolutional Layer:

$$c_i = s(w.x_{i:i+h-1} + b) \tag{2.2}$$

where c_i = new feature,
$w \in R^{hk}$ = filter,
$i : i + h - 1$ = window size of h words,
$b \in R$ = bias term,
s = non-linear activation function, such as Rectified Linear Units (ReLU), sigmoid or hyperbolic tangent (tanh) functions.

For the sentence, the possible windows are $x_{1:h}; x_{2:h+1} \cdots ; x_{n-h+1:n}$, that leads to the feature map as:

$$c = [c_1, c_{2,\cdots c_{n-h+1}}] \in R^{n-h+1} \tag{2.3}$$

Max-pooling Layer: Corresponding to one particular filter, it selects maximum value

$$\hat{c} = \max[c] \tag{2.4}$$

Softmax Function:

$$z = [\hat{c}_1, \hat{c}_2 \cdots \hat{c}_m] \tag{2.5}$$

$$y_j = softmax(W.\mathbf{z} + b) \tag{2.6}$$

where m = number of filters
W = final feature vector
y_j = Resultant Output

2.8.2 Long Short-Term Memory Network

LSTM is a subcategory of RNN. It is capable of learning long-term dependencies between data. This feature makes LSTM more suitable for predictive applications as it requires maintenance of the information for a longer time. RNN cannot overcome this problem easily. LSTM solves the gradient vanishing problem as well. In an LSTM architecture, as showcased in Figure: 2.10, each network consists of a memory block, i.e., cells. Each cell transfers two states to the next one, cell and hidden state. A gate containing different weights acts as a layer of neurons. LSTM implemented gates to control the memorization process, solving the problem of long-term dependencies.

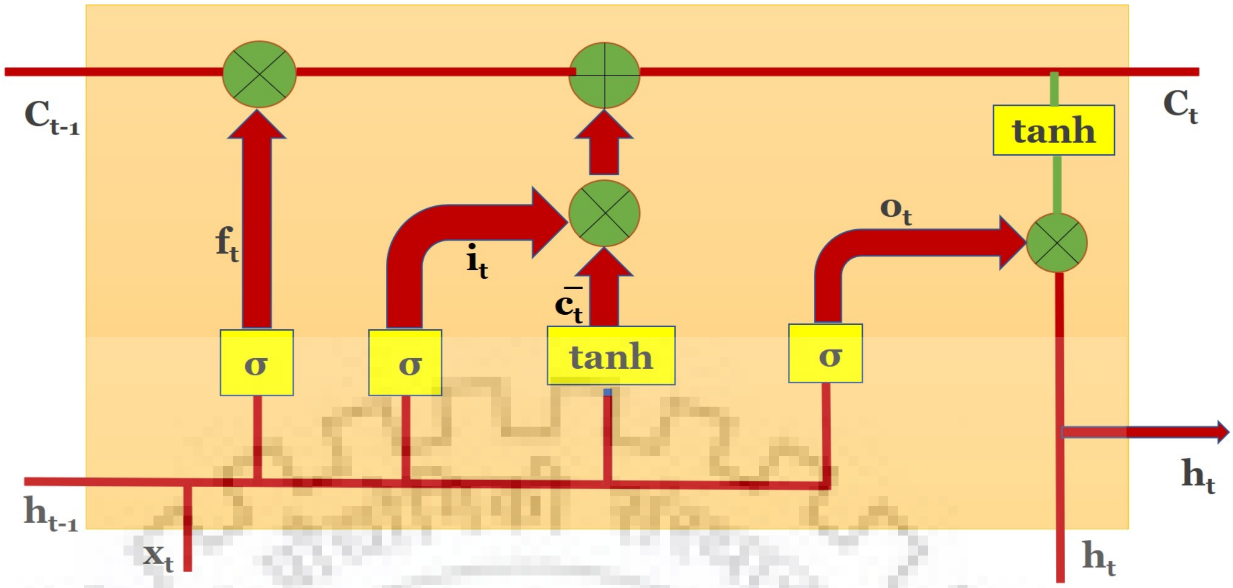


Fig. 2.10: LSTM Architecture.

Mathematically, the equations that describe the LSTM are:

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \tag{2.7}$$

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \tag{2.8}$$

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \tag{2.9}$$

$$\tilde{C}_t = tanh(x_t U^g + h_{t-1} W^g) \tag{2.10}$$

$$C_t = \sigma(f_t * C_{t-1} + i_t * \tilde{C}_t) \tag{2.11}$$

$$h_t = tanh(C_t) * o_t \tag{2.12}$$

where x_t is the input vector
h_{t-1} is the previous cell input
C_{t-1} is the previous cell memory
h_t is the current cell output
C_t is the current cell memory (Varsamopoulos et al., 2018).

2.8.3 *Bidirectional Long Short-Term Memory Network*

Bidirectional LSTM (BiLSTM) is a further development of LSTM. It combines a forward hidden layer and a backward layer, which allows it to access both preceding and succeeding contexts. In constrast to BiLSTM, LSTM exploits the historical context, i.e., information flows only from backward to forward. BiLSTM retains information by utilizing both ways of direction, input data of preceding, and succeeding sequence (Tam et al., 2021). The diagrammatic illustration of the BiLSTM model is depicted in Figure: 2.11. The h_t is a hidden state, where *t* nominates the t^{th} moment. x_t, and y_t is the input and output state, respectively (Liu and Guo, 2019). The figure showcases that at each time step *t*, a hidden forward layer $\overrightarrow{h_t}$ is computed based on the previously hidden state $\overleftarrow{h_{t-1}}$, and the current input x_t. Similarly, a hidden backward layer $\overleftarrow{h}$ is computed based on future hidden state $\overleftarrow{h_{t+1}}$, and current input x_t. The forward and backward context representations $\overrightarrow{h_t}$, and $\overleftarrow{h_t}$ are then concatenated into a long vector at time step *t* as (Do et al., 2019):

$$\overleftrightarrow{h_t} = [\overrightarrow{h_t}; \overleftarrow{h_t}] \tag{2.13}$$

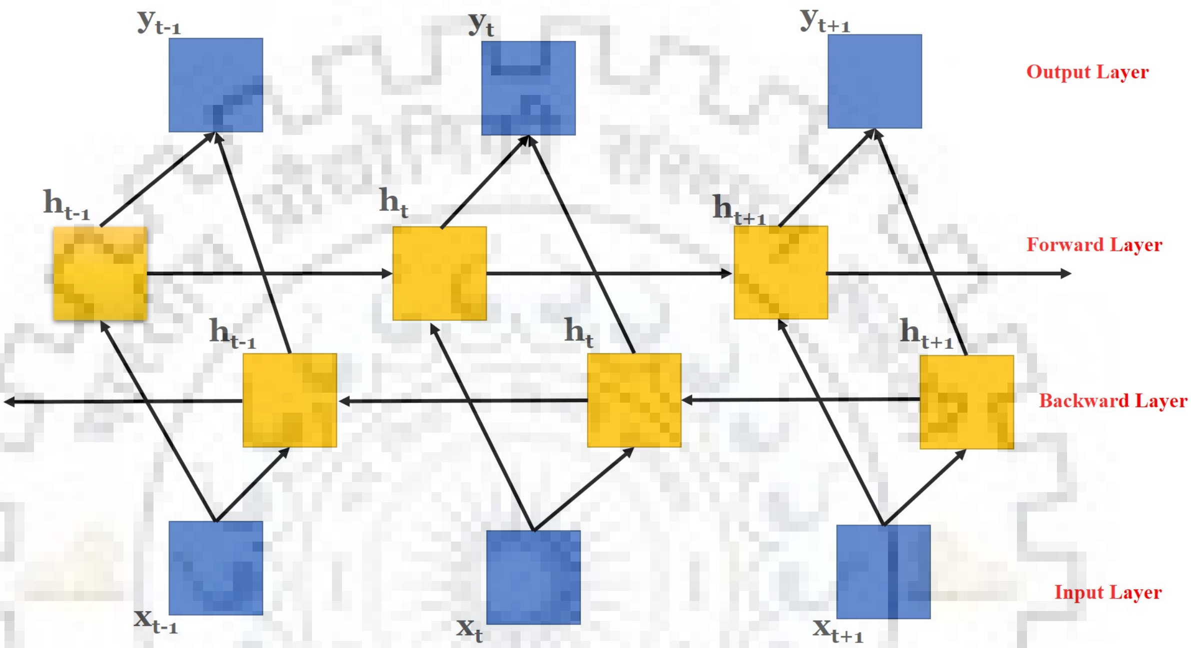


Fig. 2.11: Bidirectional Long Short-Term Memory Model.

The chapter deals with the conceptual clarity of the topics researched upon in the dissertation. The chapter leads with the KM concepts, i.e., KM cycle and KM models, translating it to the path of DMC, and DL tecniques. The chapter describes in-detail different existing KM models, with different phases of DMC. It provides mathematical foundation of DL algorithms with the pictorial representation. The chapter provides a highly concise overview that supplements the literature review of KM for DM via SM.

3. LITERATURE REVIEW

The existing body of literature is analyzed according to structured literature review method. The method is relevant to the dissertation to achieve high validity as well as reliability, the structured literature review is conducted. The validity in this context means the degree of accuracy in identifying and handling sources, including a comprehensive selection of scientific databases and search terms. Reliability is the replicability of the search process and can be achieved by thoroughly documenting the procedure and by making the selection criteria explicit.

The conducted literature review procedure is described in the following: In the first step of literature review, we take as many relevant research results into consideration as possible, an exhaustive literature search within a multitude of different databases using different search terms is conducted. The presented literature review remain both exhaustive and selective because of the broad spectrum of research contributions.

As a second step of literature review, it is recommended to work with key concepts and working definitions (Hofmann et al., 2015). Thus, the domain of DM, SM, and KM is explored. A first analysis of domain-specific sources of DM is made apparent that terminology within this domain is particularly manifold. A distinction is made amongst different kinds of disasters, denoted as, ”crisis”, ”emergency”, ”disaster”, and ”catastrophe.” All of these events have in common that they are limited in foreseeability and predictability; are non-routine; occur suddenly; and threaten the assets.

The next step of a structured literature review involves the selection of databases, the development of search strategies, and an initial evaluation of the literature. Since high-impact journals and conferences are recommended, the considered databases are Thomson Reuters Web of Knowledge, IEEE Explore Digital Library, Wiley Online Library, AIS Electronic Library, Elsevier/Science Direct, Emerald, Springerlink, dblp Computer Science bibliography, Inderscience, Sage, Taylor & Francis, Association of Computing Machinery Digital Library, Scopus, ArXiV, Emerald Insight, ERIC Institute of Education Sciences, Association for Information Systems eLibrary, ResearchGate, Google Scholar, and others. The thesis collection from MIT, Stanford, Oxford, Shodhganga is also explored. Only English language articles have been chosen for the study.

This set of databases, journals, and conference proceedings allow a literature search within a broad range of international scientific journals as well as in high-ranking conference proceedings. The defined set is searched for all contributions that have the identified keywords within their title and/or abstract. The keywords from DM, SM, and KM are combined with the aid of logical operators as follows: ("emergency management" OR "disaster management" OR "crisis management" OR "catastrophe management" OR "knowledge management" OR "social media" OR "social networks").

While this query is applied in Science Direct, IEEEXplore Digital Library, ACM Digital Library, it was necessary to adjust it for other databases. Not every search engine provide the functionality to search in title and abstract simultaneously and others did not allow the combination of keywords by multiple logical operators. Thus, simplified search queries are used, e.g., ("disaster management" AND "social media" AND "knowledge management) in Scopus. The collection of literature ended in 2022 and 6188 articles in total.

As next step of a structured literature review, duplicates, and irrelevant contributions, i.e., false positives, are removed from the list of search results by a preliminary screening of titles and abstracts. In particular, the research contributions to only such contributions that deal with KM of SM and DM are considered. We also included contributions that explicitly focus on the applicability of Qualitative Data Analysis (QDA) and tools within the domain of DM, and challenges/enablers in a detailed manner. In contrast, we exlude the contributions that solely used AI tools and methods for other research areas. We also excluded articles that used DM only for exemplary application of their research approach without discussing DM-related requirements in detail.

This chapter is divided into seven sections. In the first section, 3.1, SM challenges and enablers are explored. Second section 3.2 reviews the research methods and tools for QDA. Section 3.3 explicates the SM usage in DMC. In section 3.4, the usage of AI for prediction and decision-making is elucidated. Section 3.5 reviews the applications of DL algorithms for DM. In section 3.6, fusion-based DL models are explored. Finally, section 3.7 explicates the DL-based studies using SM for DM.

3.1 Social Media Challenges and Enablers

Several studies have investigated the challenges and enablers for the usage of SM for DM. Table: 3.1 summarises the reviewed challenges and enablers for SM usage in DM. The *social* and *technological* challenges for the use of SM in emergency management has been explored in (Luna and Pennock, 2015), where social challenges comprise of stakeholders' universality, social stratification, economic status based on age, gender, education level, ethnicity; and sarcasm being a linguistic barrier leading to misinformation skyrocketing during an emergency. Technological challenges embody physical and logistical resources en masse with data management plus security. Lack of standard policies, formats or regulations, and SM budget and monetization is another technical challenge.

Kavanaugh et al. (2012) conducted an exploratory study on the use of SM for crisis management by government officials, businesses, and the public at large by conducting focus-group interviews of 25 officials, and the results were analyzed using Wordle, Perl script considering Twitter, Facebook, YouTube, and Flickr platforms. The author categorizes the challenges into *information* and *organization* factors. *Information* factors are sub-categorized into technology, communications, and information. *Organization* factors are sub-categorized into legal, policy, costs, and training.

Hiltz and Kushma (2014) discuss the limitations, barriers, and enhancements of SM, especially Twitter and Facebook, by public sector managers of the United States of America (USA). The authors conducted semi-structured interviews with emergency managers. The significant challenges result out to be lack of personnel time to work on SM usage for DM, inadequate policies and guidelines to use SM for DM, lack of trustworthiness on the SM data, and overloading of information. Gunessee et al. (2017) conducted a study on Chennai's 2015 floods discussed the barriers, enablers and motivators of local citizens and spontaneous volunteers during disaster relief operations, and it is found that SM acts as an enabler in digital volunteerism. Park and Johnston (2017) report the data inaccuracies, privacy, security, and lack of technical specialization of SM as limitations in digital volunteer networks for crisis response.

Tagliacozzo (2018) explores a mix of qualitative (semi-structured interviews) as well as quantitative (multiple choice questionnaires) research methods to analyze how SM underpins and hampers the Government to Citizen (G2C) communication during post-disaster reconstruction, revealing security issues, lack of policies and guidelines, concerned legal problems, etc. as potential barriers. (Petersen et al., 2018) conducts a qualitative study employing semi-structured interviews with French stakeholders to know about SM usage during the Paris attacks of 2015. The study mentions that changed real-time communication behavior, real-time situational update, and helping one another via social media act as enablers. The communication and data analysis challenges on the use of SM for emergency management have been reviewed in (Simon et al., 2015). The design challenges and solutions such as data modeling, time vs. space trade-offs, and the need for a valuable and usable system for supporting the SM data while managing crisis are described in (Anderson et al., 2015).

An explication in rumor research by comparing French and English tweets sent during Paris attacks 2015 identifies the language as a barrier. A focus on the technical enablers and their classification in crisis information management has been reviewed (Wilson et al., 2017). In addition, SM data mining, crisis and event recognition, location recognition, identification of false information, situational awareness, communication to and within the public, organizational use of communication systems, collaboration systems for volunteers and organizations, collective sense-making, volunteered geographic information are the identified enablers (Wilson et al., 2017).

Lack of staff and information overload have been considered barriers (Plotnick and Hiltz, 2018). In addition, less affluent and lesser-educated people, poorly functional infrastructure, the authenticity of the message, scalability of SM sites, security as well as privacy, lack of local language websites, lack of basic computer skills are considered as barriers in Taiwanese study (Huang et al., 2016). SM enables public participation at local and national levels, publicized meetings and events, information to residents about public safety, and networking in Philippine Local Government Units (Alampay et al., 2018). Lack of dedicated staff, network failure, misleading information, lack of clarity on responsible authorities as barriers and robustness, SM's capability of dealing with velocity and high traffic data, cost-effectiveness are identified enablers of SM for emergency management (Anikeeva et al., 2015).

Lambert (2020) explores the Facebook platform to analyze risk communication qualitatively. It identifies the ability to exchange information, transparency, trust, and support during disaster as enablers of SM usage. Lieneck et al. (2022) review barriers of covid-19 vaccine promotion on SM usage in the USA. The study identifies the spread of misinformation, decreased vaccine acceptance amongst SM users, and lack of SM regulations as barriers.

Article	Platforms Considered	Technique	Findings
(Luna and Pennock, 2015)	-	Interactions among stakeholders	**Social Challenges** - Dissemination of target information - Detection of irony - Misinformation Spreading - Information Verification, and - Validation in real-time **Technology Challenges** - Data Management - Data Privacy - Data Standardisation
(Kavanaugh et al., 2012)	- Twitter - Facebook - YouTube - Flickr	- FGD - Participant Questionnaires (25)	**Information Challenges** - Communications - Technology - Information **Organization Challenges** - Policy - Legal Issues - Costs - Training
(Hiltz and Kushma, 2014)	- Twitter -Facebook	- Semi-structured interviews (11)	**Challenges** - Lack of personnel time to work on SM usage for DM - Inadequate policies and guidelines to use SM for DM - Lack of trustworthiness on the SM data, and - Overloading of information
(Gunessee et al., 2017)	- Twitter	- Content Analysis	**Enablers** - Decentralized decision-making **Challenges** - Ineffective coordination - High-Risk Exposure
(Alampay et al., 2018)	- Facebook - Twitter	- Round Table Discussion	**Enablers** - Public Participation - Publicized Meetings - Networking - Information Dissemination

(Park and Johnston, 2017)	-	Webster and Watson approach	**Challenges** - Data and Information System Integration - Privacy and Security - Data inaccuracy - Lack of technical specialization
(Tagliacozzo, 2018)		- Qualitative Field Notes - Quantitative Multiple-choice questionnaires (N= 56)	**Challenges** - Security issues - Lack of policies and guidelines - Legal issues
(Petersen et al., 2018)	- Twitter	- Semi-structured interviews (7)	**Enablers** - Changed communication behaviors - Situational awareness - Helping one another
(Bertot et al., 2012)	-	-	**Challenges** - Privacy - Security - Accuracy - Policy laws
(Simon et al., 2015)	- Twitter - Facebook - Flickr	- Review	**Challenges** - Communication - Data Analysis
(Anderson et al., 2015)	- Twitter	- Interviews	**Challenges** - Data modeling - Space vs. Time tradeoffs
(Comunello and Mulargia, 2017)	- Twitter - Facebook	- Semi-structured interviews (16)	**Challenges** - Organizational Issues - Lack of Personnel - Lack of Legislation

(Wilson et al., 2017)	- Twitter	- Mixed methods	**Challenges** - Language - Rumor spreading **Technical Enablers** - Data Mining - Crisis and Event Recognition - Location Recognition - Identification of False Information - Situational Awareness - Communication to and within the Public - Organizational Use of Communication Systems - Collaboration Systems for Volunteers and Organizations - Collective Sense-Making - Volunteered Geographic Information
(Plotnick and Hiltz, 2018)	- Twitter	- Semi-structured interviews - Survey	**Challenges** - Information Overload - Lack of staff
(Huang et al., 2016)	-	-	**Challenges** - Less Affluent and Lesser Educated People - Infrastructure - Authenticity - Scalability - Security - Privacy - Lack of Local Language Websites - Lack of Basic Computer Skills
(Lambert, 2020)	- Facebook	- Semi-Structured Interviews	**Enablers** - Exchanging Information - Transparency - Trust - Support during disaster
(Lieneck et al., 2022)	-	Theme-based approach	**Challenges** - Spread of misinformation -Decreased vaccine acceptance amongst SM users - Lack of SM regulations

Tab. 3.1: Summary of Challenges and Enablers of Using Social Media in Disaster Management.

3.2 Research Methods and Tools for Qualitative Data Analysis in Social Media for Disaster Management

There are many different methods in qualitative research and various ways to analyze qualitative data. One of the methods being in-depth interview is a research method followed to explore respondents' thoughts, perspectives, and feedback on the ideas and outcomes of the study, enabling more profound and detailed information via intensive discussions, with a lesser number of respondents than surveys (Link et al., 2016). Semi-structured interviews and surveys are research methods used by USA county-level emergency managers to understand the barriers that may prevent the use of SM for DM (Plotnick and Hiltz, 2018). Semi-structured interviews have been conducted in the exploratory study with Italian expert informants and officers operating at municipality and province level to explore the barriers for SM usage in emergency communication (Comunello and Mulargia, 2017). Facebook group discussion is another research method used to communicate risk in an earthquake study investigated for qualitative analysis (Lambert, 2020).

In addition to the research methods, numerous techniques analyze the qualitative data. One well-known process is called grounded theory, comprising inductive and deductive approaches. Inductive approach analyses the data for generating novel insights and theoretical constructs, while on the contrary, the deductive method starts with a theory/ hypothesis, culminating at the confirmation or rejection of the theory/ hypothesis (Badam et al., 2017). Truelove et al. (2019) utilize an inductive thematic approach for QDA by the Snapchat platform use. In another research, Gray et al. (2017) use content analysis to track the changes in storm Desmond of 2015 considering the Twitter platform. A qualitative and a quantitative study by conducting an online survey with closed and open-ended questions to identify the emergency services' attitudes towards the usage of SM. Qualitative analysis of the data is done using an inductive approach of grounded theory along with open coding to categorize the data (Reuter et al., 2016). Considering the Zika health crisis in the United States, a mixed-methods study has been conducted to identify the government's communications engagement with the public on the Twitter platform using the grounded theory approach (Hagen et al., 2018).

Some tools are available for QDA. Atlas.ti is one of the most prominently used software. In a Chinese study on the use of SM for governmental use, qualitative software tool Atlas.ti has been used for coding and analysis activities (Zheng, 2013). Grounded theory analysis using an inductive approach, abductive thematic network analysis, and a deductive critical discourse analysis methods of QDA are discussed using Atlas.ti software, considering Mauritius, Kenya, and Zambia countries (Rambaree, 2012). Computer-Assisted Grounded Theory Analysis with Atlas.ti software is explicated in (Friese, 2016). A research study involving QDA of focus groups and surveys assessing challenges in low-income families to inform a life skills-based obesity intervention has analyzed the results with Atlas.ti software (Bhushan et al., 2017).

In addition to Atlas.ti software, several other QDA software is available, like NVivo, KNIME,QDA Miner, Qualrus, Transna and Leximancer, and HyperRESEARCH. QDA Miner complements Atlas.ti with further quantitative features like cluster, factor, and correspondence analysis (Scientific Software Development GmbH, 2013). Incorporating the NVivo software for qualitative analysis of Twitter data, a research study on risk communication has been done, considering the heavy snow of 2010 and riots of August 2011 (Panagiotopoulos et al., 2016). Challenges on using SM in qualitative research utilize NVivo software to analyze the data (McKenna et al., 2017). (Ahmed, 2011) investigates the potential of using SM in Australian natural disasters, conducting semi-structured interviews and focus groups with disaster officials, and ends up analyzing the data using NVivo software. MAXqda is another software with qualitative, quantitative, and mixed research analysis capabilities. To identify the barriers and wishlists for the usage of SM in DM, QDA Miner has been used in (Hiltz and Kushma, 2014) to code the qualitative data.

3.3 Social Media and Disaster Management Cycle

Communication plays a vital role during disasters. Takahashi et al. (2015) examines communication on Twitter data during and post Typhoon Haiyan in Philipines. Tang et al. (2015) examines the role of SM, i.e., Facebook, Twitter, and YouTube, in historic California drought of 2014. Lambert (2020) investigates Facebook data for disaster risk communication of the Alaska earthquake of 2018. Yang et al. (2019) consider the 2013 earthquake in Ya'an, China, to analyze public emotions on SM. Kirac and Milburn (2018) propose a general framework post-Haiti earthquake of 2010, utilizing data from Facebook, Twitter, blogs, and SMS text messages for disaster response logistics planning. Domdouzis et al. (2016) present a DM system to prevent crime during and post-disaster utilizing Twitter and Facebook data. Agarwal and Toshniwal (2020) identify leadership characteristics during natural disasters using Twitter data of Hurricanes, Typhoons, Floods, and Earthquake.

Squicciarini et al. (2017) investigated the sentiments of tweets during Hurricane Sandy and visualized them on a geographical map. Sit et al. (2019) identify disaster-related tweets considering Hurricane Irma. Basu et al. (2017) map the resources during a natural disaster, considering the Nepal Earthquake of 2015 utilizing WhatsApp messages.

Yuan et al. (2021) have utilized SM data to study disaster resilience considering Hurricane Florence. Kim and Hastak (2018) apply social network analysis to transform the data collected from the Facebook page of the Louisiana flood of 2016 into knowledge by exploring the patterns generated by aggregated interactions of online users after a disaster. Yabe and Ukkusuri (2019) utilize the Hurricane Sandy dataset to study the human behavior of evacuees post-disaster. Dutt et al. (2019) use Twitter data for helping in post-disaster relief operations by matching needs and availabilities.

3.4 Artificial Intelligence for Prediction and Decision-Making

Nguyen et al. (2019) build a framework to forecast people's needs during disasters focusing on spatial, temporal, and atmospheric factors, utilizing enormous SM and weather data. Yabe and Ukkusuri (2019) propose a novel method by integrating information from heterogenous networks on SM to improve the predictability of evacuees' returning behavior post-disaster, utilizing the Twitter dataset of Hurricane Sandy. Singh et al. (2017) propose an algorithm for identifying victims asking for help on Twitter in flood. It achieves a classification accuracy of 81% and location prediction accuracy of 87%. The study utilizes geo-tagged tweets and builds a Markov chain to predict the location. Chen et al. (2015) forecast the appearance and disappearance of smog disasters by analyzing the factors utilizing Twitter data. Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF) are some of the AI-based algorithms are considered for forecasting purposes.

Kibanov et al. (2017) mine tweets to inform Peatland fire and haze DM. In addition, it analyzes online users' behavior during extreme haze events. Chikaraishi et al. (2020) conducted an exploratory study on the applicability of ML algorithms to predict short-term traffic during disasters, considering landslides that occurred in June 2018 in the cities of Hiroshima, Kure and Higashi-Hiroshima. Since landslides ruined the towns, minimizing the congestion is essential for effective disaster response. Neppalli et al. (2016) perform experiments on tweets to study retweetability in the context of Hurricane Sandy so that pertinent information is flown during disasters, stopping the spreading of misinformation. It results in how likely the tweet is to be retweeted using ML approaches.

Ranjan and Gupta (2021) propose an approach automatically detecting fake news on the social networks using DL techniques. The author has considered the Facebook platform and LSTM techniques in the research study. The system has attained considerably higher accuracy than existing state-of-the-art techniques. Qu et al. (2018) present an efficient recommender framework utilizing a Deep Graph-Based Network for friend recommendation on SM platforms, considering Weibo datasets. The author uses Back-Propagation Neural Network to predict accurately.

Zuheros et al. (2021) develop a decision-making model for a restaurant choice considering TripAdvisor reviews utilizing CNN and LSTM models. Subramani et al. (2018) automatically identify domestic violence victims, adopting CNN, RNN, LSTM, bi-LSTM, and Gated Recurrent Units (GRU) models. The model achieves an accuracy of 94% in contrast with traditional ML algorithms. Zhang et al. (2021) propose a model to predict creep, fatigue, and creep-fatigue life predictions under high-temperature conditions. The results illustrate that the DL model Deep Neural Network (DNN) performs better than conventional ML algorithms in prediction accuracy.

Gao et al. (2021) propose a multi-step LSTM method to predict the ship trajectory using AIS data. The method simultaneously predicts the short and long distance and predicts each moment's accurate position, which provides more comprehensive and accurate data for ship collision forecasts. Giannakas et al. (2021) propose a DNN classification framework to predict the team's academic performance in software engineering. The authors have used different activations functions, i.e., (Sigmoid, ReLU, and Tanh), and optimizers (Adagrad, Adadelta). The results show the superior performance of the framework with 86.57%.

3.5 Applications of Deep Learning for Disaster Management

Researchers have employed DL models for diverse applications in the past few years, including DM. Nguyen et al. (2019) forecast people's needs considering Hurricane disasters, utilizing LSTM and CNN models. The models are implemented on Hurricane Sandy'12, Hurricane Harvey'17, and Hurricane Irma'17 tweets. Ofli et al. (2020) utilize CNN architecture on a real-world disaster-related dataset gathered from Twitter, i.e., CrisisMMD for disaster response. Muhammad et al. (2019) propose a computationally efficient CNN-based architecture for fire detection, localization, and semantic understanding of fire scenes.

Caragea et al. (2016a) identify the flood informative messages using CNN from SM data. Pereira et al. (2020) assess the severity of floods by exploring SM data with deep CNN. Tian et al. (2018) employ the LSTM model to obtain sequential semantics for disaster-related classification. Zhai et al. (2020) adopt the sentiment analysis to assess the situational awareness in Hurricane Florence, using CNN and Latent Dirichlet Allocation (LDA) approach. Huang et al. (2020b) explore disaster-related tweet posts for rapid disaster response, utilizing CNN architecture.

Zeroual et al. (2020) present a comparative study employing RNN, LSTM, bi-directional LSTM, GRUs, and Variational AutoEncoder (VAE) for worldwide forecasting of Covid-19 cases. The author considered Italy, Spain, France, China, the USA, and Australia countries for daily confirmed and recovered cases. The results highlight the superiority of VAE when compared to other techniques. Liang et al. (2020) propose a DL-based survival model to predict the risk of developing critical illness at the time of Covid-19 patient's admission. The model is constructed using three-layer feedforward neural network for survival modeling.

Chikaraishi et al. (2020) conducted an exploratory study on the possibility of short-term traffic during heavy rain and subsequent landslides in Hiroshima in July 2018. The authors considered RF, SVM, XGBoost, deep feed-forward neural network, and shallow feed-forward neural network. Amin and Ahn (2021) explore the CNN algorithm to teach earthquake disaster avoidance. Chaudhuri and Bose (2020) explore the role of DL algorithms for post-disaster decision support by collecting data from earthquake-hit regions. The experimental results reveal that DL algorithms perform better than ML algorithms. Table: 3.2 presents a tabular format of the works pertaining to SM usage in DM utilizing DL methods.

Article	Aim of the Work	Disasters	Methods Used	Evaluation Metric
Chakraborty et al. (2020)	Sentiment Analysis	Covid-19 Tweets	- LR, - Multinomial Naive Bayes (NB), Linear SVM, - RF, and - AdaBoost classifiers	- Accuracy - 0.81
Sit et al. (2019)	Identify disaster-related tweets and their semantic, spatial, and temporal context	Hurricane Irma	- LSTM - LDA	- Accuracy - 0.7478, - F1-score - 0.7514
Snyder et al. (2019)	Identify relevant tweets to support real-time situational awareness	-Colorado Wildfires - Boston Bombings - NY Train Crash	-CNN - RNN -LSTM	Avg F1-score - 0.75
Basu et al. (2019)	Extracting resource needs and availabilities for aiding post-disaster operations utilizing SM	- Nepal Earthquake 2015 - Italy Earthquake 2016	- CNN - SVM - NB - Gradient Boosted Decision Tree (DT)	F-score @ 100 - 0.191
Yu et al. (2019)	Utilizing DL for real-time situational awareness	Hurricane Irma, Sandy, and Harvey	- CNN - SVM - LR	Accuracy - 0.81
Kundu (2018)	Classification of tweets during disasters on need and availability based	- FIRE 2016 - FIRE 2017	- CNN - Deep-LSTM	F-score - 0.9159
Devaraj et al. (2020)	Identifying SM based requests for help	Hurricane Harvey	- CNN - SVM - LR - NB - DT - Ridge Classifier - Multilayer Perceptron	F-score -0.86

Tab. 3.2: Comparison of the relevant works pertaining to Disaster Management. The insights from these papers are used in the experiments conducted in our research work.

3.6 Fusion Deep Learning Model

Zhang et al. (2020b) propose a novel methodology for predicting and analyzing financial time series by constructing a hybrid prediction model for stock markets. The author has presented Complementary Ensemble Empirical Mode Decomposition- Principal Component Analysis- LSTM networks. Petmezas et al. (2021) propose a novel hybrid neural network model to detect Atrial Fibrillation using ECG datasets. The authors extract the features via CNN and input them into the LSTM network, achieving a sensitivity of 97.87% and specificity of 99.29%.

Abbasimehr and Paki (2020) propose a fusion model, combining LSTM, CNN, and Bayesian optimization algorithms, to predict the confirmed Covid-19 cases. The experimental results exhibit the superiority of the DL models against the benchmark model. Chen et al. (2019) propose a hybrid CNN-LSTM model to forecast typhoon formation. Extensive experiments show that the proposed model outperforms LR, LDA, RF, and DT.

Li et al. (2021) propose a hybrid DL model, i.e., CNN-LSTM, to predict next-day traffic flow. The results indicate the superior performance of the proposed model against CNN, LSTM, Multilayer Perceptron (MLP), and statistical model ARIMA. Ajao et al. (2018) propose a framework for fake news identification from the Twitter platform using a hybrid of CNN and RNN models. The proposed work achieves an accuracy of 80%. Zhu et al. (2020) propose a combination of CNN-LSTM model to classify human activities based on Micro-Doppler Radar, achieving an accuracy of 98.28%.

Li et al. (2020) construct a CNN-LSTM hybrid model to forecast the next 24 hours PM2.5 concentration in Beijing. The model extracts the features related to air quality with CNN and LSTM reflects the long-term dependencies from input time series data. The proposed model performs excellently. Another study Alhussein et al. (2020) offers a CNN and LSTM combination model for forecasting individual household electric loads. The proposed hybrid model has shown an improvement of 5.98% compared to the LSTM model. Kanjo et al. (2019) employ the CNN-LSTM hybrid model to classify human emotions, outperforming the traditional fully connected neural networks, with an average accuracy=83% and F-score=82%.

3.7 Deep Learning-Based Studies using Social Media for Disaster Management

Behl et al. (2021) explore the Twitter platform, considering Nepal-2016 and Italy-2016 earthquake for disaster relief operations through sentiment analysis. The research study achieves an accuracy=83% on the Covid-19 dataset. Nguyen et al. (2016) propose a DNN utilizing CNN to identify informative tweets. Burel and Alani (2018) introduce an open-source web API, i.e., Crisis Event Extraction Service, to classify crisis-related events from SM. The API is backed with CNN and is validated against traditional ML algorithms.

Xia et al. (2020) extract and acquire the outlier knowledge of Covid-19 data to understand the public opinions based on microblog data. The paper utilizes LDA architecture for knowledge acquisition, achieving an accuracy of 89.9%. Yang et al. (2019) introduce a DL method to extract fine-grained sentimental information of citizens from Chinese SM in disaster analysis, considering the Ya'an earthquake of 2013. The research study attains an F1-score of 88.89%.

Table: 3.3 reviews some of the DL-based research studies utilizing Twitter data for DM.

Article	**Aim of the Work**	**Disasters Considered**	**Methods Used**
Sit et al. (2019)	Identify disaster-related tweets and their semantic, spatial, and temporal context	Hurricane Irma	- LSTM - LDA
Snyder et al. (2019)	Identify relevant tweets to support real-time situational awareness	-Colorado Wildfires - Boston Bombings - NY Train Crash	-CNN - RNN -LSTM
Habdank et al. (2017)	Relevancy Assessment of tweets	Fire incident, Ludwigshafen, Germany	-NB; SVM; RF; DT; Neural Networks (NN)
Caragea et al. (2016b)	Identify disaster-relevant tweets	- Philipinnes floods(2012); - Colorado floods (2013); - Queensland floods(2013); - Sardinia floods (2013); - Alberta floods (2013); - Manila floods (2013)	CNN
Nguyen et al. (2016)	Identify informative tweets for crisis response	Nepal earthquake (2015)	CNN
Abel et al. (2012a,b)	Filter disaster-relevant information from social web streams	Everyday-life incidents in Netherlands	-
Feng and Sester (2018)	Extraction of flood-related content from Twitter	Pluvial Floods	CNN; LR; NB; SVM; RF

Tab. 3.3: Comparison of the relevant works pertaining to Disaster Management.

3.8 Usage of Social Media in Different Stages of Disaster Management Cycle

SM is a virtual commonplace amongst all the domains, including DM. Conventionally, DM is conceptualized into three periods, i.e., pre, during and post-disaster. SM plays multiple roles in DMC. Disasters are of two types- natural, and man-made. According to National Disaster

Management Authority (2019), *natural disasters* include landslides, tsunamis, drought, floods, earthquakes and volcanic eruptions, cyclones, thunderstorms. We have considered the literature that demonstrates how SM relates to different periods and phases of DMC.

The importance of SM in disasters has been reviewed by several scholars Akter and Wamba (2017); Gaspar et al. (2019); Goswami et al. (2016); Simon et al. (2015). Abedin et al. (2014) review the usage of SM data in all the phases of DMC: prevention (mitigation), preparedness, response, and recovery. Sarker et al. (2020) review disaster resilience through big data. The study reveals big data technology, i.e., SM usage in different phases of DMC. Tarasconi et al. (2017) explore SM usage for natural disasters, considering three phases of DMC- awareness/early warning, response, post-disaster. Oktari et al. (2020) reviews the use of KM practices applied in all the phases of DMC. Kaewkitipong et al. (2016) conducted a study to analyze the information sharing activities using SM within local communities to deal with the Thai flood crisis of 2011. The authors consider all the periods of disaster.
Takahashi et al. (2015) examines communication on Twitter data *during* and *post* Typhoon Haiyan in Philipines. Squicciarini et al. (2017) investigated the sentiments of tweets *during* Hurricane Sandy and visualized them on a geographical map. Sit et al. (2019) identify disaster-related tweets considering Hurricane Irma. Basu et al. (2017) map the resources *during* a natural disaster, considering the Nepal Earthquake of 2015 utilizing WhatsApp messages.

Yuan et al. (2021) have utilized SM data to study disaster resilience *during* Hurricane Florence. Kim and Hastak (2018) apply social network analysis to transform the data collected from the Facebook page of the Louisiana flood of 2016 into knowledge by exploring the patterns generated by aggregated interactions of online users post-disaster. Yabe and Ukkusuri (2019) utilize the Hurricane Sandy dataset to study the human behavior of evacuees *post-disaster*. Dutt et al. (2019) use Twitter data for helping in *post-disaster* relief operations by matching needs and availabilities.

In the past few years, researchers have employed DL models for diverse applications, including DM. Nguyen et al. (2019) forecast people's needs *during* Hurricane disasters, utilizing LSTM and CNN models. The models are implemented on Hurricane Sandy'12, Hurricane Harvey'17, and Hurricane Irma'17 tweets. Ofli et al. (2020) utilize CNN architecture on a real-world disaster-related dataset gathered from Twitter, i.e., CrisisMMD for disaster response. Caragea et al. (2016a) identify the flood informative messages using CNN from SM data *during* disaster events. Tian et al. (2018) employ the LSTM model to obtain sequential semantics for disaster-related classification. Zhai et al. (2020) adopt the sentiment analysis to assess the situational awareness in Hurricane Florence, using CNN and LDA approaches. Huang et al. (2020b) explore disaster-related tweet posts for rapid disaster response, utilizing CNN architecture. Table: 3.4 presents a tabular format of the works pertaining to SM usage in DM utilizing DL methods.

3.9 Social Media Based Systems for Disaster Management

Nguyen et al. (2016) propose a DNN utilizing CNN to identify informative tweets. Yang et al. (2019) introduce a DL method to extract fine-grained sentimental information of citizens from

Article	**Aim of the Work**	**Disasters**	**Methods Used**	**Evaluation Metric**	**Period**
Chakraborty et al. (2020)	Sentiment Analysis	Covid-19 Tweets	- LR, - Multinomial NB, Linear SVM, - RF, and - AdaBoost classifiers	- Accuracy - 0.81	During
Sit et al. (2019)	Identify disaster-related tweets and their semantic, spatial and temporal context	Hurricane Irma	- LSTM - LDA	- Accuracy - 0.7478, - F1-score - 0.7514	During; Post
Snyder et al. (2019)	Identify relevant tweets to support real-time situational awareness	-Colorado Wildfires - Boston Bombings - NY Train Crash	-CNN - RNN -LSTM	Avg F1-score - 0.75	During
Basu et al. (2019)	Extracting resource needs and availabilities for aiding post-disaster operations utilizing SM	- Nepal Earthquake 2015 - Italy Earthquake 2016	- CNN - SVM - NB - Gradient Boosted DT	F-score @ 100 - 0.191	Post
Yu et al. (2019)	Utilizing DL for real-time situational awareness	Hurricane Irma, Sandy, and Harvey	- CNN - SVM - LR	Accuracy - 0.81	During, Post
Kundu (2018)	Classification of tweets during disasters on need and availability based	- FIRE 2016 - FIRE 2017	- CNN - Deep-LSTM	F-score - 0.9159	During
Devaraj et al. (2020)	Identifying SM based requests for help	Hurricane Harvey	- CNN - SVM - LR - NB - DT - Ridge Classifier - Multilayer Perceptron	F-score - 0.86	During
Chowdhury et al. (2013)	Tweet4act: Using Incident-Specific Profiles for Classifying Crisis-Related Messages	Small-scale incidents	Dictionary-based period classifier; SVM; DT; RF; MaxEnt	Precision (P)=0.69, Recall (R)=0.93, F-score= 0.71	Pre, During, Post

Tab. 3.4: Comparison of the relevant works pertaining to stages of DMC. The insights from these papers are used in the experiments conducted in our research work.

Chinese SM in disaster analysis, considering the Ya'an earthquake of 2013. The research study attains an F1-score of 88.89%.

In Table: 3.5, we explain the summary for studies relevant to the Covid-19 disaster, with SM-based papers utilizing DL models on Covid-19 related datasets. We have tabulated the reference of paper, published year, main discussed content of the article, results, and limitation of the paper. Naseem et al. (2021) analyze the Covid-19 tweets and present a new large-scale sentiment dataset CovidSenti, consisting of 90,000 tweets from February 2020 to March 2020. The authors classify the sentiment of Covid-19 tweets utilizing CNN, BiLSTM, SVM, DT, NB, and RF, achieving an accuracy of 0.869. Behl et al. (2021) use the Twitter platform for disaster response through sentiment analysis, considering Covid-19 and natural hazard crises. The study achieves an accuracy of 83%.

Chakraborty et al. (2020) analyze the sentiments of Covid-19 tweets considering multinomial NB, SVM, LR, RF, and Adaboost. The proposed model achieves an accuracy of 81%. Jelodar et al. (2020) classifies the sentiments and discovers topic in Coronavirus online discussions on Reddit. It reaches an accuracy of 81.5% with the LSTM, and LDA. Abdelminaam et al. (2021) automatically detect fake news on Twitter, utilizing modified-LSTM and modified-GRU. The authors achieve an accuracy of 83.32%, P of 84.45%, R of 83.32%, and F1-score of 83.7%.

There is a dearth of automated SM-based DM systems. However, after various cataclysms, a few applications have been developed. In Table: 3.6, we explain the existing SM-based systems of DM. We have tabulated the paper's reference, published year, the paper's main aim, platform considered, methods used, results, limitations, datasets, and country of implementation of the system. Yasin Kabir et al. (2020) presents STIMULATE, i.e., real-time information acquisition and learning platform for DM. It employs a Twitter platform, considering Hurricane Harvey and Irma datasets, and deployed in the United States of America. Morrow et al. (2011) explore the Ushahidi platform, which utilizes the Twitter platform after the earthquake Haiti in 2010. Emergency situations can be tracked using the platform with a geographical map.

Irmamiami [1] is an application to depict different activities in the Miami area related to Hurricane Irma. The platform considers Twitter, web, email, and SMS to generate disaster-based reports. It utilizes the Ushahidi source code. Wladdimiro et al. (2016) propose a platform to support real-time data in disaster scenarios with the Twitter platform. The authors build a case study to detect the need-based messages in an earthquake at Santiage de Chile. Imran et al. (2014) present a platform to classify the tweets as *informative* or *non-informative*, considering the 2013 Pakistan earthquake, and deployed in Qatar. Burel and Alani (2018) introduce an open-source web API, i.e., Crisis Event Extraction Service, to classify crisis-related events from SM automatically.

[1] https://irmamiami.ushahidi.io/views/map

Study	Aim	Platform	Methods	Results	Limitations
Naseem et al. (2021)	Sentiment analysis of Covid-19 related tweets	Twitter	CNN; BiLSTM; SVM; DT; NB; and RF	Accuracy = 0.869 (86.9%)	Ony accuracy metric is used for evaluation.
Behl et al. (2021)	Twitter for disaster relief through sentiment analysis	Twitter	MLP; CNN; and LR	Accuracy 83%	Under-performance because of the limitations on smaller training dataset.
Chakraborty et al. (2020)	Sentiment analysis of Covid-19 tweets	Twitter	Multinomial NB; SVM; LR; RF; and Adaboost	Accuracy 81%	Study is limited to only one DL technique
Jelodar et al. (2020)	Sentiment classification and topic discovery of Corovirus online discussions	Reddit	LSTM; LDA	Accuracy 81.5%	Study limited to only one DL technique
Abdelminaam et al. (2021)	Automated detection of misleading Covid-19 related information	Twitter	Modified-LSTM; and Modified GRU	Accuracy= 83.82%; P= 84.85%; R = 83.82%; and F1-score = 83.7%	Only one SM platform is considered.

Tab. 3.5: Studies pertaining to Covid-19 and SM.

Study	Aim	Platform	Methods	Results	Limitations	Datasets	Country
Yasin Kabir et al. (2020)	STIMULATE - a System for Real-time Information Acquisition and Learning for DM	Twitter	BiLSTM, and CNN	Hurricane Harvey, and Hurricane Irma	The system considers flood disasters, with need-based tweet classification.	Hurricane Harvey and Irma; Nepal Earthquake; California Earthquake; Typhoon Hagupit; Cyclone PAM; CrisisLex	United States
Morrow et al. (2011)	Ushahidi	Twitter	-	Emergency situations can be tracked with geographical map	Limitation in information use from UHP website	Earthquake	Haiti
Wladdimir et al. (2016)	Platform to support real-time data in disaster scenario.	Twitter	-	Basic supplies like water, and electricity; disaster information; crimes, and looting; and missing people information.	Only one disaster is considered.	Earthquake	Santiago de Chile

Study	Aim	Platform	Methods	Results	Limitations	Datasets	Country
Irmamiami[2]	Depicts different activities in Miami area related to Hurricane Irma	Twitter; web, email, sms	-	Disaster-based reports	Only one disaster Hurricane is considered.	Hurricane Irma	United States
Imran et al. (2014)	Platform to automatically classify crisis-related microblog information.	Twitter	-	Classification of messages as *informative* , and *non-informative*.	Earthquake is the only considered disaster.	2013 Pakistan Earthquake	Qatar
Burel and Alani (2018)	Crisis event extraction service (CREES) – Automatic detection and classification of crisis-related content on SM	Twitter	CNN	Three types of reports: crisis vs non-crisis, type of crisis, and type of information	Performance drops identifying fine-grained event-related information	CrisisLex T26	United Kingdom

Tab. 3.6: Existing Social Media Based Systems for Disaster Management.

The identified research gaps after reviewing literature are as follows:

1. While data are available on challenges and enablers related to SM usage for DM, the existing challenges constitute social, technological, informational, and organizational perspectives. Enablers comprise decentralized decision-making, changed, i.e., rapid communication behavior, helping one another, identification of false information, ease of exchanging information, transparency, trust, and support generation during a disaster. There is a dearth of data in the Indian context despite it all. No published articles found that focus groups' viability to investigate the challenges and enablers of SM usage for DM in the Indian scenarios utilizing grounded theory. Therefore, the first objective is identified as to determine the challenges and enablers in the use of SM for DM.

2. Although DL techniques have received considerable attention during the last decades, most studies focus on a single DL technique to identify the disaster-related SM message. With the growth in data and the improvement of demand, the DL model built for data identification is increasingly complex. It is no longer a single model but a more complex hybrid model. The lower efficiency of existing methods could not meet the demand in DM yet. Meanwhile, due to the difficulty of the DL model in stability, it is essential to develop a more effective model to identify disaster-related SM message. The second objective is the

result of this research gap, i.e., developing a framework to identify the relevancy of a SM message to a disaster.

3. The existing works detect period of the disaster using ML techniques. The approach lacks at capturing non-linearity of the data. Moreover, the high complexity of the SM-based data is also a major issue. DL techniques have been found to be the most effective in dealing with non-linearity of data. Therefore, a DL-based framework is required to determine the stage of disaster from a SM message. The third objective is the result of this research gap, i.e., developing a framework to determine the stage of DMC from a SM message.

4. There are a few observed issues and potential pitfalls in interpreting recently published work. None of them [Table: 3.6] have proposed a web interface to investigate SM-based messages for employing the nature of the message. In addition, the existing works have specified their analysis as specialized to particular regions and cannot be generalized to those approaches globally. Regarding Covid-19 studies in Table: 3.5, most times, they focus on sentiment analysis or fake news detection. However, they did not pay attention to determine the nature of the SM message, due to which the most significant messages, i.e., help-seeking and help-offering, are not explored. Hence, it is difficult to gain an understanding of the disaster situation according to this perspective. Although the existing applications provide partial functionalities to collect and process the SM data, they lack decisive components for an automatic SM-based DM system that can detect help-seeking people, assistance offering people from SM posts and update regarding situational information as well as general opinions location-wise. Hence, there is a need of web-based DM system that can identify the nature of the SM message, locationwise. The forth objective is born out of this research gap, i.e., developing a web-based DM system to determine the nature of SM message for decision-making.

4. CHALLENGES AND ENABLERS FOR USING SOCIAL MEDIA FOR DISASTER MANAGEMENT

4.1 Introduction

SM is the lifeline of the current generation. It has connected the world like never before. It has led the citizens to behave more as partners rather than passive public service consumers (Linders, 2012). It has opened new channels of communication. According to Global Digital Report 2021 (Simon Kemp, 2021), the global population was 7.83 billion at the start of 2021, and 5.22 billion people are using mobile phones, equating to 66.6% of the total population. At this moment, 4.66 billion people use the internet, i.e., 59.5%. The same report pinpoints the number of SM users, mentioning that 4.20 billion active SM users are there, i.e., 53.6% of the population. The same report pinpoints that 98.8% of SM users access via mobile phones.

The usage of SM is mushrooming post-haste in different domains (Kanagarajoo, 2018; Mols and Pridmore, 2019; Subramani et al., 2018; Truelove et al., 2019). According to authors Kavota et al. (2020), only 8% of DM cases have used SM, and it is the lowest in the available domains. As a reminder, a disaster is an event, natural or man-made, sudden or progressive, that negatively impacts society, affecting the lives and infrastructure. It is worth mentioning that the terms' disaster' and' emergency' are used synonymously in the research study.

According to the Centre for Research on the Epidemiology of Disasters and United Nations Office for Disaster Risk Reduction (2021), there are 389 reported disasters in 2020. In addition, disasters claim the lives of 15,080 people, 98.4 million people are affected globally, and US171.3 billion dollars are spent on economic damage. Numerous authors have studied the relationship between SM and DM and have highlighted the importance of SM for DM on a positive note (Kim et al., 2016; Snyder et al., 2019; Yuan et al., 2021). Despite the various benefits of SM in daily life and crisis, no study has been conducted on the identification of challenges and enablers of SM usage for DM. The present study tries to do so by conducting an FGD. In other words, this study seeks to bridge the knowledge gap by exploring SM usage for DM, as applied in India, which faces multiple disasters.

The central argument of this research is that the DM is one of the critical domains of national as well as international interest that attracts SM attention, but it does come with idiosyncratic challenges as well as enablers. They need to be identified with an in-depth analysis in the Indian context, and a deep understanding of the particular issues faced during DM. The chapter is important for dissertation as it brings together researchers from multidisciplinary fields of DM for better decision-making and problem solving in the event of a disaster. It closes the gap between those who suffers from the disaster and who are capable of providing relief to the victims by contributing to managerial thinking. This study intends to advance the knowledge about SM usage in disasters and how SM technology fits into DM practices. The Research Questions (RQs) are as follows:

1. What is the role of SM technology in DM?
2. What are the challenges faced in SM usage for DM?
3. What enables SM usage in disaster?
4. How does the usage of SM differ from traditional media that take place in DM?

To the best of our knowledge, there is no research study identifying the challenges and enablers of SM usage for DM. In addition, no approach considers the FGD data utilizing inductive thematic analysis for the aforementioned concern, which is the novelty of our approach.

FGD has been employed to conduct the study as it gathers fruitful data from participants' rich experience working in the DM domain. It helps in attaining an in-depth understanding of the study. The grounded inductive approach is chosen in contrast to the deductive method. The inductive approach aims to generate a theory grounded in the data without any preconceived notions. It focuses on exploring new phenomena or looking at a different perspective on previously researched phenomena.

We believe that this is the first study to describe the process and results of the FGD for DM by demonstrating the breadth of barriers and enablers that can emerge from FGD implementation; offer a methodological contribution that can enable trans-disciplinary research outcomes and provide practical guidance for related outreach.

This chapter seeks to add expert, and institutional understanding to the body of evidence, regarding SM usage during disasters, limiting the solutions to overcome these barriers through QDA with expert informants. The study is important internationally due to the uniqueness of this country's DM system, with idiosyncratic challenges and enablers on SM usage in DM.

Following the introduction, the rest of the chapter is organized as follows. Section: 4.2[55] explains the methodology used in the research. Section 4.3 [Page 62] elucidates a discussion on the findings. The study is concluded in section: 4.4 [Page 70].

4.2 Methodology

This is a qualitative FGD study. The aim is to explore and understand the disaster professionals' views about SM usage during disasters and the factors that led them to have that perception. To accurately capture this kind of information, a qualitative study is an appropriate choice. The flow chart of the research process is depicted in Figure 4.1. The existing literature falls short of focusing specifically on the novel and complex factors which act as challenges and enablers of SM for DM.

The methodology developed is in a concoction of existing literature and the efficacy of FGD QDA using Atlas.ti software following inductive thematic approach. The study aims at the inductive qualitative research using FGD to provide valuable insight into the complex and little-known challenges and enablers of the use of SM in DM. FGD technique was chosen as opposed to other methods because it allows the researcher to ask questions, probe for clarifications of ideas, and directly converse during the session.

The participants chosen for FGD are homogeneously working in disaster management domain but are from heterogeneous backgrounds like civil, architecture, mechanical, management, and computer science. The transcripts are transcribed manually by the moderator. After acquiring validation from the participants, raw data is categorized using an inductive thematic approach in Atlas.ti software. Atlas.ti software has been chosen over other tools as it offers a helpful flat code list while working on the inductive thematic approach. The results are finalized after expert validation. To maximize credibility, confirmability, and dependability, a reflective journal was kept by the researcher that documented the decisions, thoughts, and assumptions that shaped the research process throughout the study.

4.2.1 Conduct of FGD and Data Collection

FGD is a widely used qualitative research method that captures extempore comments, unvarnished remarks, and sundry firsthand descriptions through group dynamics. The stepwise process of conducting FGD is depicted in Figure: 4.2. The primary researcher contacted the scholars and professors working in the DM area to recruit the participants. The participants were chosen from the Centre of Excellence in Disaster Mitigation and Management (CoEDMM) of the Indian Institute of Technology Roorkee (IITR), as it works on a diverse range of disasters. Luckily, all the contacted persons agreed to become participants of FGD. Participants share their stance in their own words providing context and nuance.

FGD is conducted with 10 participants [Table 4.1] ranging from age 21 to 42 years in the English language at the CoEDMM, IITR, India. The typical size of a FGD is 6-12 participants. The technique is unique due to its ability to generate data based on the synergy of the group interaction, to discuss differences between participants. In addition, it gives equal time and opportunity to project one's perspective Eeuwijk and Angehrn (2017). 8 out of 10 are male participants, and 2 are female, all from India. 5 out of 10 participants were first-year Master of Technology students of Disaster Management. 2 of them were second-year Master students of DM. The remaining participants were doctoral first year, second year, and a professor in DM of CoEDMM. The

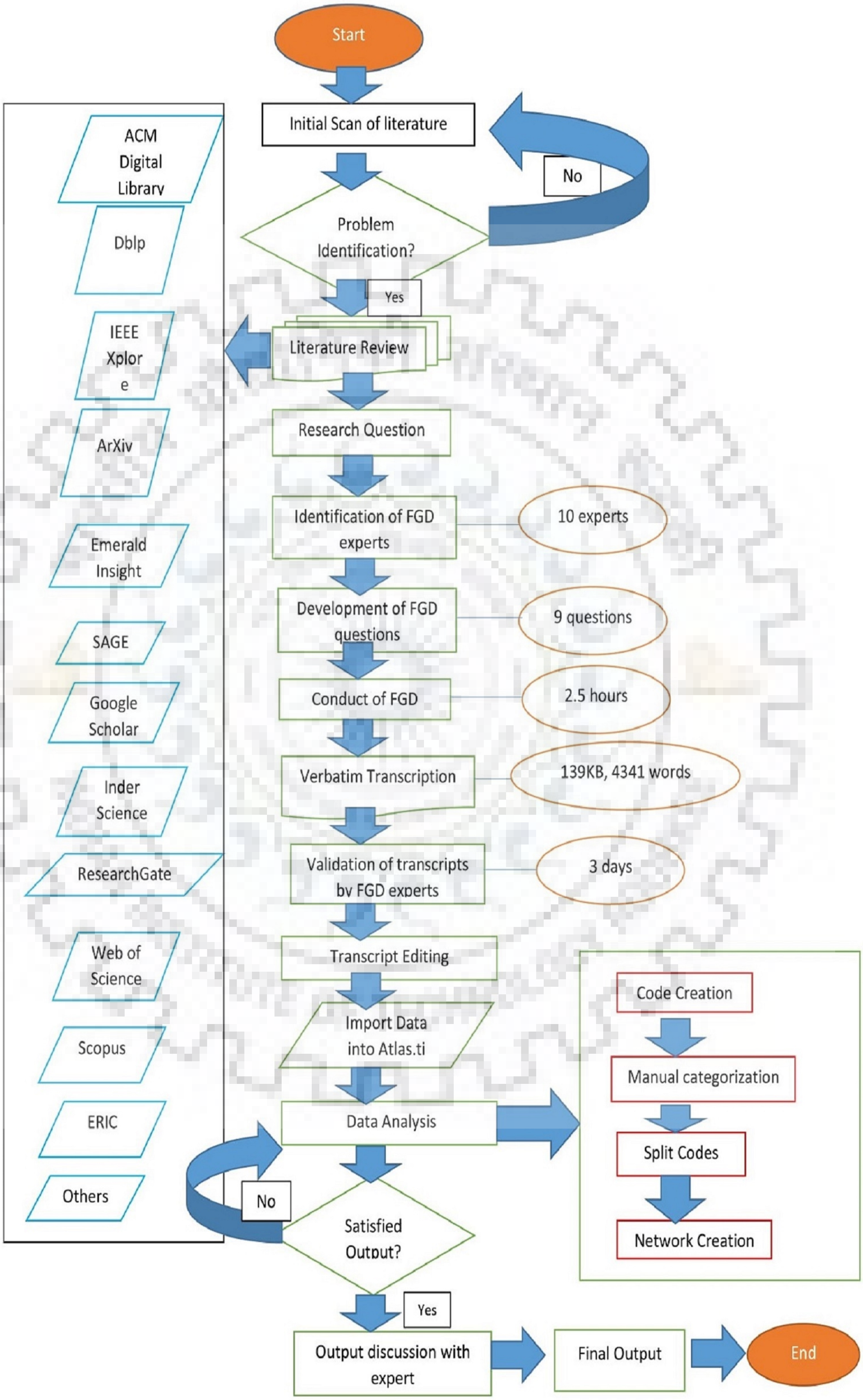


Fig. 4.1: Flow Chart of the Research Process.

<table>
<tr><td>Age</td><td>21-42 years</td></tr>
<tr><td>Gender</td><td>Male: 8
Female: 2</td></tr>
<tr><td>Year of study</td><td>Master's1styear: 5
Master's 2^{nd} year: 2
Doctoral 1^{st} year: 1
Doctoral 2^{nd} year: 1
Professor: 1</td></tr>
</table>

Tab. 4.1: Participant Characteristics (N = 10).

background of the participants was in civil engineering, architecture, computer science, chemical engineering, management studies. Master students with civil background worked in natural hazards like landslides, earthquakes, and cyclones. Participants with architecture background work in urban flooding.

Similarly, a second-year doctoral participant with a computer science background was working in the algorithmic design of DM systems. The professor with management studies had expertise in DM's financial and managerial aspects. The moderator is the researcher and the note-taker as well. There were no ethical issues identified during data collection. The moderator develops the questions for FGD.

The FGD lasted for 2.5 hours. At the start of the session, the moderator explains the purpose of the discussion and initiates the formal consent process, including the permission to digitally video-record the session for later transcription purposes. It is worth mentioning that the primary author takes care of participants' privacy, including recording pieces of equipment and participant information sheets.Participants filled out a small questionnaire regarding their background. The structured focus group agenda begins with baseline queries about SM and DM. The participants are given questions in the form of hard copy at the start of the session and a time of 15-20 minutes to recollect their thoughts on the topic.

Table 4.2 mentions the questions of FGD. Questions are formulated corresponding to the objectives of the research study. The questions are motivated to study the overlapping integration of DM, SM, and disaster informatics. Government is the decision-maker in DM processes. Hence, understanding the Indian government's capability in using SM for DM is required to have an all-inclusive Indian context. FGD was conducted in 3 phases. The first four questions in Table 4.2 are used to increase participation and last for around 50 minutes. The next phase comprised of Questions 5-7 which lasted for 1 hour. The last phase was of Questions 8-9 and lasted for 30 minutes, followed by conclusion remarks by the author. Initially, there were general views, but with the moderator's intervention, there were brainstorming arguments with increased comfort and participation in repeated rounds to reach a common convergence point. The moderator encouraged diverging viewpoints and equal participation by following clockwise and anti-clockwise sequences in FGD. The moderator progresses to the next question when information gained in

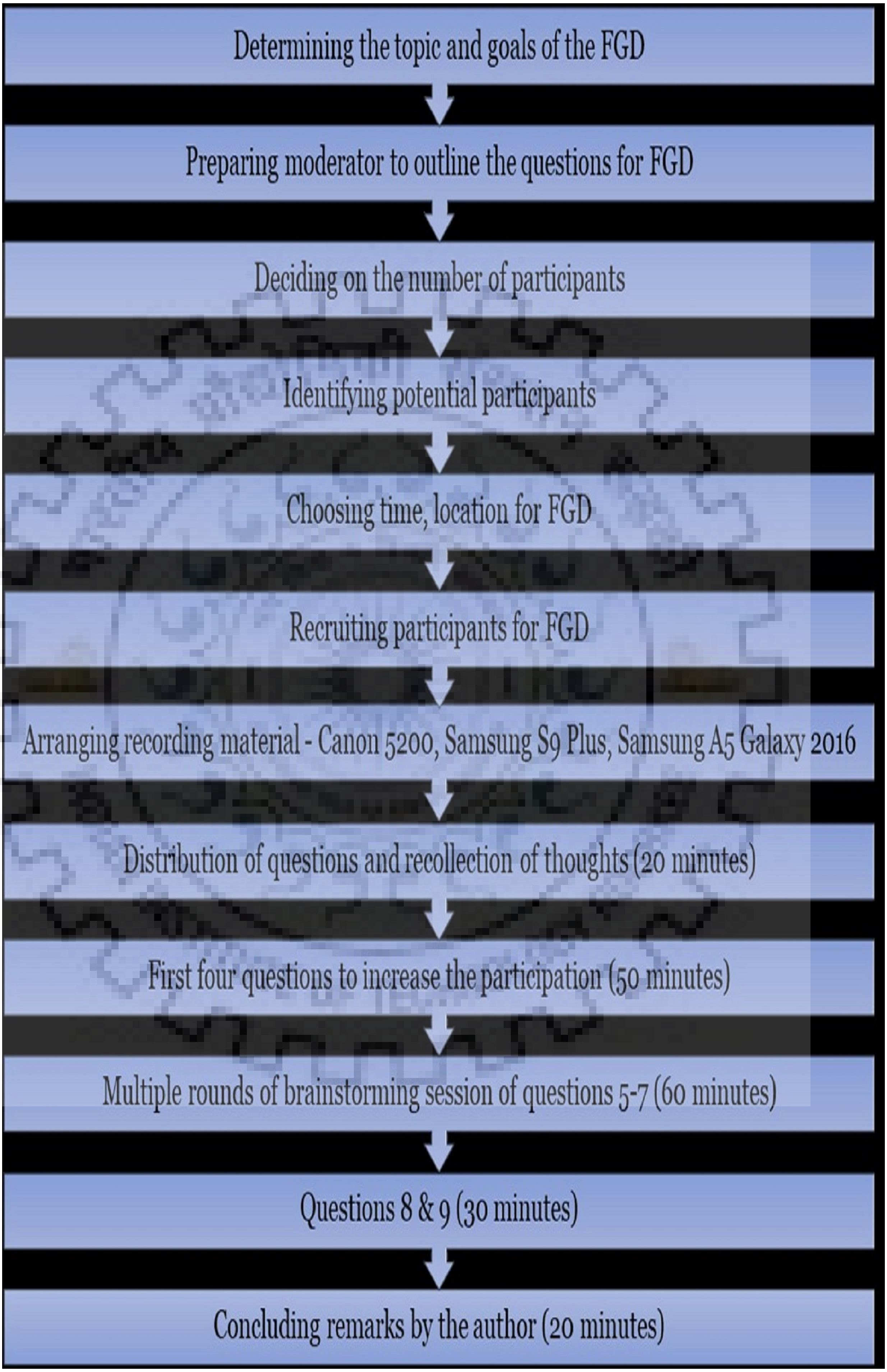


Fig. 4.2: Stepwise Process to Conduct FGD.

the discussion reaches the saturation state, where no new information is being added to the data.

Sr No.	Question
1.	How familiar are you with SM?
2.	Which platforms do they use most often?
3.	How do you feel if SM can be used in an emergency?
4.	Have you used it in an emergency? If yes, please explain the kind of emergency.
5.	What are the challenges/ barriers to using SM in DM?
6.	What do you think are the enablers of SM use in disaster?
7.	What motivates you to use SM for DM?
8.	What do you have to say on its implementation w.r.t. other traditional media?
9.	Do you think our Indian government agencies/ NGOs are capable of using SM for DM effectively?

Tab. 4.2: Questions for FGD.

4.2.2 Data Analysis

Figure: 4.3 illustrates the data analysis employed in the research study using the grounded theory following an inductive approach. The FGD has been video recorded using Canon 5200, Samsung S9 Plus, and Samsung A5 galaxy 2016 with the permission of the participants and transcribed verbatim in a week yielding nine pages (4341 words, 139 KB in size) of raw data. The transcription was sent to all the participants to check if the transcription was verbatim, and they were given the time of 3 days.

After the validation by participants, the transcripts are converted into Atlas.ti format. The transcribed data is systematically coded and qualitatively analyzed using ATLAS.ti 8 software in an iterative process using an inductive approach. ATLAS.ti 8 helps create a 'visual play-ground' with codes, categories, concepts, and networks to build the logico-empirical patterns that emerge from the collected data. The analysis results are presented in a graphical layout [Figures 4.4 - 4.8].

The participants' notes are analyzed first to represent an individual's views. Next, the FGD data is coded, and the initial codes are assigned. The analysis starts with open coding, a process by which all potentially useful and relevant data for answering the RQs are coded. The clusters of similar codes are viewed as Network View to create themes of concepts for eventually looking at a logical pattern that explicates the challenges and enablers on the use of SM for DM. This gives the liberty to compare all codes, inspect frequency of mention, remove redundancies, merge similar-meaning codes, and group codes into categories. Careful consideration is taken to avoid duplication of themes noted by individual participants.

Using the panoply of coding functions - Open coding, Quick coding, List coding, and In-vivo coding - available in Atlas.ti 8, the data is categorized into themes. Once a coding scheme is developed, we check the codes to support code generation and clustering. We organize and present the nodes using 'Network Function' to create categories of concepts for eventually looking at a logical pattern. The 'Network View' function helps build the complex phenomena, deep-rooted

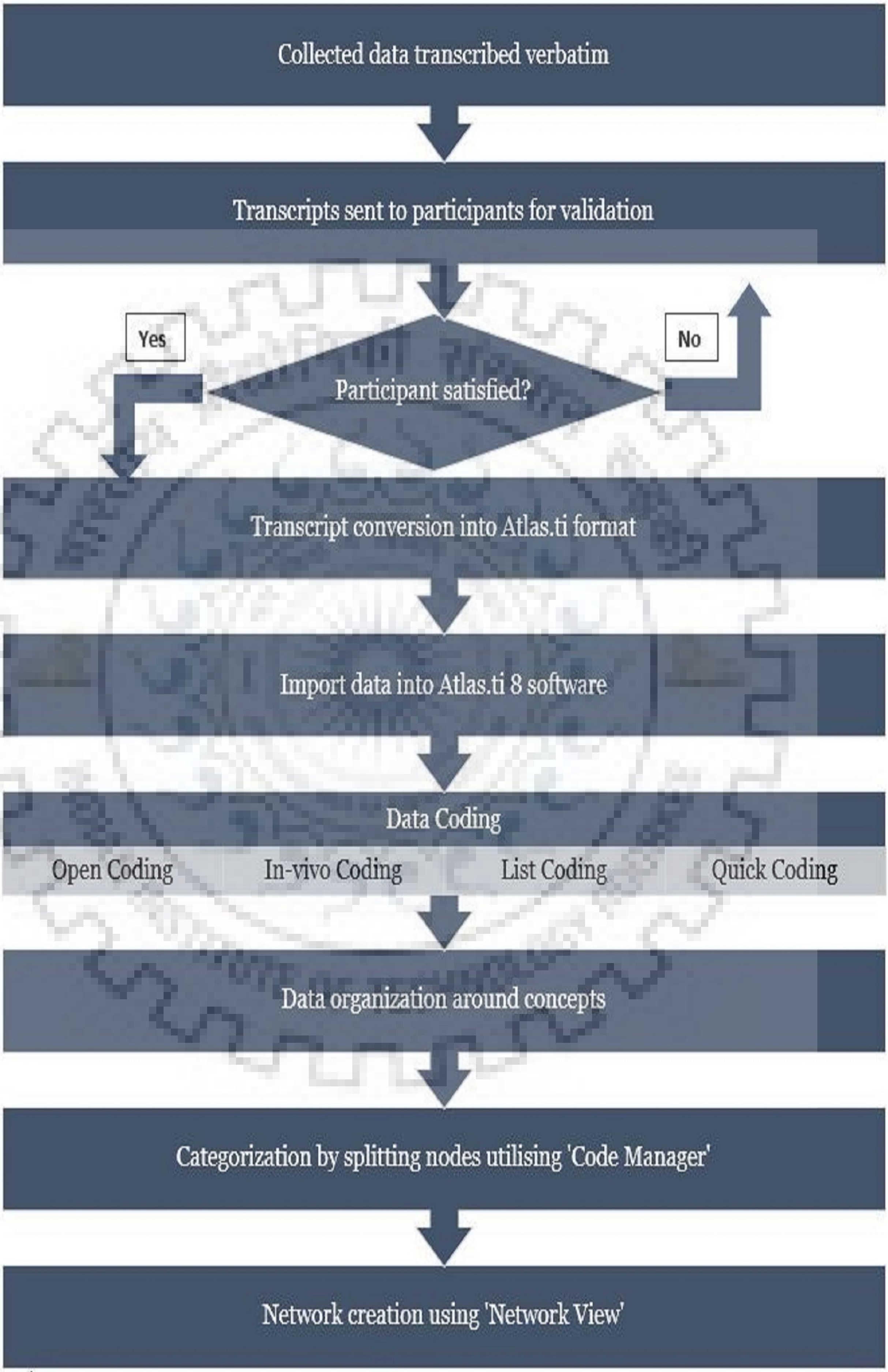


Fig. 4.3: Data Analysis using Grounded Theory

in the gathered data, by creating linkages (Rambaree, 2012). The results are displayed without any identifiable information of the participant.

4.3 Discussion

4.3.1 SM Platforms Usage

SM, being Internet-based technology, features user-generated content and personalized profiles. In the era of nomadic computing, where services come to the users whenever and wherever it is needed, all the participants are aware of SM platforms. They use [Figure: 4.4] Instagram, WhatsApp, Facebook, 9GAG, LinkedIn, Twitter, YouTube, Hangouts, Pinterest, Google Blog, Vimeo. Participant #3 mentions former SM network Orkut. There were similarities in the responses; 10 out of 10 participants used WhatsApp, Facebook, and YouTube. Another participant, #6, states: "*I access SM primarily on mobile phones and laptop as being a hostler, I hardly get a chance to watch through radio and television*." Figure: 4.4 illustrates the network results, and the numbers indicate the respective codes of the transcripts.

4.3.2 Motivation

Participants use different SM platforms for sundry reasons [Figure: 4.5]. SM is predominantly used for disseminating, listening, and gathering information. Most of them use for socializing, staying connected to friends, and being updated with the current activities in society and the world. For instance, floods hit Kerala, an Indian state in The Republic of India, in August 2018. Many participants stayed updated regarding flood information from the text, photos, and videos shared on SM.

Another instance is the #MeToo movement, a campaign for women empowerment by exposing sexual exploiters that have created awareness. Many of the participants use it as a leisure activity for entertainment. Some use it for seeking jobs and networking purposes on LinkedIn. Participant #7 states: " *I use SM to collect data from SM platforms for research purposes,*" while another participant, #2, mentions: " *I use it for educational purpose by listening to lecture videos.*" Participant #1 uses it to get different ideas for art and illustration. A participant, #8, used SM for planning her travel to Kerala as per the situational awareness using SM. Almost all the participants use SM for diverse reasons, but none use SM primarily for DM, signifying the lack of understanding to use the SM platform for DM.

4.3.3 Practical implementation of using SM in emergency

All the participants have not used SM in an emergency. A few participants have shared instances of using it in an emergency. sm in an emergency. One participant, #5, states: "*I have not used much of SM during a disaster, but I use it often for situational updates post-disaster.*" One participant shares an incident of arranging blood during a medical emergency using WhatsApp. One more instance of a social mishap and WhatsApp messages circulation helped get through the situation. WhatsApp message circulation of Cyber emergency of Facebook data leak helped

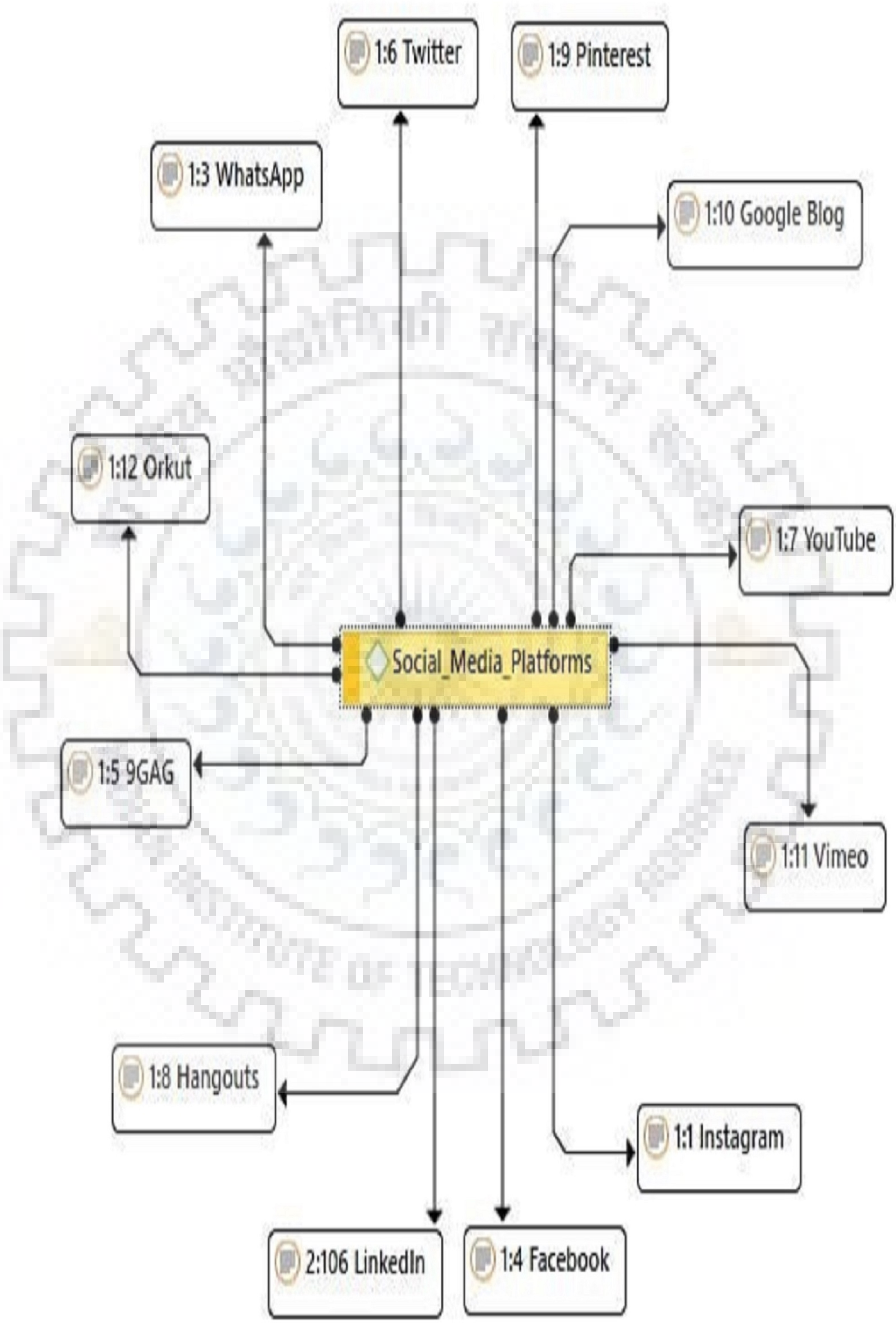


Fig. 4.4: SM Platforms used by participants

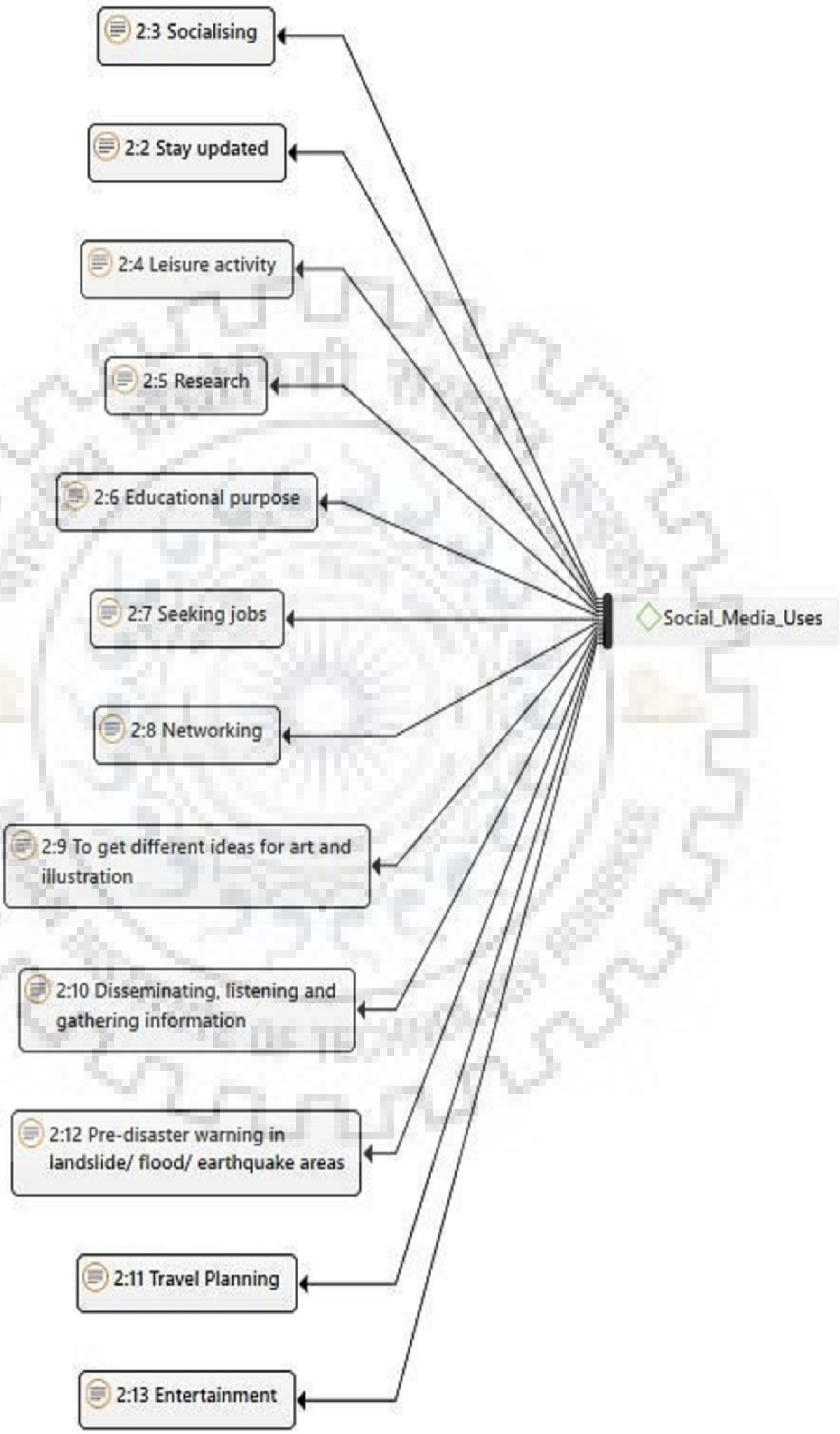


Fig. 4.5: Participants' purpose of using SM

many participants in protect their Facebook IDs and passwords. Former SM platform 'Orkut' is credited by a participant for generating Earthquake awareness and pre-disaster warning for Gaziabad Earthquake a few years back. Participants mention the use of Facebook and YouTube during the Kerala floods. They believe it helped them to stay updated with real-time situational awareness. A scholar participant #9 mentions that SM can be used at sites other than disaster sites because of alleged network issues during a disaster. It can also be used for mass communication during preparedness and mitigation.

4.3.4 Challenges

The research aims to identify the challenges of using SM in DM. Key findings on the challenges [Figure: 4.6] have been categorized into six themes: *Regulatory, Software, Authenticity, Physical, Cultural, and Demographic*. Each challenge is discussed in turn.

Regulatory

Regulatory barriers suggest that there is no *regulatory* authority of SM for DM, and there is a considerable gap between governmental use and citizens' use of SM for DM. A participant, #7 states: "*SM is not considered to be the first responder as it is not an official platform for DM, prime purpose being socialization and not DM.*" There is a lack of rules/ regulations/ laws on the use of SM for DM, due to which there is a misuse of freedom of speech/ freedom of expression. Another participant, #6, states: " *SM is required to be used as a tool for DM, and for that, it needs to be regulated.*" A lack of skilled/ trained professionals has been found for handling SM in DM. The use of SM is dependent on the type of disaster. In the case of an earthquake, there are significantly lesser chances that people can use it during disaster.

Software

Numerous *software* limitations have been identified. There is a lack of clarity on which SM platform to be used for DM. One of the participants, #1, states: "*I feel data redundancy leads to more loss in information.*" In addition to this, there is a lack of clarity on the use of appropriate hashtags and keywords for the message to reach the respective authority. A standard framework for DM using SM is missing. There is an issue with the restriction on the word limit of the message. There is no standard WhatsApp number for DM. The absence of a statistical database and paucity of cross SM platforms have been identified. Transparency of data poses a security threat.

Authenticity

There is a question mark on the message's *authenticity* on SM as there is no credibility check on the message. A participant, #2 mentions: "*Being a free source, it may lead to spreading of the fake news leading to misinformation and mass-hysteria. For instance, the WhatsApp forwarding*

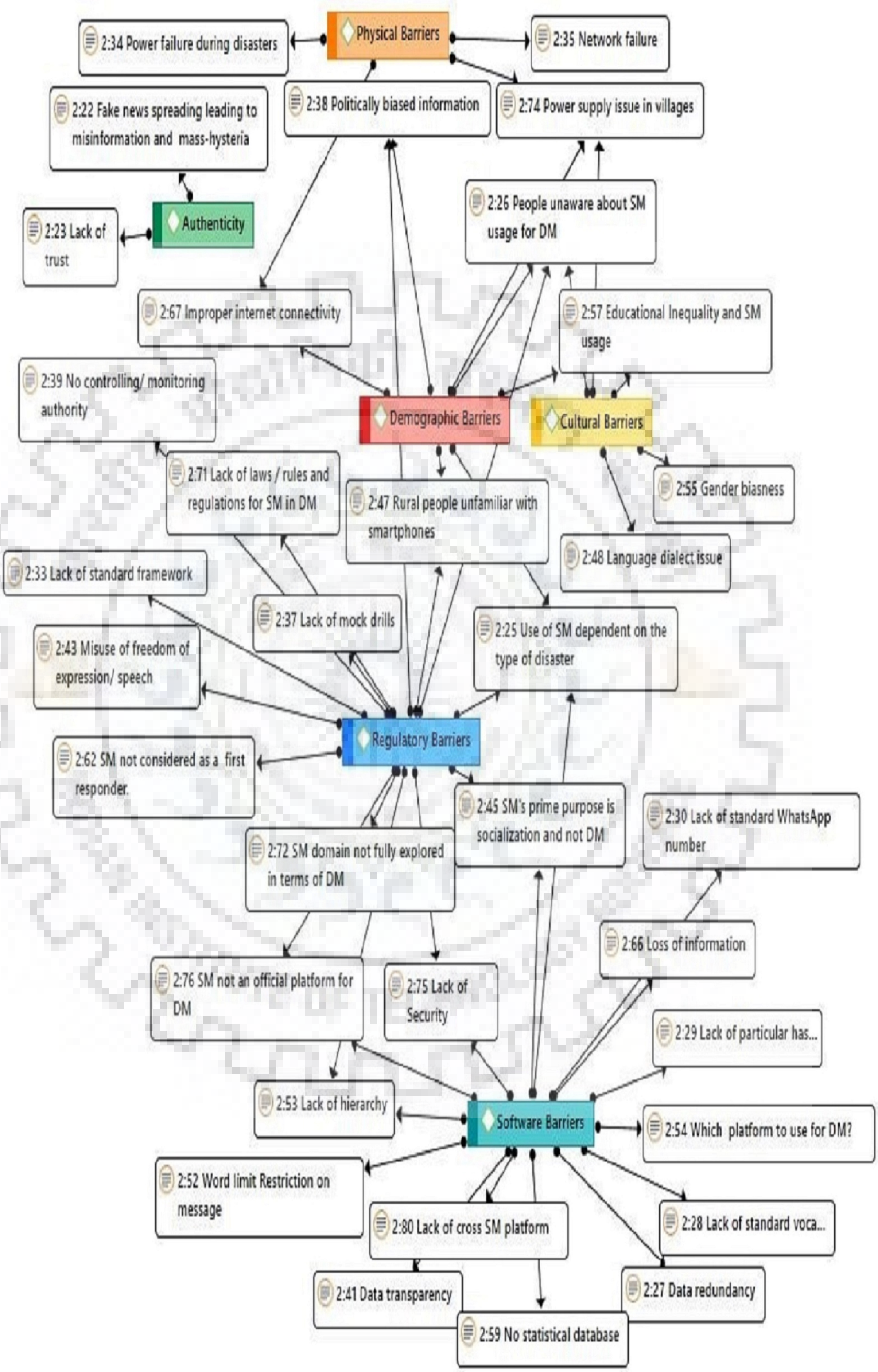


Fig. 4.6: Identified Challenges on the use of SM for DM.

feature led to mass lynching [1], *communal messages get viral on SM".*

Physical

Physical barriers constitute the power supply/ failure issue. Improper internet connectivity in rural areas is another vital issue. Network coverage and network failure are major drawbacks during disasters in DM. One participant states: "*During the time of any disaster, the first thing we don't have is power supply, even we have network problem because of the damage to the towers. So if we do not have any platform, how can we use SM during a disaster*". Another participant states: "*There was complete network failure during Kedarnath tragedy.*"

Cultural

Cultural barriers comprise gender biases in Indian society with inequality in the level of education in the juxtaposition between males and females along with rural and urban exposure to education. There are cultural reasons for not sharing information on SM. One of the participants, #5, succinctly explains: " *Poor communities have limited access to basic phones, so they hardly use smartphones and SM.*" Rural people are ignorant of smartphones which implicitly makes them unaware about SM or its usage for DM.

Demographic

Language is a *demographic* barrier as most SM is in the English language. According to a participant, #7," *because of lesser-educated society, people are a pro at regional languages but English.*" Moreover, a difference in the regional dialect of the language has also been identified as a barrier as dialect change makes it difficult to understand the context of the message.

4.3.5 Enablers

Logico-empiricial findings on enablers of SM for DM are exhibited in Figure: 4.7. They are elaborated as follows:

Rise in mobile penetration

The exponential rise in mobile penetration has enabled the use of SM via mobile internet, which implicitly augments the use for DM. An instance by a participant, #9 explains *"the case of floods where heavy rainfall news can be shared on SM for awareness.*"

Democratic participation

Democratic participation and engagement of Indian citizens enable SM's usage for DM as the Constitution of The Republic of India supports the freedom of speech. According to partici-

[1] https://www.livemint.com/Technology/XQonkoebctmCSyqzbJY8BN/WhatsApp-to-limit-message-forwarding-after-India-mob-lynchin.html

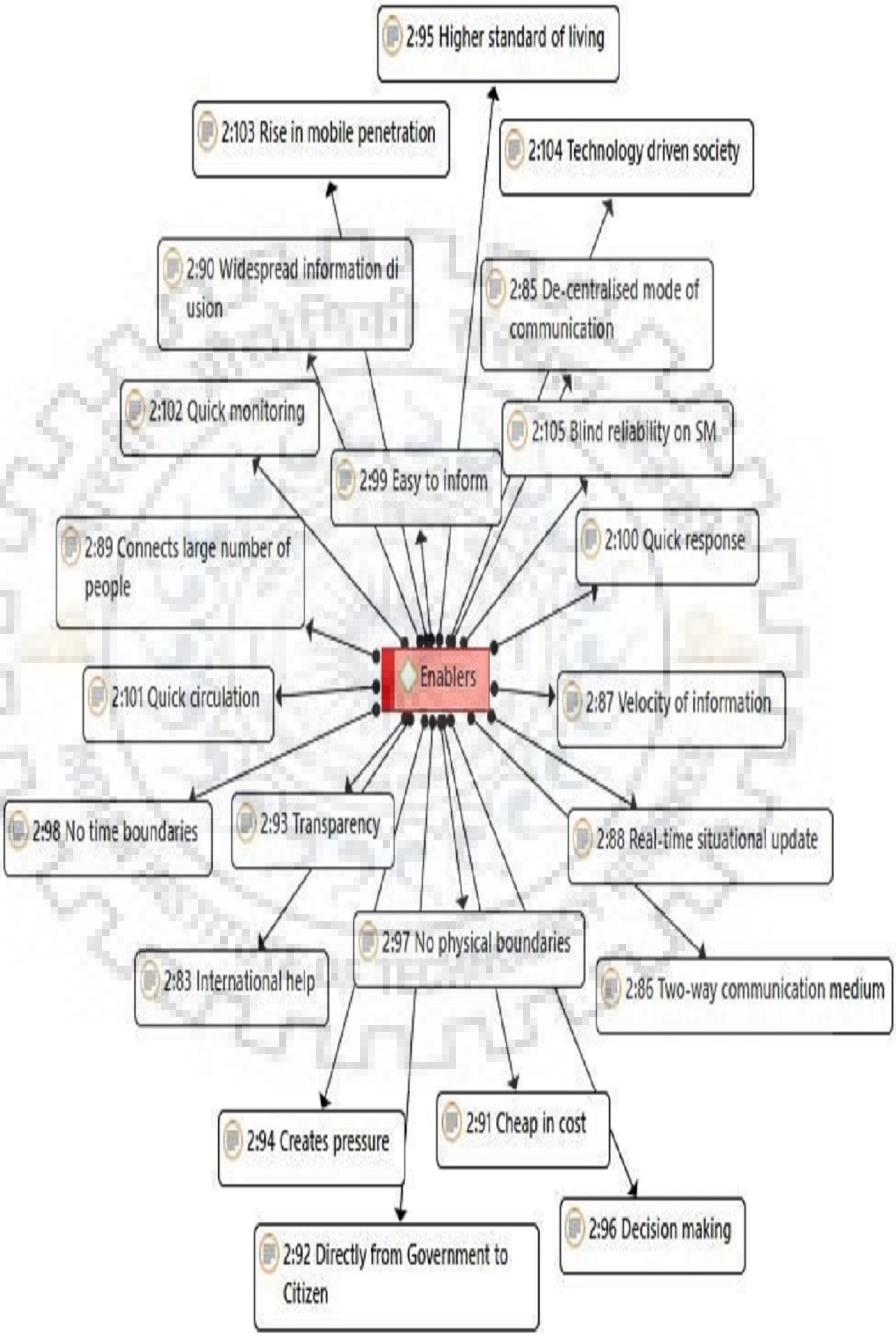


Fig. 4.7: Enablers of SM for DM

pant #1 ", the freedom of speech/ expressions enhances SM usage for DM". Moreover, there is richness in diversity helping people, evolve with the technology.

Increase in living standards

Over the last few years, the living standards of citizens of The Republic of India have been significantly ameliorated, leading to the augmentation of SM users. Participant #5 succinctly explains: "*Blind reliability on SM and technology-driven society enables the use of SM for DM.*"

Two-way real-time communication

According to participant #6, "*SM is a two-way communication medium, provides real-time information diffusion,*" enabling the use of the platform in DM.

Global reach

Over the recent years, SM garners global attention, and broader audience reach creates pressure on the government, triggering the government to act during disasters, enabling the use of SM for DM. Participant #7 states: " *SM connects to a more significant number of people on a single platform, so any news related to disaster can be spread to a larger audience quickly."*

Expeditious decision making

According to participant #4, "*Velocity of information being too high on SM enhances the prompt decision making and real-time situational update in DM.*" Expeditious decision-making augments the usage of SM for DM.

No space-time constraint

Communication regardless of space-time constraint fosters the usage of SM for DM. In addition to it, participant #7 states: "*The ease of using SM on fingertips plus the user-friendly features saves time and commuting distance.*"

Cheaper source of information

Participant #9 states: "*In this day and age, Internet being a cheaper source enhances the use of SM for DM."* More affordable internet opens doors to a wealth of information and knowledge regarding disasters.

4.3.6 Use of SM in juxtaposition with traditional media

Figure: 4.8 elucidates that the SM is more influential, being a two-way communication medium, connecting a more significant number of people simultaneously, generating a quicker response with no physical and time boundaries. There is blind reliability on SM. SM having a wider audience gives voice to every citizen to raise their point. SM is way more manipulative in contrast to traditional media. SM generates voluminous data with higher velocity as compared to traditional media. Participant #3 states: "*WhatsApp has been used in the time of disaster. So implementation has already taken place as earlier methods were confined to specific areas ,e.g., loudspeaker warning*". Another participant, #6, mentions: " *SM has been an emerging domain in DM. Its usage at Indian governmental level is still in its infancy in terms of DM compared to traditional media.*"

Building on the existing body of knowledge related to the challenges and enablers of SM usage for DM (Gunessee et al., 2017; Hiltz and Kushma, 2014; Kavanaugh et al., 2012; Lambert, 2020; Lieneck et al., 2022; Park and Johnston, 2017), the chapter showcases the diverse SM platforms in use with the participants' experiences of using it in an emergency. The findings of the FGD depicts the *Regulatory, Software, Authenticity, Physical, Culturalm and Demographic* challenges. In addition, different enablers are explored.

4.4 Conclusion

SM plays an important role in DM. The research study answers the call of RQs and understands the existing challenges and enablers, with the research tools and methods for QDA. The literature review highlights the usage of software for QDA purpose. The research study incorporates the Atlas.ti software for data analysis. The literature reviews about the research methods, i.e., in-depth interviews, semi-structured interviews and FGD. The research study utilizes FGD to accentuate the interpretability of findings. The literature review explores inductive and deductive theory for QDA. The research article focuses on inductive thematic theory for the findings.

The literature review explicates the social, technological, information, organization, technical barriers. The research findings classify the challenges into six themes. The literature review elucidates the enablers as public participation, networking, publicized meetings, information dissemination. The research study finds numerous enablers from the FGD data.

The research study understands the usage and role of SM in different catastrophic times through participants' experiences. In this chapter, we provide several implications by presenting findings from the multi-round analysis of FGD with Indian participants in reference to the usage of SM for DM. Informative data has been gathered from FGD within a homogeneous group. FGD generated data is filled with participant's rich working experience in DM domain. The findings of this study build our understanding of ways that SM is used by citizens and its range of practical purposes. The study contributes by explaining "what" challenges and enables the usage of SM for DM. In addition to it, SM usage with respect to the traditional media has been explored.

Atlas.ti software has been employed to categorize the extracted six barriers, namely *Regulatory, Software, Authenticity, Physical, Cultural, and Demographic*, from the FGD data in the Indian context. In addition to challenges, enablers have been drawn out from the FGD data. Identified enablers are the rise in mobile penetration, democratic participation, increase in living standards, two-way real-time communication, global reach, expeditious decision-making, no

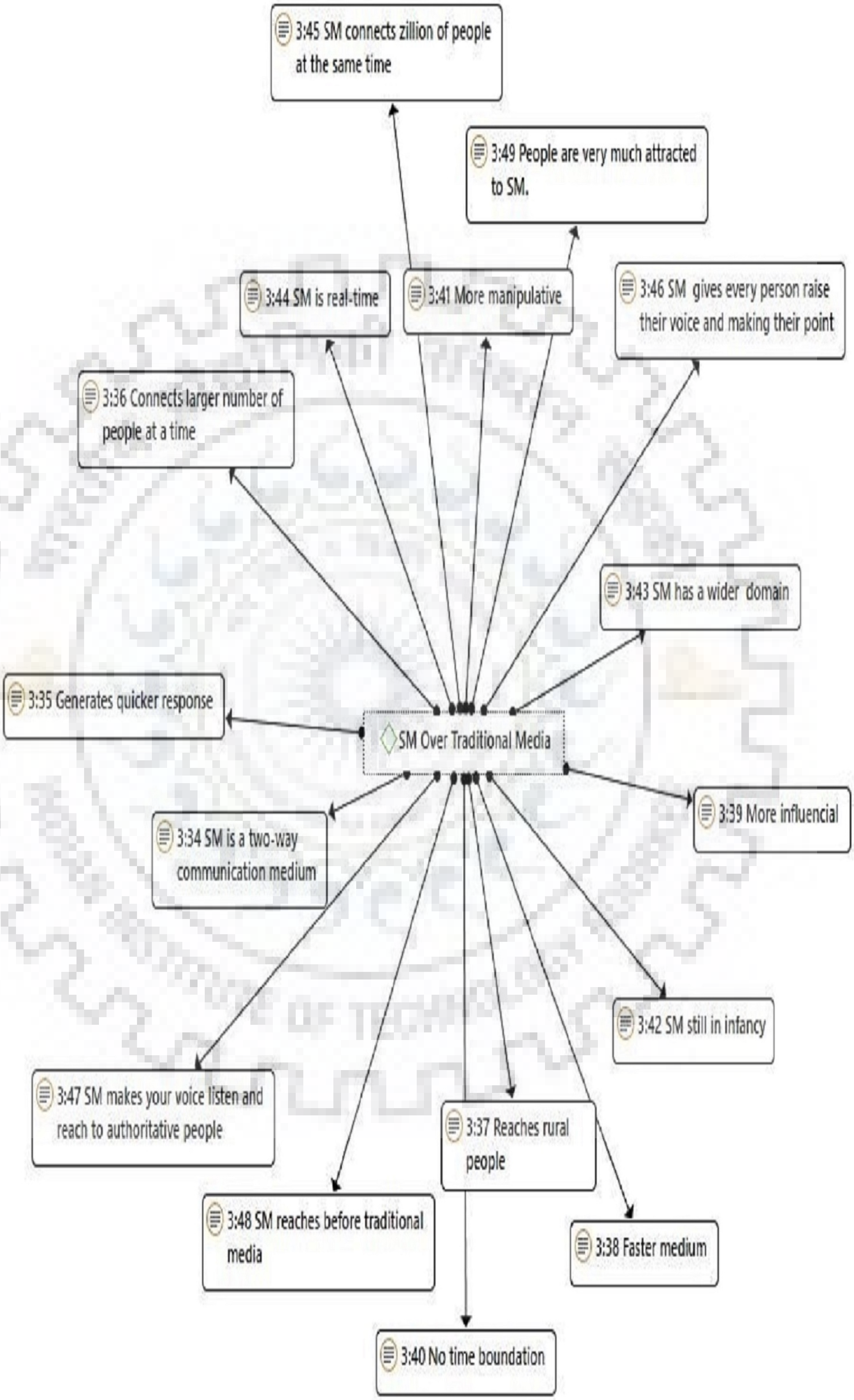


Fig. 4.8: SM Usage benefits over Traditional Media for DM

time-space constraint, cheaper source of information. In addition to it, we explore SM usage with respect to traditional media, understanding the positive edge of SM over traditional media. The research shed new light on the understanding of SM as a vital player in DM and contributed to enlarging the scope of advanced research on the relationship between SM and DM.

5. A FRAMEWORK TO IDENTIFY THE RELEVANCY OF A SOCIAL MEDIA MESSAGE TO A DISASTER.

5.1 Introduction

The automatic identification of information related to disasters is the most compelling and necessitous problem globally. Disasters negatively affect the lives of people. The occurrence of a disaster does not affect physically by ruining infrastructure, landscape, residences, and business properties; it also affects the emotional well-being of an individual. Apart from short-term losses, disaster has long-term repercussions in terms of job losses, financially insuperable infrastructure damage, and post-disaster trauma (Jamali et al., 2019).

Modern SM applications have considerably penetrated the lives of people, providing an insight into people's opinions, beliefs, and sentiments. SM connects people of different genders, races, ethnicity, and origin. According to Global Digital Report 2021 (Simon Kemp, 2021), the global population was 7.83 billion at the start of 2021, and 5.22 billion people are using mobile phones, equating to 66.6% of the total population. According to Global Digital Report 2021 the global population being 7.83 billion at the start of 2021, and 5.22 billion people are using mobile phones, equating the number to 66.6% of the total population. At this moment, 4.66 billion people use the internet, i.e., 59.5%. The same report pinpoints the number of SM users, mentioning that 4.20 billion active SM users are there, i.e., 53.6% of the population.

Currently, Twitter is a significant example of SM usage, with millions of users scattered around the globe. As of 2020, Twitter had more than 450 million active users. According to Internet Live Statistics 2021, Twitter disseminates 500 million tweets each day, approximating it to 200 billion tweets in year [1]. Luckily, the data is relatively available in the public domain for research purposes. However, extracting relevant information immersed in noisy, irrelevant, and unstructured SM data during disasters is a challenge in SM studies. Haplessly, apart from the volume, velocity, and variety of the SM data, the complex nature of the catastrophe, makes the aforementioned data less applicable on disaster-based studies, leading to decreased system performance and inaccurate results. People switch to SM to seek and offer help related to food, water, electricity, shelter, donation, etc., at the time of disaster. Therefore, there is a need for an intelligent framework that can automatically identify the relevancy of the SM message to a disaster so that timely decisions can be taken.

[1] https://www.internetlivestats.com/twitter-statistics/

Although DL techniques have received considerable attention during the last decades, most studies focus on a single DL technique to identify the disaster-related SM message. With the growth in data and the improvement of demand, the DL model built for data identification is increasingly complex. It is no longer a single model but a more complex hybrid model. The lower efficiency of existing methods could not meet the demand in DM yet. Meanwhile, due to the difficulty of the DL model in stability, it is essential to develop a more effective model to identify disaster-related SM message. Therefore, a hybrid LSTM-CNN model is proposed for determining the disaster-relevant SM messages. By fusion of LSTM and CNN models, the structure and classification abilities are improved, creating greater synergy than the individuals of their own. Therefore, the main goal of this work is to develop iRelevancy, i.e., an intelligent framework for disaster relevancy identification of SM-based messages using DL techniques so that accurate decisions can be taken at the right time. The objective is important for dissertation as it transforms the raw data into a knowledgable format. Hence, the objective of this research paper is to propose a framework to answer the following RQs:

1. How can DL be used to determine the relevance of an SM message to a disaster?
2. Which DL model is effective in classifying disaster-related data?
3. How effective is the proposed model in terms of performance?
4. Can the same DL model be used for different disasters?
5. How a DL-based model performs on different disasters?

The main contributions of this research paper are as follows:

1. We propose a cyclone Fani disaster dataset comprising SM message and the label of the relevance, i.e., *relevant (0)* or *irrelevant (1)* to disaster.
2. We present a hybrid DL-based framework to identify the disaster relevancy of an SM message.
3. We conduct extensive simulations against other neural network models and existing research studies to confirm the superiority of the proposed model.
4. We demonstrate the technical efficacy by determining the disaster-relevance of SM message of cyclone Fani.
5. We exhibit the identification of other cyclonic disasters' SM message disaster-relevance.
6. We illustrate the restating of the proposed model to other disaster events. We apply our proposed model technique to SM messages posted during the Covid19 event in 2020. We see that our method is effective for different disaster types.

Although the approach explored in this research can be applied to any disaster, cyclone Fani was selected for this case study due to the abundance of data available and the magnitude of the damage it caused. Using cyclone Fani as a case study, this research inspects the solutions to the RQs. The rest of the paper is as follows. Section: 5.2 expounds on the proposed framework. Section: 5.3 offers the conclusions about the proposed work.

5.2 Methodology

For the proposed framework on relevancy identification, we go through the following phases: (1.) collecting the cyclonic and pandemic data from the Twitter platform; (2.) manual labeling of the cyclone Fani dataset for the framework; (3.) pre-processing of the data; (4.) designing of the hybrid model; (5.) evaluation of the proposed model; (6.) comparison of the proposed framework with state-of-the-art methods and existing research studies; (7.) predictions on cyclonic and pandemic dataset. Figure: 5.1 illustrates the iRelevancy framework of the research study.

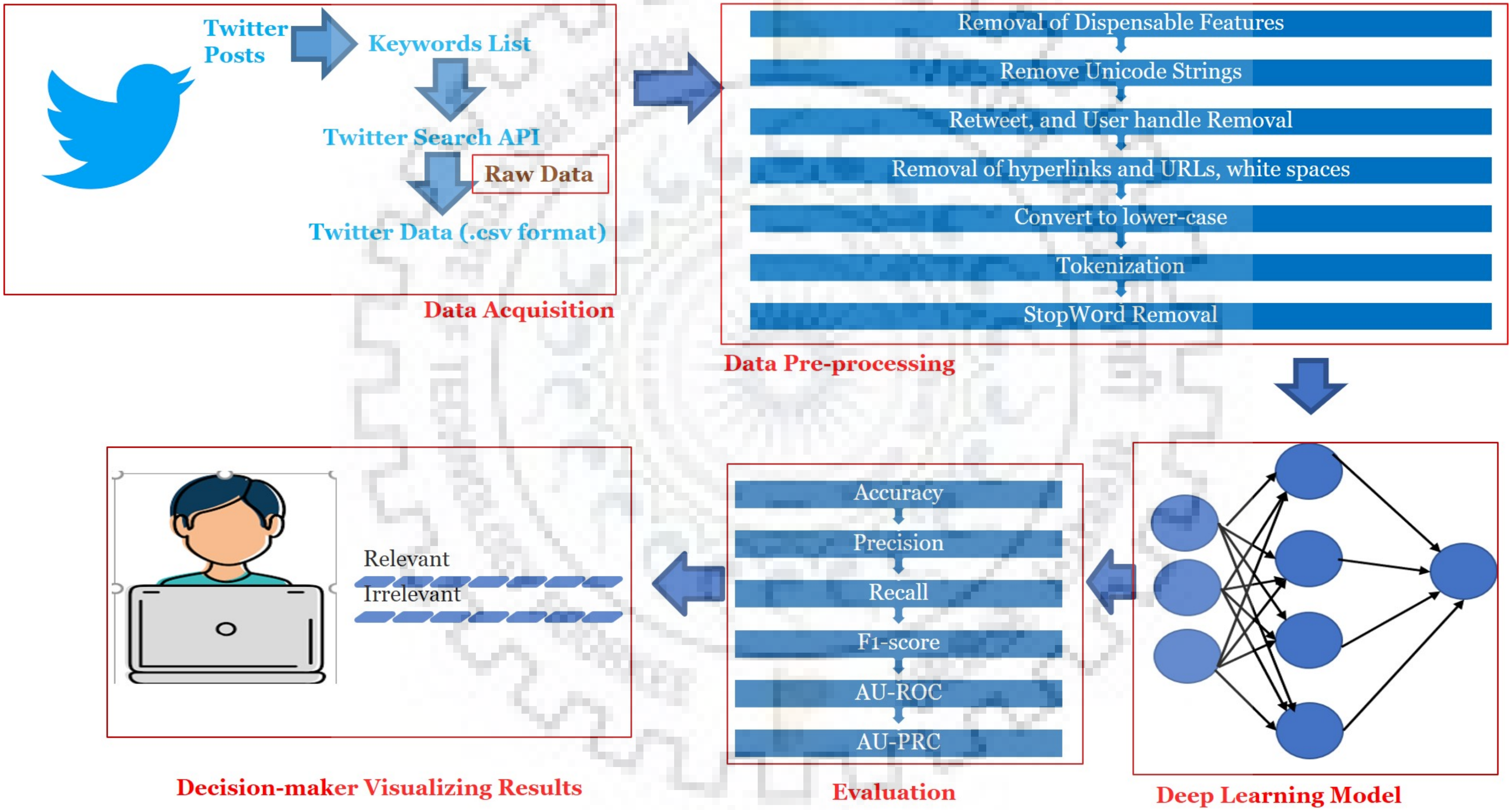


Fig. 5.1: iRelevancy Framework to identify disaster-relevancy of a social media message.

5.2.1 Data Acquisition and Dataset Description

The data for the aforementioned disasters are collected from the Twitter platform using a search application programming interface by our web crawler developed in Anaconda Integrated Development Environment using Python language. The data contains numerous features like created_at, text, lang, id, user object. Further, the user object collected features are screen_name, userid, id_str, name, location, url, description, verified, followers_count, friends_count, listed_count, favorites_count, statuses_count, created_at, and profile_image_url. The fetched data is saved in csv format. The pictorial representation is showcased in Figure: 5.2. Twenty features have been fetched in the form of raw and unstructured data. Table: 5.1 represents the extracted features of the tweet.

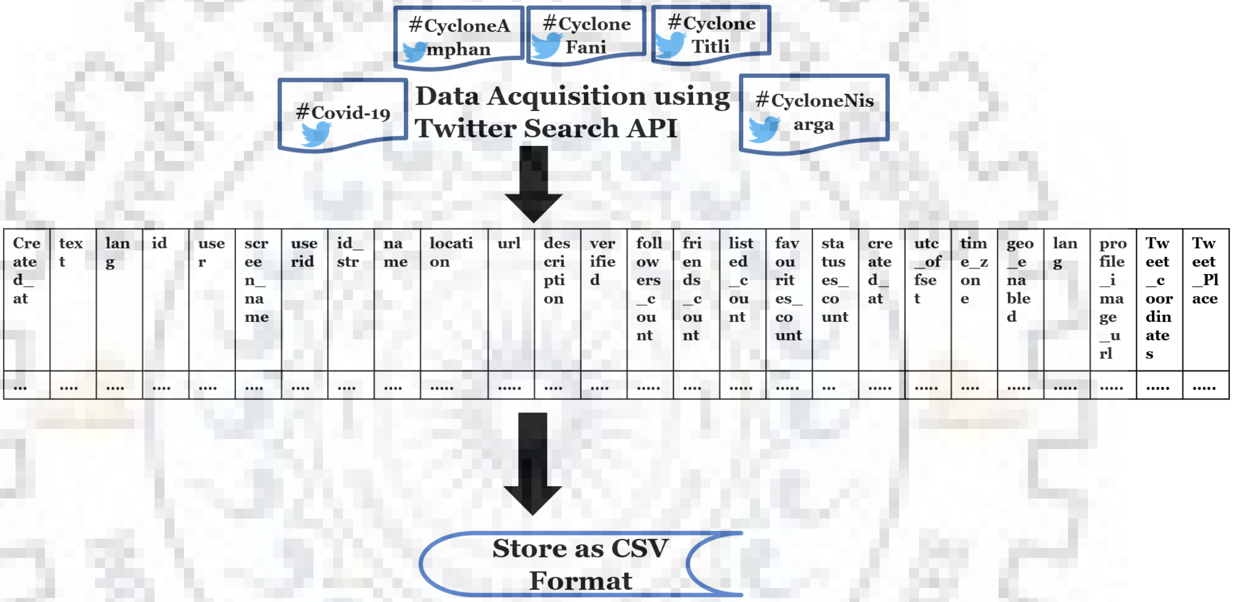


Fig. 5.2: Data acquired from the Twitter platform using Search Application Programming Interface and stored in csv format.

The disasters considered in this research study include cyclone Fani, cyclone Titli, cyclone Nisaraga, cyclone Amphan, and ongoing Covid-19. In May 2018, an extremely severe cyclone happened over the Bay of Bengal, from April 26 to May 04, 2019. 16.53 million got affected. In addition, 1053 primary schools and more than 5 lakh trees were severely damaged (Society, 2019). Another very severe cyclone Titli made landfall over the Srikakulam district of Andhra Pradesh in October 2018. The government of Andhra Pradesh reported a loss of Rs.3680 crore along with seven deaths. Cyclone Fani is reported as Category-4, and cyclone Titli as Category-2 cyclone (Mishra et al., 2021). The descriptive statistics of the gathered data are specified in Table: 5.2.

The third cyclonic event, being cyclone Nisarga, made landfall near Alibaug in June 2020 (Anjali Marar, 2020). Super cyclone Amphan struck the country devastatingly on May 16-20, 2020, causing mammoth loss to resources and properties (Goswami et al., 2021). Twenty million people have been affected in West Bengal alone, with about 100 deaths (Basheer Ahammed and Pandey,

Sr No.	Feature	Description	Type
1.	created_at	Universal Coordinated Time of creation of message	String
2.	text	The actual UTF-8 text of the message	String
3.	lang	Language of the message	String
4.	id	Unique identifier of the message	Int64
5.	user	The user who posted the message	User Object
6.	screen_name	User handle or alias of the user	String
7.	userid	Integer representation of the unique identifier of this user	Int64
8.	id_str	String representation of the unique identifier for this user	String
9.	name	User-defined name of the user	String
10.	location	User-defined location of the user	String
11.	url	URL provided by user in association with his profile	String
12.	description	User-defined string describing the account	String
13.	verified	Indicates whether the profile is verified or not	Boolean
14.	followers_count	The number of followers this account has	Int
15.	friends_count	The number of users, this account is following	Int
16.	listed_count	The number of public lists this account is member of	Int
17.	favorites_count	The number of SM messages this account has liked	Integer
18.	statuses_count	The number of messages this account has posted	Int
19.	created_at	The Universal Coordinated Time of creation of user account	String
20.	profile_image_url	A https-based url pointing to user profile image	String

Tab. 5.1: Feature Information of Social Media Data from Fetched Disaster Datasets

Disaster	Hashtag Used	Dates	# of Tweets
Cyclone Fani	#CycloneFani	May 03-10, 2019	1,72,907
Cyclone Amphan	#CycloneAmphan	May 16-21, 2020	8214
Cyclone Titli	#CycloneTitli	October 08-12, 2018	96,286
Cyclone Nisarga	#CycloneNisarga	June 01-04, 2020	81,475
Covid-19	#Covid-19	August 08-15, 2020	76,953

Tab. 5.2: Descriptive Twitter Statistics of Disaster Events

2021). The last incident is Covid-19. World Health Organization (WHO) declared Covid-19 as a pandemic on March 11, 2020. It has been reported 39.6 million cases as of October 2020, with more than 1.1 million deaths (Wang et al., 2021).

5.2.2 Manual Annotation

The annotation task aims to label the tweets as *Relevant* or *Irrelevant*. We mark the tweets as *Relevant* or *Irrelevant*, after removing duplicates, according to the following criteria:

- *Relevant Tweets*: The tweets contain information about the need, situation, availability of vital resources like food, water, electricity, etc.
- *Irrelevant Tweets*: The tweets that do not contain any information regarding the disaster.

The manually labeled tweets of the cyclone Fani dataset are shown in Table: 5.3, and tweet distribution of annotated SM messages is illustrated in Figure: 5.3. Out of 1,72,907 tweets in total, 65,140 tweets are remained after removing duplicates. Table 5.4 depicts the tweet distribution statistics. The human-annotated data is used to train the classifier. This contributes in proposing a cyclone Fani disaster dataset, comprising of the SM message and the label of relevancy.

Tweet	**Manual Label**
b" #CycloneFani An appeal to all to #HelpRebuildOdisha, please contribute to Chief Minister's Relief Fund as much as you can \xf0\x9f\x99\x8f\xe2\x80\xa6 https://t.co/BSHLSvEceU"	Relevant (0)
b'Rahul Gandhi addresses election rally in #Bhiwani, criticises PM Modi for implementing #GST, #demonetization; PM ad\xe2\x80\xa6 https://t.co/rRk1QWjHA6'	Irrelevant (1)
b'#CycloneFani #PowerSupplyUpdate : 33KV Jagatsinghpur feeder restored. Power supply to district headquarter hospital'	Relevant (0)

Tab. 5.3: Manual Segregation of Relevant and Irrelevant Tweets.

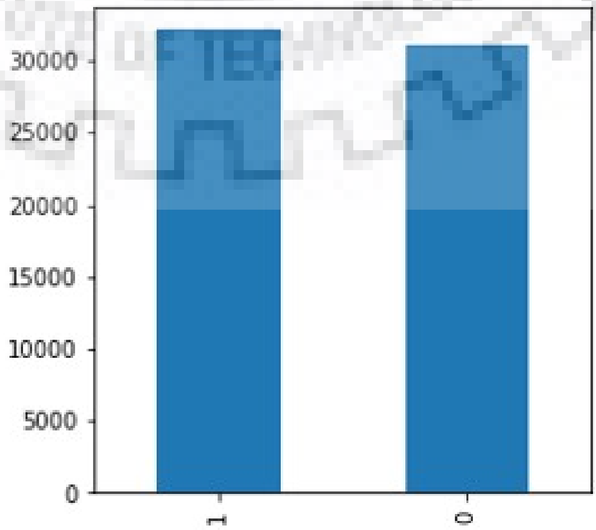


Fig. 5.3: Tweet distribution of cyclone Fani dataset. *0* represents *relevant* messages, and *1* illustrates *irrelevant* messages.

Label	# of Tweets
Relevant (*0*)	33,265
Irrelevant (*1*)	31,875

Tab. 5.4: Tweet Distribution Statistics. *0* represents *relevant* messages, and *1* illustrates *irrelevant* messages.

5.2.3 Pre-processing of Tweets

The collected real-world dataset is raw, unstructured, and meaningless. Therefore, data needs to be pre-processed in order to ensure the reliability of the knowledge discovery from the data. In this phase, we: (1). eliminate dispensable features, where only the tweet text is retained; (2). remove Unicode strings (b' and \x02) and non-English characters ('!', '#,' etc.), as they are remnants of the crawling procedure; (3). retweets repeat the old information as they have nothing new to offer. Therefore, the tweets starting with 'RT' are eliminated; (4). user handles are removed, considering the privacy of the user; (5). the special elements, i.e., URLs and hyperlinks, are removed; (6). additional white spaces are removed as they do not provide any semantic information; (7). capital letters are replaced by lower-case letters; (8). generate tokens by tokenizing tweet words; and finally (9) English language stopwords are removed as it does not add any meaning and reduce the classifier's performance. Figure: 5.4 illustrates the pictorial representation of pre-processing.

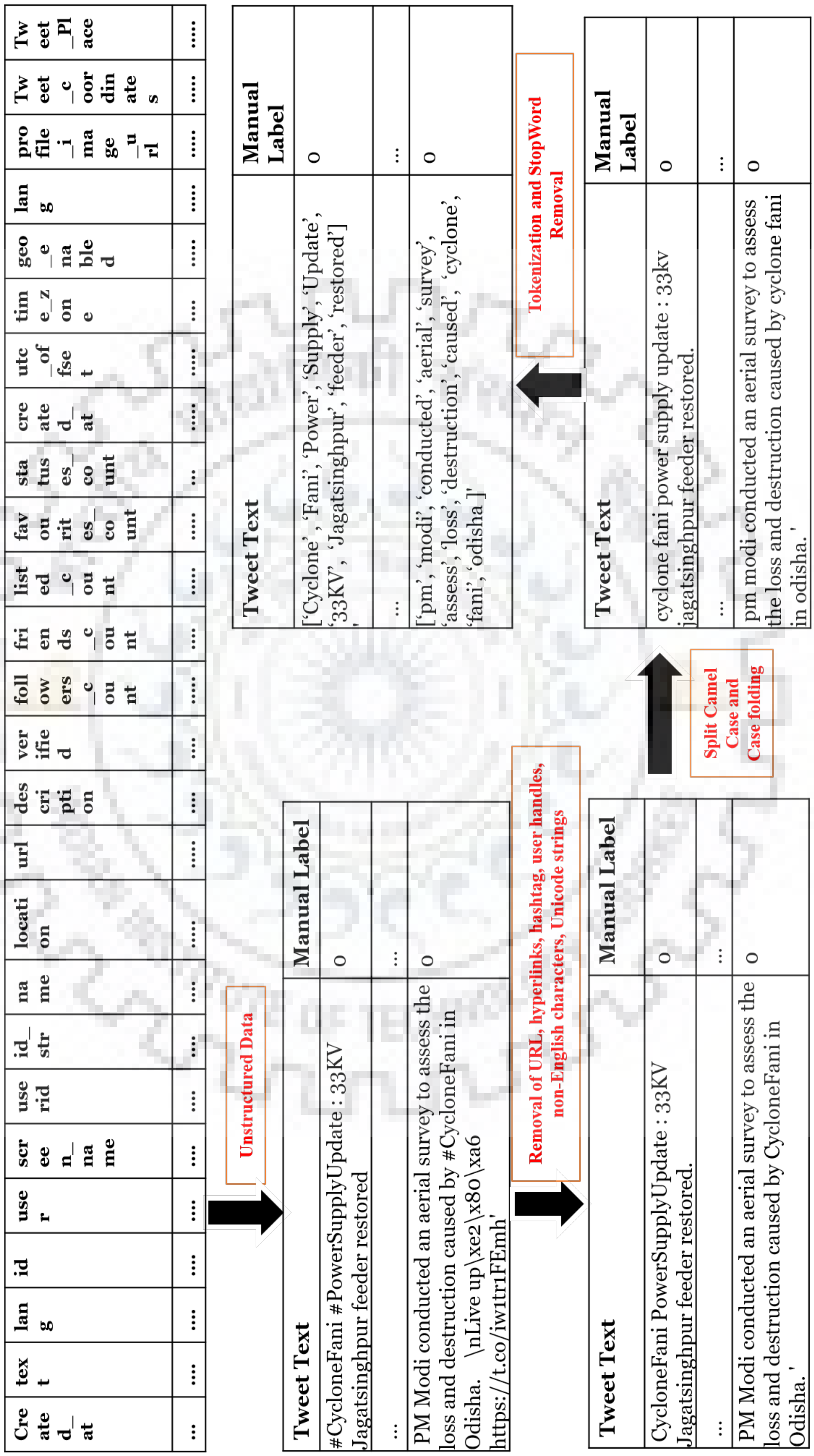


Fig. 5.4: Preprocessing of Tweets

5.2.4 Model Design

We randomly split the cyclone Fani dataset into three subsets: *training, validation, and testing* in the ratio of 60:20:20 Alam et al. (2018) Tam et al. (2021). The *training* data is used to train the model, and the *validation* dataset is used to tune the hyperparameters and provide an unbiased evaluation to select the best model. The *testing* dataset is used for testing and prediction purposes. In this chapter, we develop a DL-based hybrid model composed of multiple layers of the cascade. The hybrid model is proposed as it brings the best features of both the models. Section 3.6 describes the superior performance of the hybrid DL models in diverse applications of DM. The pseudocode of the proposed model is summarized in Algorithm 1. Figure: 5.5 provides a pictorial representation of the proposed model.

Algorithm 1: Pseudo-code for Proposed Model

Input: Training dataset, Validation dataset, Manual label of messages
Output: Label Prediction of SM message

```
Begin
For number of training iterations do
Construct sentence matrix using embedding layer
Employ LSTM to learn the contextual features of text sequences
Reduce the dimensionality by maxpooling layer
Apply ReLU activation function f(x)= max(0,x) to convolution layer
Add dropout to prevent overfitting
Feed the comprehensive representations into the softmax classifier to get the class labels
Update parameters of the model using the loss function with the Adam method
The performance metrics (Accuracy, P, R, F-score, AU-ROC, and AU-PR-Curve) are
  calculated using equations in Table: 7.2
End
```

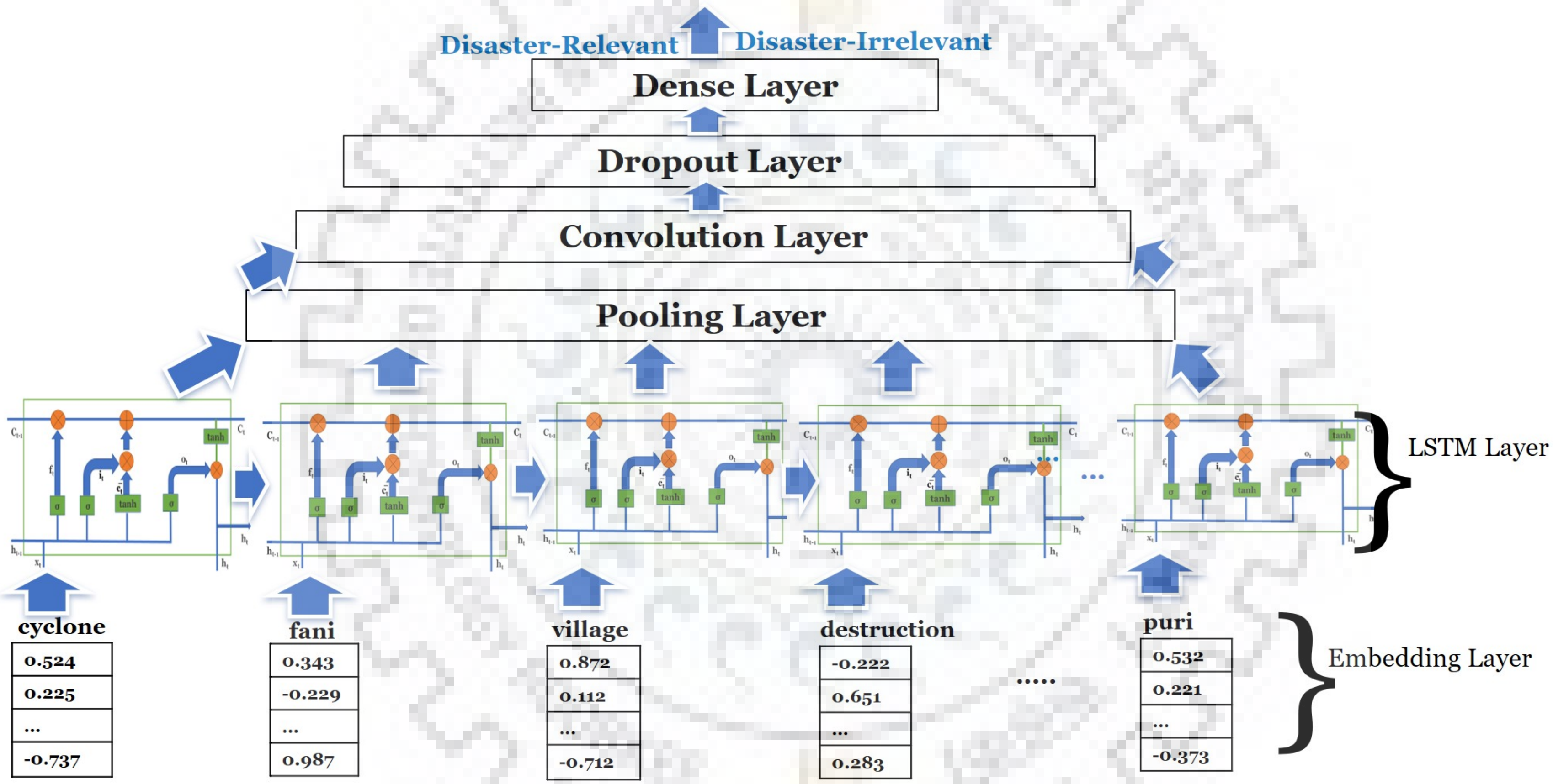


Fig. 5.5: Proposed architecture for SM message classification into *disaster-relevant* and *disaster-irrelevant* messages. The model consists of embedding, LSTM, pooling, CNN, dropout, and dense layers.

The model consists of an LSTM layer, which receives the word embeddings for each token in the tweet as inputs. The LSTM layer receives the input from the embedding layer and generates a new encoding for the input. Subsequently, the LSTM layer's output is pooled to a smaller dimension. The output of the pooling layer is fed into the convolution layer, where it extracts the local features and is activated by a non-linear function ReLU. Dropout layers are used to overcome the overfitting issues. The dense layer ultimately outputs the tweet as *relevant* or *irrelevant* to the disaster. Fig. 5.6 presents the architecture of the proposed model generated by the Anaconda Integrated Development Environment.

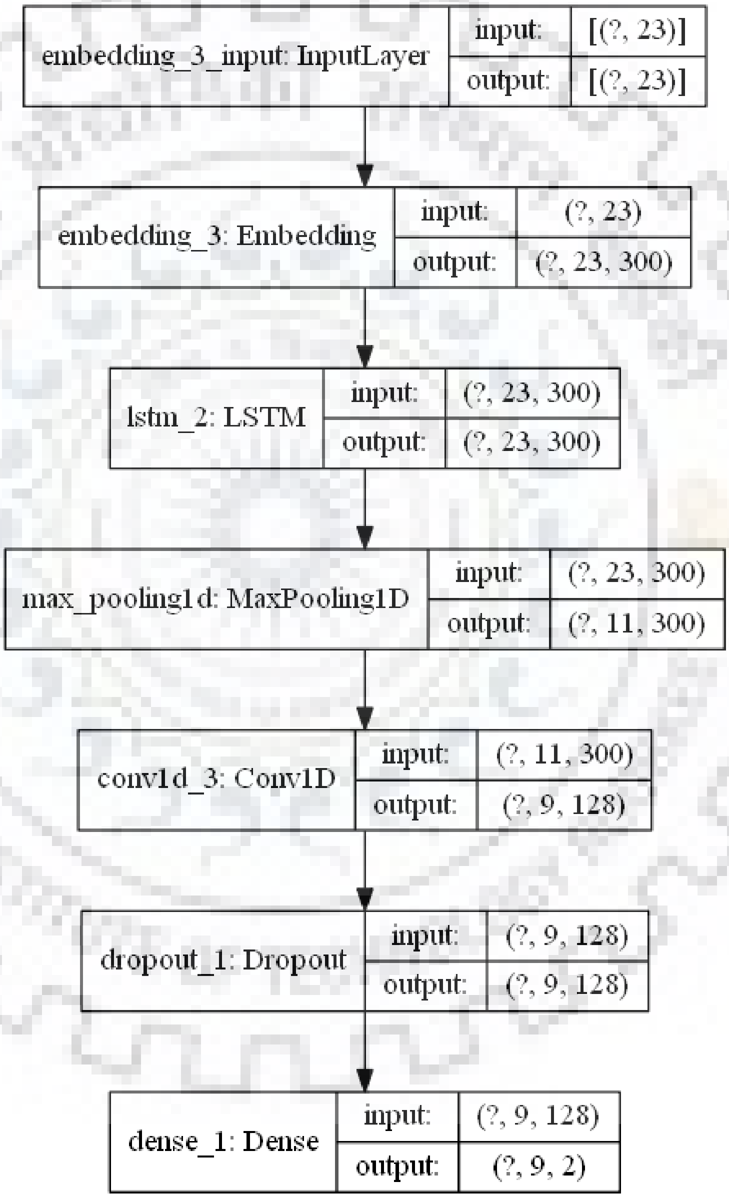


Fig. 5.6: Architecture of the proposed model generated by the Anaconda platform.

The model is compiled using a loss function and an optimizer. The optimum number of hyperparameters are selected using grid search. The selected hyperparameters with their definitions and values are specified in Table: 5.5. The working environment of the proposed approach is as follows:

- *Hardware Configuration*: Intel(R) Core(TM) i5-8250 8th Generation, 8GB RAM, 512 GB

SSD, and NVIDIA GeForce MX150.

- *Software Configuration*: Microsoft Windows 10, 64-Bit, Python 3.8.3.

Hyperparameter	**Definition**	**Value**
Epochs	Iteration count	25
Embedding Dimension	Size of vector used to represent each of the word embedding	300
Maximum Sequence length	Maximum tweet-length	23
Activation Function	Calculates the weighted sum of its input, adds a bias and then decides whether to activate the neuron or not	ReLU
Dropout Rate	Some hidden layer neurons are discarded with the 40% probability	0.4
Optimizer	The method used to update the weights in order to reduce the error	Adam(1e-3)
Batch Size	Data are grouped into batches prior to feeding it into DL model	1024
l2 regularizer	A regularizer to apply penalties on layer parameters during optimization to prevent overfitting	0.001
Loss Function	Function to assess model prediction	Binary CrossEntropy

Tab. 5.5: The key parameters (hyperparameters) used in study.

5.2.5 Performance Evaluation

Experiments are performed and in order to evaluate the performance of the proposed model trained for each of the configurations, the focus is not only on one performance metric; instead, multiple metrics are considered. Specifically, the ”confusion matrix” illustrated in Figure: 5.7 (also known as ”error matrix”) and its related metrics are used for assessing the metrics. The X-axis represents the human-annotated labels, and the Y-axis represents the prediction results. Accuracy, precision, recall, f1-score are the used performance metrics. Furthermore, the area under ROC and Area under PR-Curve are deployed to measure the completeness and robustness of the framework. The performance metrics are summarized in Table: 5.6.

	Actual Relevant	Actual Irrelevant
Predicted Relevant	True Positive (TP)	True Negative (TN)
Predicted Irrelevant	False Positive (FP)	False Negative (FN)

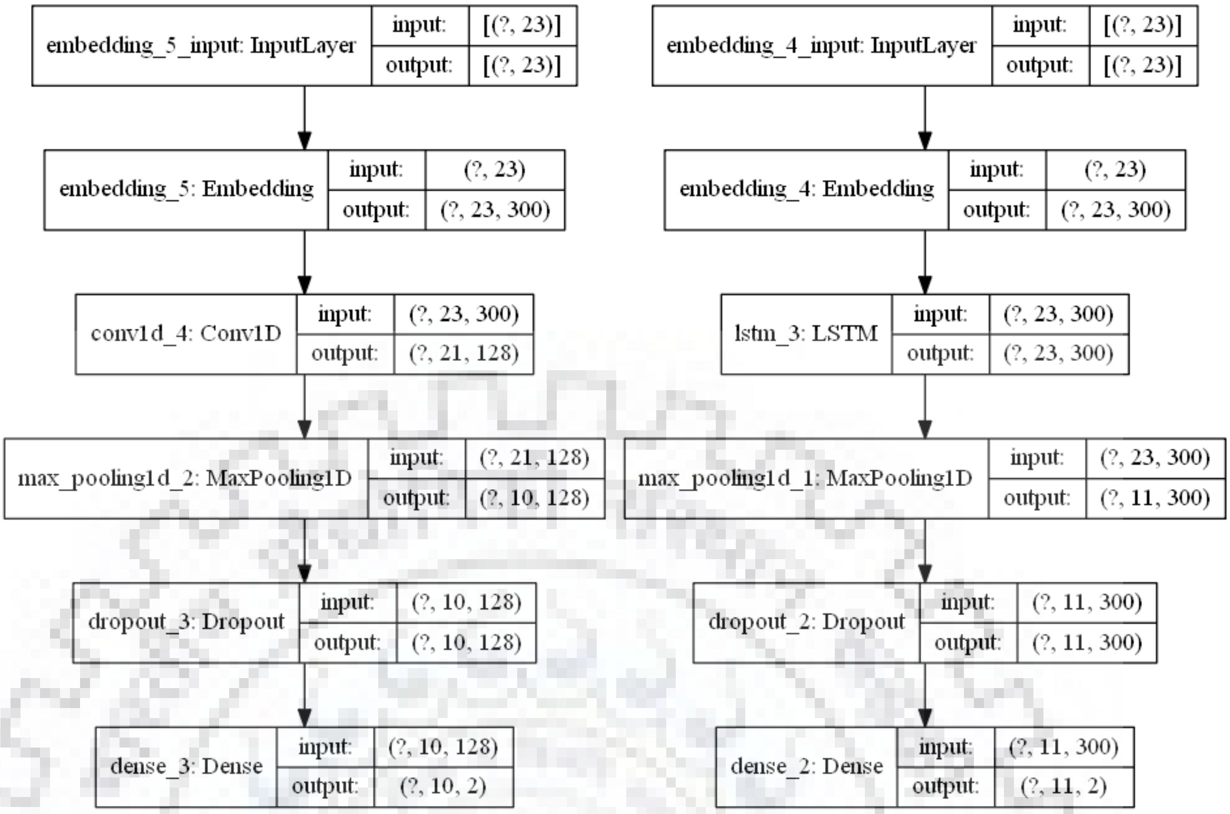
Fig. 5.7: Confusion Matrix (Han et al., 2011)

Name	**Description**
True Positives (TP)	Number of relevant tweets that are correctly classified to a specific relevant category.
False Positives (FP)	Number of relevant tweets that are incorrectly assigned to a specific irrelevant category.
True Negatives (TN)	Number of irrelevant tweets that are correctly assigned to the relevant category.
False Negatives (FN)	Number of irrelevant tweets that are inaccurately classified by the classifier.
True Positive Rate (TPR) (Sensitivity)=TP/(TP+FN)	The proportion of tweets that belongs to the specific relevant category and are really classified into that category.
False Positive Rate (FPR)=FP/(FP+TN)	The proportion of tweets that is incorrectly classified to a particular relevant category that it should not belong to.
Accuracy =(TP+TN)/(TP+TN+FP+FN)	%age of correct predictions by the classifier.
P=TP/(TP+FP)	A measure of exactness(%age of predicted relevant tweets that are actually relevant)
R=TP/(TP+FN)=TPR	A measure of completeness(%age of relevant tweets labeled as such)
F-measure =$(a^2+1)P*R/(a^2(P+R)$	Harmonic mean of Precision and Recall
AU-ROC	Area under ROC
AU-PRC	Area under PR-Curve

Tab. 5.6: Measures of Performance for Proposed Study.

5.2.6 Comparison with Other Models

In this section, we train DL models - CNN and LSTM. The comprehensive descriptions of CNN and LSTM are narrated in Kim (2014) and Hochreiter (1997), respectively. The architectures of models are displayed in Figures: 5.8a and 5.8b, generated by the Anaconda platform. The

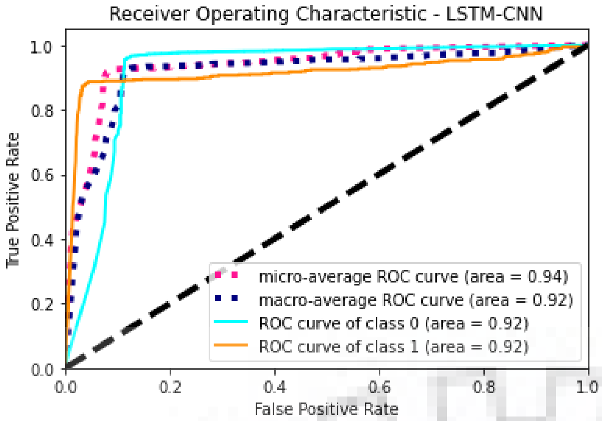


(a) CNN　　*(b)* LSTM

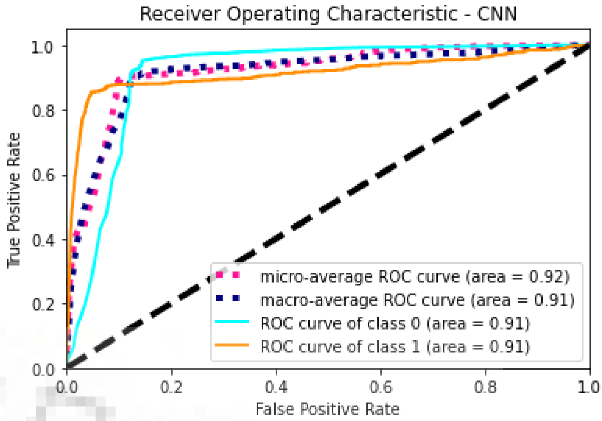
Fig. 5.8: Architectures of CNN and LSTM models for Disaster Relevancy

model performance results are displayed in Figures: 5.9a- 5.9c. As shown in Figure: 5.9, the proposed model [Figure: 5.9a] with an Area Under Curve (AUC) of 0.94 performs better than CNN(AUC=0.92) and LSTM(AUC=0.93) models, [Figure: 5.9b and Figure: 5.9c, respectively]. A higher level of AUC-PR-Curve(0.93) [Figure: 5.10a] than other models(0.92 for CNN [Figure: 5.10b], and 0.92 for LSTM [Figure: 5.10c]) is illustrated in Figure: 5.10.

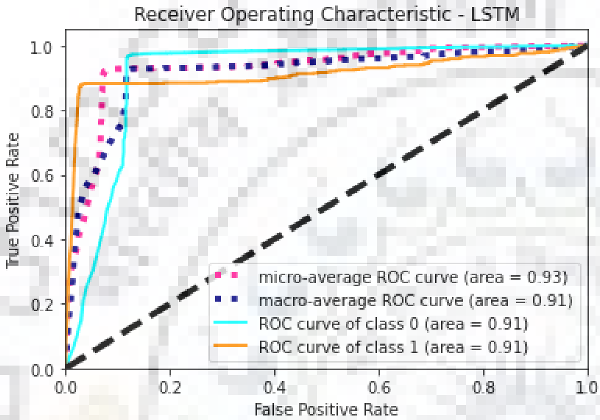
The performance of the proposed model against CNN and LSTM models is depicted in Table:5.7. The proposed model attains a higher value of accuracy, i.e., 0.92 than the other models (0.90 for CNN, and 0.91 for LSTM); along with values of precision(0.93 for the proposed model, 0.90 for CNN, and 0.92 for LSTM). The proposed model has a higher recall value(0.95); CNN and LSTM have 0.90 and 0.92, respectively. Computing the F1-score with the formula mentioned in Table: 5.6, the F1-score of the proposed model is 0.93, which is higher than CNN, and LSTM has 0.90 and 0.92, respectively.



(a) ROC curve of proposed model with 0.94 value.

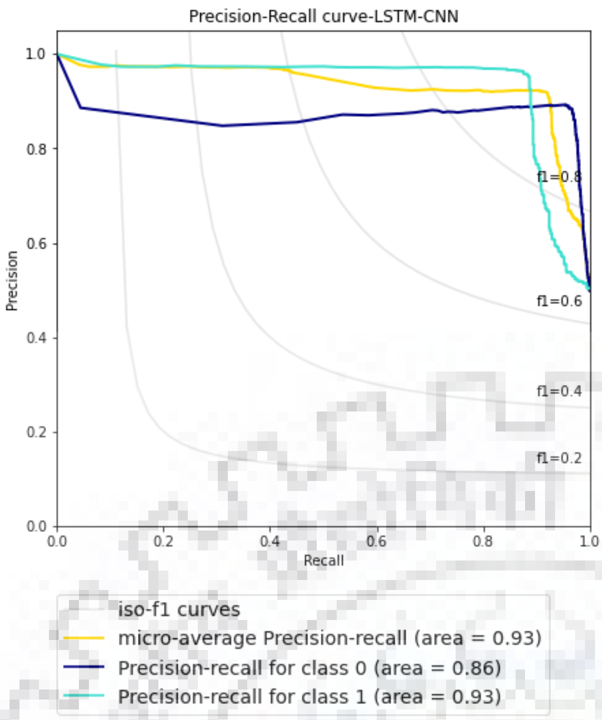


(b) ROC curve with CNN model with 0.92 value.

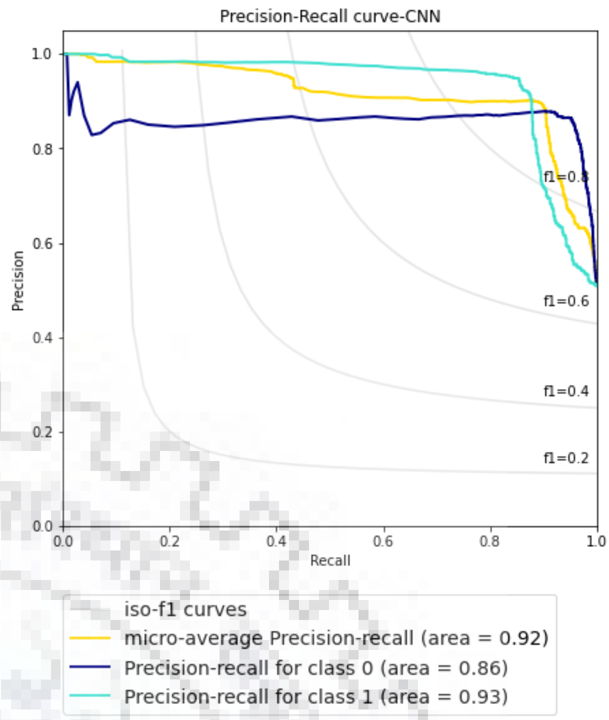


(c) ROC curve with LSTM model with 0.93 value.

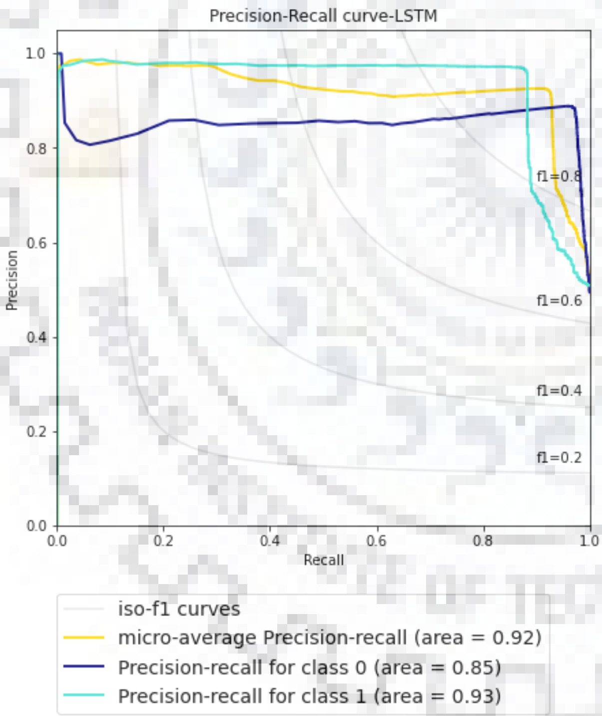
Fig. 5.9: ROC curves of all models



(a) PR-Curve of proposed model



(b) PR-Curve of CNN model



(c) PR-Curve of LSTM model

Fig. 5.10: PR Curves of all models

Approach	Accuracy	Precision	Recall	F1-score	AUC	PR Score
Proposed Model	0.92	0.93	0.95	0.93	0.94	0.93
CNN	0.90	0.90	0.90	0.90	0.92	0.92
LSTM	0.91	0.92	0.92	0.92	0.93	0.92

Tab. 5.7: Comparison results for the proposed model against other classifiers.

5.2.7 Comparison with Existing Studies

To further evaluate the performance of the proposed method, we compare the proposed method with the existing research studies that are applied for disaster-relevancy identification of an SM message, as presented in Table: 5.8. Although comparing our results with existing works is not straightforward due to differences in the used dataset, and experimental setup, our study differs due to its methodological contribution. Studies suggesting different methods previously made using DL methods are reviewed in Table: 5.8 in terms of accuracy, P, R, F-score. Accordingly, it is seen that the proposed method is comparable with previous studies in terms of accuracy, P, R, and F-score. The most recent paper on related research is of 2019, which is used in the comparison. The performance of the proposed approach is presented in the last row of Table: 5.8. It can be observed that the performance of the proposed method is better than the previous studies, considering all the evaluation metrics.

Study	Method	Accuracy	Precision	Recall	F-score	AU-ROC	AU-PR-Curve
Sit et al. (2019)	CNN; LSTM; SVM; LR	0.74	0.76	0.78	0.75	-	-
Snyder et al. (2019)	CNN; LSTM; RNN	-	0.76	0.74	0.75	-	-
Habdank et al. (2017)	MLP; RF; SVM	0.88	0.86	0.89	0.87	-	-
Nguyen et al. (2016)	CNN	-	-	-	-	0.78	-
Abel et al. (2012a,b)	-	-	0.23	0.61	-	-	-
Verma et al. (2011)	NB; Maximum Entropy (ME)	0.89	-	-	-	-	-
Li et al. (2015)	NB	-	-	-	-	0.73	-
Caragea et al. (2016b)	CNN	0.77	-	-	-	-	-
Feng and Sester (2018)	CNN; RF; SVM with linear kernel; SVM with RBF kernel; NB	-	0.89	-	-	-	-
Proposed	LSTM-CNN	**0.92**	**0.93**	**0.95**	**0.92**	**0.94**	**0.93**

Tab. 5.8: Comparison of the proposed framework with the previous studies. The datasets used in the research studies are different.

5.2.8 Decision-Making

The superiority of the proposed model is verified by comparing multiple metrics as well as previous studies; thus, we adopt the trained proposed model to identify the relevance of the SM message for decision-making as we divided the dataset into three subsets: *training, validation, and testing*. The *testing* dataset is used to identify the disaster-relevance of SM message. After the training process, the trained model is saved and called when the testing process comes into play, as illustrated in Figure: 5.11.

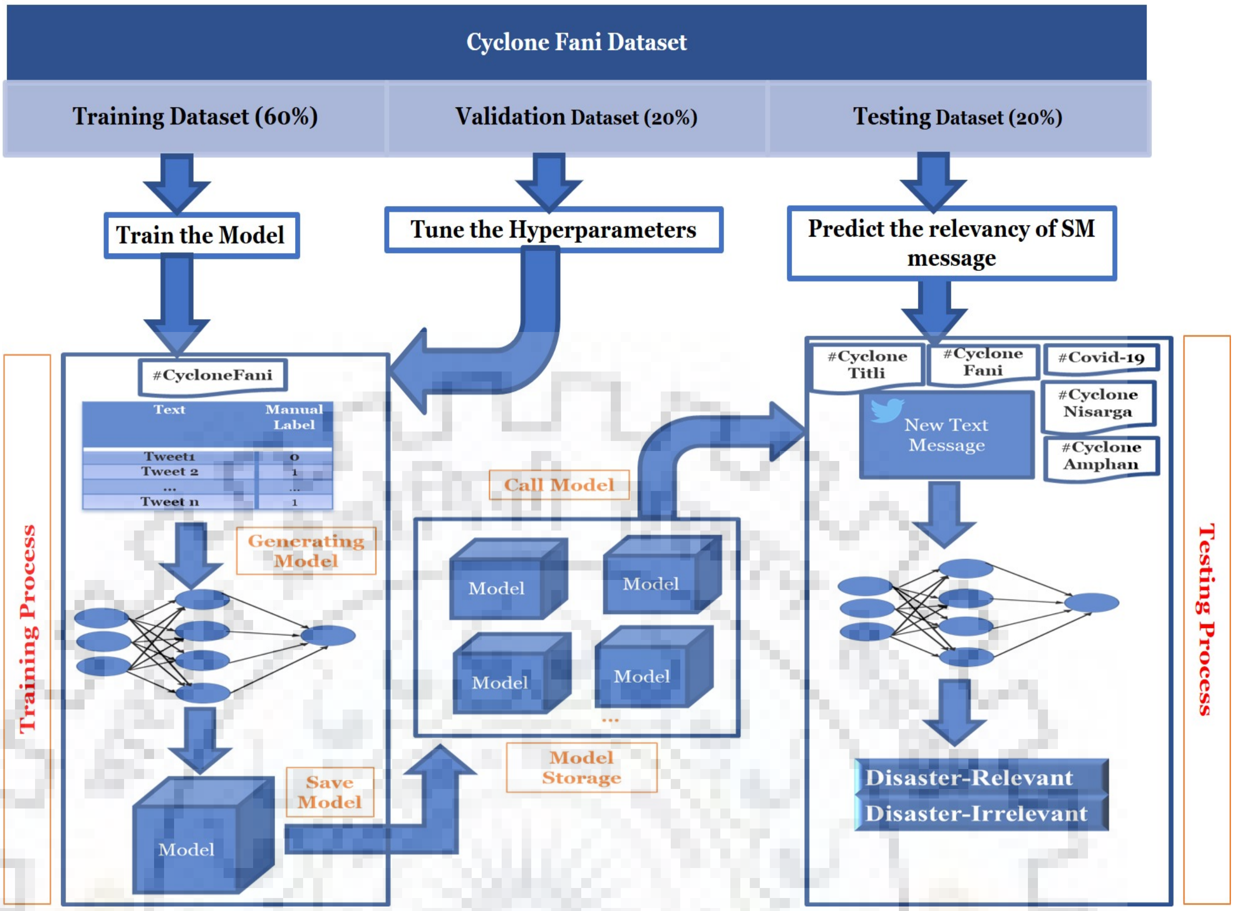


Fig. 5.11: Trained model is saved and called upon for the testing process, which is further used for identification of the SM message as *disaster-relevant* or *irrelevant*.

Table: 5.9 depicts the experimental results of the cyclone Fani dataset. *0* represents the *relevant* message, and *1* illustrates the *irrelevant message*. The table presented the manual label and predicted label, where blue color gives the correct prediction, and red color depicts the misclassification of the SM message. The results show that rows 1, 2, 3, 4, 5 represent the correct classifications, whereas the last row, Table: 5.9, is a misclassification.

Tweet no. 27131(row 1) states the *movement of cyclone storm towards Vizag within 10 hours*, making it a *relevant* message to disaster. Tweet 39785 (row 2), and tweet no 63007 (row 3) mentions about the fake account, and country development, making them *irrelevant* to the disaster. Tweet 25038 (row 4) talks about *vehicular movement in certain regions in cyclone Fani*, and tweet no 22459 (row 5) updates about the direction of cyclone Fani, categorizing them into *disaster-relevant* messages. Lastly, tweet no. 39147 (last row) is misclassified as *disaster-relevant* message.

Tweet No.	Tweet Text	Manual Label	Predicted Label
27131	@ncbn The cyclonic storm moving northwest ward drift towards the Vizag Andhra Pradesh coast by with in 10 hours. #CycloneFani @vizaggoap'	0	0
39785	I wonder who's paying these fake accounts? \n#Dubailagoon #Schon_the_con #Xanadufraud #Xanaducrooks #xanaduscam\xe2\x80\xa6 https://t.co/5e3J1KGyW9",,,,,	1	1
63007	Security and poverty free Nation, makes the country developing. You(Namo)tried a lot to connect each and every stat\xe2\x80\xa6 https://t.co/D489Q6txqO'	1	1
25038	#CycloneFani \nVehicular movement banned in most vulnerable Ichapuram, Kaviti, Vajrapu Kothuru, Kanchili & Sompet bl\xe2\x80\xa6 https://t.co/sAe3HUznpz'	0	0
22459	ESCS Fani about 40 km south-southwest of Puri at 0730 IST. To cross Odisha coast between Gopalpur and Chandbali clo\xe2\x80\xa6 https://t.co/E58JDjSgRv'	0	0
39147	A wind of positivity in the time wholesale negative narrative by 1 side & over the top nationalism campaign from th\xe2\x80\xa6 https://t.co/BKvCEA5HbS'	1	0

Tab. 5.9: Tweets with predicted relevancy label of the Cyclone Fani disaster. 0 represents *Relevant*, and 1 illustrates *Irrelevant*. The blue color represents the correct predictions, and the red signifies incorrect predictions.

In order to explore the technical efficacy of the iRelevancy, we identify the relevancy of the SM message of other cyclonic disastrous events. We consider cyclone Amphan, cyclone Nisarga, and cyclone Titli. Table: 5.10 depicts the predictions of disaster-relevance of SM message on different cyclonic events. The green color represents disaster-relevant messages, and the brown color represents disaster-irrelevance.

Tweet no.1 (row 1) mentions *road damage in cyclone Titli with amount of loss*, and is identified as a *disaster-relevant* message. Tweet no.2 (row 2) states the *affected families in cyclone Titli*, and tweet no.3 (row 3) mention *ravaged Srikakulam district* bacause of cyclone Titli, classifying them as *disaster-relevant* messages. Tweet no.5 (row 5) and tweet no. 6 (last row) mentions *electricity restoration* in cyclone Nisarga and *help and relief* in cyclone Amphan, respectively, identifying them as *relevant* tweets. Tweet no. 4 (row 4), and tweet no. 89 (penultimate row) are classified as *irrelevant tweets*.

Sr No.	Tweet Text	Predicted Label
1.	b'RT @naralokesh: 369 kms. of roads have been damaged by **#CycloneTitli** amounting to a loss of Rs. 300 Cr. All fallen trees and traffic interr\xe2\x80\xa6'	0
2.	b'RT @ncbn: Around 86% of the 2,45,881 affected families have received food commodities as on 17th Oct (17:30 PM). Efforts are being made to\xe2\x80\xa6' **#CycloneTitli**	0
3.	b'RT @naralokesh: **#CycloneTitli** has ravaged Srikakulam dist beyond recognition. It\xe2\x80\x99s time to help our brothers & sisters in distress. Pls don\xe2\x80\xa6'	0
4.	b'RT @VenkataRamanaJu: Mr @pawankalyan seems like your latest movie about loving #Uttarandhra to the core has failed miserably in the time of\xe2\x80\xa6'**#CycloneTitli**	1
5.	**#CycloneNisarga**Electricity has ben restored in 784 villages	0
89	**#CycloneAmphan** Dinesh Karthik wants to win KKR IPL-2020	1
6.	**#CycloneAmphan** victims need help and relief	0

Tab. 5.10: Some Predicted Labels on Cyclone Amphan, Cyclone Nisarga, and Cyclone Titli datasets. 0 represents *disaster-relevant messages* in green color, and 1 illustrates *irrelevant messages* in brown color.

To investigate the reusability of the proposed framework for disaster events, other than cyclones, we predict the message relevance of the Covid-19 dataset. Table: 5.11 represents the predicted relevance label of the SM message of Covid-19. The results demonstrate that the framework works well on non-cyclonic events as well.

Tweet no.1(row 1, Table: 5.11) mentions *death toll*, classifying it as *disaster-relevant* message. Tweet no.2 and 3 are *disaster-irrelevant* messages.

Tweet No.	Tweet Text	Predicted Label
1.	b'RT @LabGrassroots: #Latest : \xe2\x80\x9cUK **#COVID-19** death toll passes 55,000 including suspected cases\xe2\x80\x9d via \xe2\x81\xa6\xe2\x81\xa6@ReutersUK\xe2\x81\xa9 https://t.co/85v8YCntfg'	0
2.	**#Covid-19** b'RT @BBhuttoZardari: The PTIs consistent assault on HEC funding is having very real consequences at a time medical universities are at the f\xe2\x80\xa6'	1
3.	b'How to use energy healing to sedate your triple warmer -\nhttps://t.co/ahB6tAIoem\n\n#energyhealing #inspiration\xe2\x80\xa6 https://t.co/poB7yIJxNO' **#Covid-19**	1

Tab. 5.11: Predicted Labels on Some Tweets of Covid19 dataset. 0 represents *disaster-relevant messages* in green color, and 1 illustrates *irrelevant messages* in brown color.

5.3 Conclusion

Automatic identification of the SM messages to disaster relevance is the need of the hour. This chapter presents iRelevancy: a hybrid framework using DL algorithms. The objective of this study is to develop a novel hybrid framework to identify the relevance of a tweet to a disaster. We sequentially complete the aforementioned task in phases: (1). acquiring cyclonic and pandemic data from the Twitter platform using search API; (2). manually inspecting and annotating the data into a *relevant* and *irrelevant* category; (3). pre-processing the raw and unstructured data; (4). development of the hybrid model; (5). evaluate the model's performance on different metrics; (6.) compare the proposed approach to CNN and LSTM models; as well as with the existing research studies; and finally, (7). identification of the tweet as *relevant* or *irrelevant* to the disaster.

In the empirical section, we conduct extensive experiments against different neural networks and previous research studies on the relevant topic to answer our RQs and validate the superiority of our model. To exhibit the technical potency of iRelevancy, we identify the relevancy of cyclonic disasters, considering cyclone Fani, cyclone Amphan, cyclone Nisarga, and cyclone Titli. Indeed, iRelevancy is successful in identifying the disaster-relevance of SM messages on datasets of varying nature with efficiency. We re-use iRelevancy on pandemic data with #Covid-19 keywords, and the results demonstrate that the proposed framework accurately identifies the SM message to be Covid-19 relevant or irrelevant to it.

We believe that our conclusions have promising implications for automatically identifying disaster-related SM messages for effective DM and combating the ever-growing challenges of SM usage in disasters.

6. AN INTELLIGENT HYBRID DEEP LEARNING-BASED FRAMEWORK TO DETERMINE THE STAGE OF DISASTER MANAGEMENT CYCLE FROM A SOCIAL MEDIA MESSAGE

6.1 Introduction

DM is a domain of national priority as India is among the top ten countries in terms of disaster losses between 1998 and 2007, estimating a loss of US$ 79.5 Billion. Success in managing a disaster is hugely dependent on accurate decisions. The right decision at the appropriate time is imperative as the Sendai framework highlights the need to build on prevention and preparedness to reduce disaster losses (Fanchiotti et al., 2020). Efficient management of disasters is enhanced with first-hand information from SM. SM data has diverse functions across all the periods of DM cycle (Anson et al., 2017).

The enormous amount of SM data is a challenge in disastrous times as the meaningful information is immersed in irrelevant content. Mining SM data is imperative for effective DM. Hence, identifying the stage of the disaster from SM message allows first responders to assess more effectively what stage disaster is and coordinate activities in a more targeted and timely manner. From a practical point of view, it is a strategic advantage for decision-makers about the stage of disaster for team configuration and mission preparation.

The fact that the success of DL is parallel to the increase in decision-making capacity has been the source of motivation for this research work. DL models represent data in multiple and successive layers. They have the ability to capture the syntactic features from SM messages automatically without extra feature extracting techniques. Therefore, to ensure that the DL model enables improved handling of SM data, we propose iStage: an intelligent framework to address the above problem. As per the authors' knowledge, no such work uses a hybrid model to determine the stage of the disaster from SM message. The study emphasizes on the efficieny, and effectiveness of the iStage. This research is in line with the ongoing efforts to use DL for providing advanced DM assistance.

The novelty of the work lies in proposing iStage: an intelligent framework to determine the stage of the disaster, i.e., *pre, during, post, or irrelevant* from SM message. More specifically, we have taken Chowdhury et al. (2013) as the base reference for the research study. As a contribution to the society, the research study has demonstrated with facts how is it predicting the messages. A few examples of typical use of SM during different stages of a disaster, i.e., *pre, during, and* *post-disaster*, are shown in Table: 6.1.

Tweet Message	**Category**
b' India set to evacuate a million people as Cyclone Fani on path to hit the eastern coastline #CycloneFani	Pre (0)
b' #fanicyclone #cyclonefani Cyclone Fani Severe storm strikes Odisha, make landfall 10km north	During (1)
We also need your help to grow up and rebuild ourselves after a devastating cyclone #CycloneFani	Post (2)

Tab. 6.1: Examples of Pre, During, and Post Disaster Tweets. These are taken from the datasets described in our paper.

These SM messages are taken from the datasets described in the section: 6.3.1. It can be observed that the SM messages contain keywords and hashtags like #CycloneFani and #FaniCyclone. Some of the messages do not have any period relevant information; instead contain irrelevant information in the tweet with the trending hashtag. Consequently, the study sets out to respond to the following research questions:

1. How can SM be used to determine the stage of disaster?
2. Which DL techniques are more effective in this context?

Identifying stage-based SM messages from an enormous dataset of SM messages is considered a tedious task. The hybrid DL model proposed by us in this paper can classify and determine the stage of the disaster from SM messages of the Twitter platform so that timely action can be taken. The study contributes to decision-making for DM by:

1. Proposing a cyclone disaster dataset comprising of SM message and the label of the stage of disaster, i.e., *pre-, during-, post-disaster* or *irrelevant* to disaster.
2. Using real-world SM data to verify that iStage can correctly determine the stage of the disaster.
3. Compare different DL models in terms of performance.
4. Computationally evaluating and extensive experimenting with different disaster-related datasets.
5. Demonstrating the technical efficacy by determining the stage of cyclonic disasters.
6. To investigate the reusability of the proposed approach to other disaster events, we apply our proposed technique to SM messages posted during Covid-19 event in 2020. We see that our method is effective for different disaster types.

We initiate this chapter with methodology, followed by an introduction. Section: 6.2 [Page: 98] describes the proposed approach, subsequently followed by the experimental work in section: 6.3 [Page: 98]. The conclusions are offered in section: 6.4 [Page: 113].

6.2 Proposed Approach

This section explains the methodology of the iStage for stage identification of a disaster, considering SM messages. The research flowchart of the iStage is illustrated in Figure: 6.1. The iStage requires SM data of disaster and determines the stage of the disaster. The raw data is acquired from SM, following the manual annotation. Subsequently, data is pre-processed involving several steps. The hybrid model consisting of LSTM and CNN layers is proposed to determine what stage disaster is in. The model is evaluated on performance metrics. In the next step, the performance is compared against state-of-the-art and existing research studies. Lastly, the predictions are carried out on datasets. We validate our proposed approach using real-world datasets from Twitter. In particular, we adapt our approach to focus specifically on cyclonic and pandemic disasters, and we show its effectiveness experimentally.

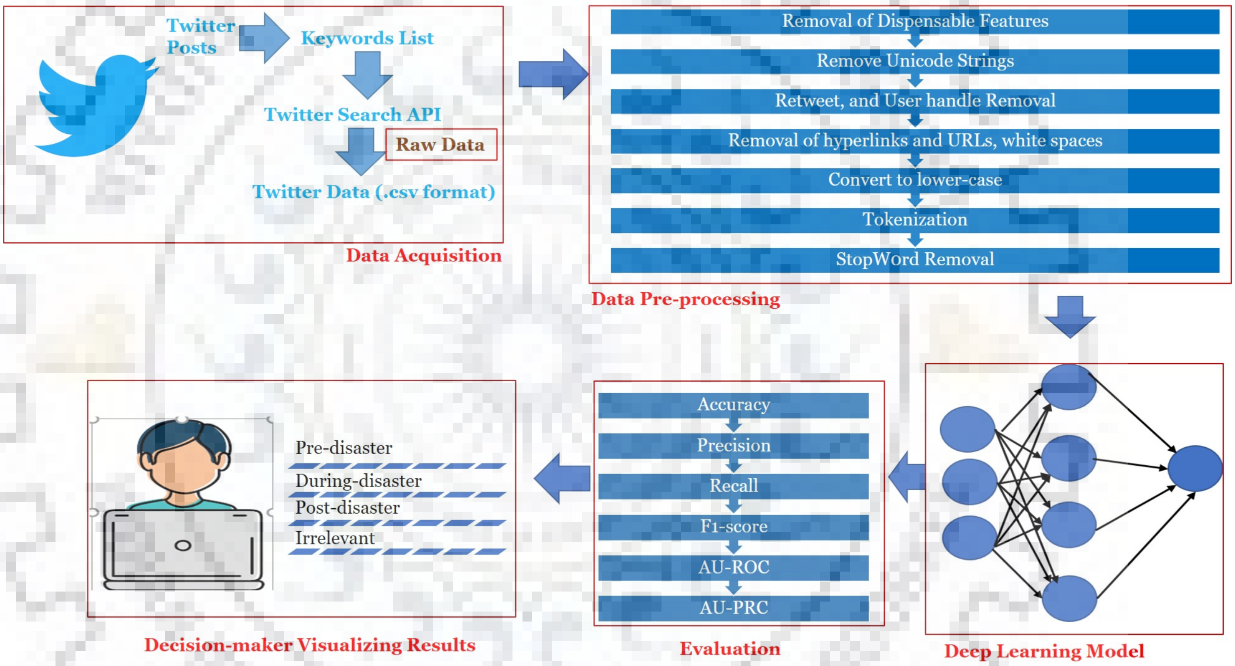


Fig. 6.1: iStage Framework to Determine the Stage of Disaster from SM Message.

6.3 Experiments

In this section, each phase of the framework is described concretely. The entire process consists of data selection and acquisition, manual annotation, data pre-processing, model building, performance evaluation, and experimental results, and analysis.

6.3.1 Data Selection and Acquisition

Twitter has been chosen over other SM platforms for several reasons. First of all, data is public. Secondly, the usage of hashtags fosters information diffusion. We have used multiple real-world datasets for our experiments in this paper. We consider SM posts during cyclone and pandemic events.

The disaster events considered in this study are:

- **Cyclone Fani**: An extremely severe cyclonic storm Fani formed over the Bay of Bengal during April 26- May 04, 2019. 16.53 million people are affected. 1053 primary schools and more than 5 lakh tree have been severely damaged (Society, 2019).

- **Cyclone Titli**: Very severe Cyclone Titli made landfall over Srikakulam district in Andhra Pradesh in October 2018. The Andhra Pradesh government reports seven deaths and a loss of Rs.3680 crore (Dash and Walia, 2020).

- **Cyclone Nisarga**: Cyclone Nisarga was the first cyclone near Mumbai city on the Arabian Sea in a long time. It made landfall near Alibaug in June 2020 (Anjali Marar, 2020).

- **Cyclone Amphan**: Super cyclone Amphan struck the Sundarbans delta of eastern India devastatingly in May 2020. It has caused mammoth losses to resources and properties (Goswami et al., 2021).

- **Covid-19**: Covid-19, originally known as Corona Virus Disease of 2019. On March 11, 2020, World Health Organisation (WHO) had declared Covid-19 as a pandemic. There have been 39.6 million cases worldwide until October 18, 2020, resulting in more than 1.1 million deaths (Wang et al., 2021).

The descriptive statistics of the datasets are depicted in Table: 6.2.

Disaster	Hashtag Used	Dates	# of Tweets
Cyclone Fani	#CycloneFani	May 03-10, 2018	1,72,907
Cyclone Amphan	#CycloneAmphan	May 16-21, 2020	8214
Cyclone Titli	#CycloneTitli	October 08-12, 2018	96,286
Cyclone Nisarga	#CycloneNisarga	June 01-04, 2020	81,475
Covid-19	#Covid19	August 08-15, 2020	76,953

Tab. 6.2: Descriptive Twitter Statistics of Disaster Events.

We used Twitter Search Application Programming Interface to gather tweets with our web crawler developed using Python on Anaconda Integrated Development Environment. Only English language tweets are considered for the study. The fetched data is raw and unstructured. A word cloud of most frequently used words in the Cyclone Fani dataset is demonstrated in Figure: 6.2. The size of each word represents descending frequency. Higher font signifying most frequency. The most frequent words of the Cyclone Fani dataset have been illustrated mathematically in Figure: 6.3.

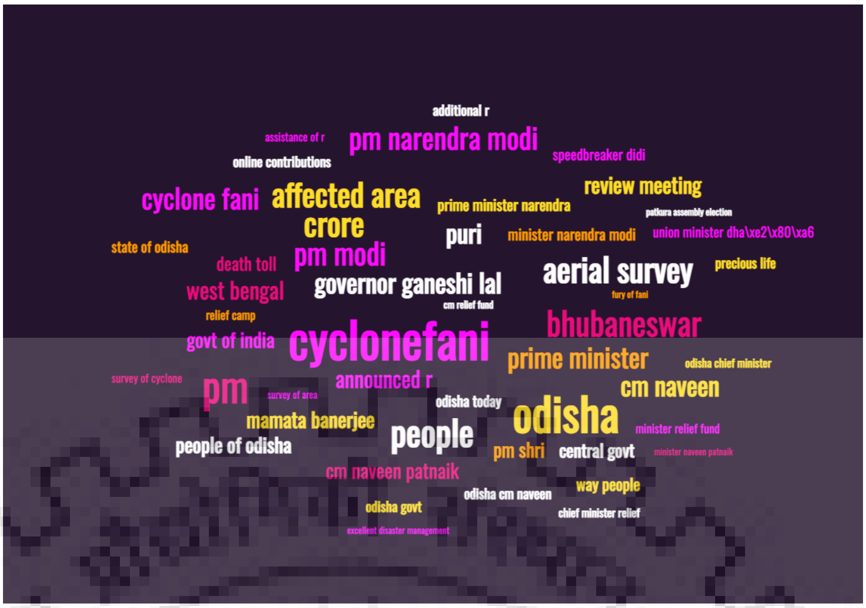


Fig. 6.2: Most Used Words represented by Word Cloud of Cyclone Fani.

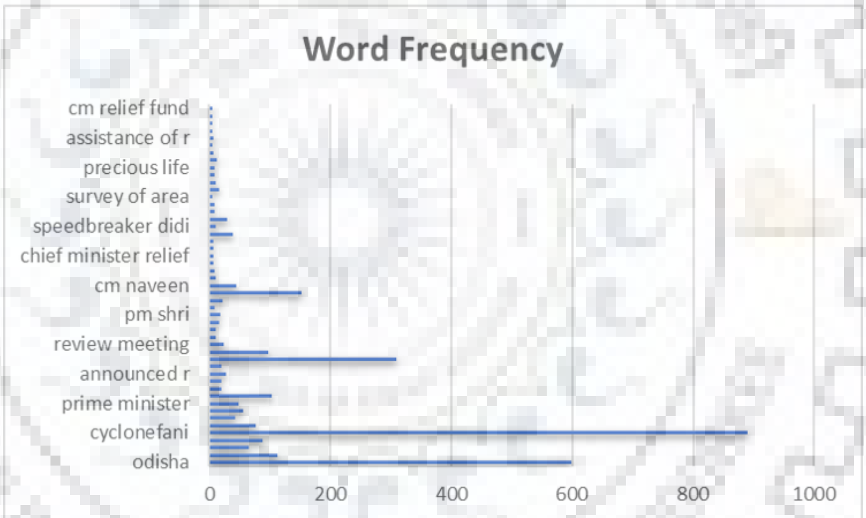


Fig. 6.3: Word Frequency of Cyclone Fani.

6.3.2 Manual Annotation

To determine the stage of the disaster, we need a labeled dataset of cyclone Fani to conduct the experiment, and to the best of our knowledge, a dataset with stage label of cyclone Fani data are not available till now. Therefore, we prepare the dataset with labeled stage annotation for the aforesaid purpose as first contribution mentioned in Table 3.3 . We divide the disaster into three stages: *pre, during, post-disaster*. We show the annotation scheme with the class label below:

0. *Pre-Disaster SM messages*: messages those contain information about the prevention, mitigation, preparedness, warning, etc.

1. *During-Disaster SM messages*: messages damage, search, first-aid, evacuation, etc.

2. *Post-Disaster SM Messages*: tweets that provide information related to rescue, rehabilitation, recovery, reconstruction, etc.

3. *Irrelevant SM Messages*: tweets that contain no information related to the disaster. The tweets contain #CycloneFani but no disaster-relevant information.

Figure 6.4 showcases the tweet distribution in each of the categories. After removing duplicates from the dataset, 1,11,746 tweets are left in the dataset. Table 6.3 depicts the tweet distribution statistics.

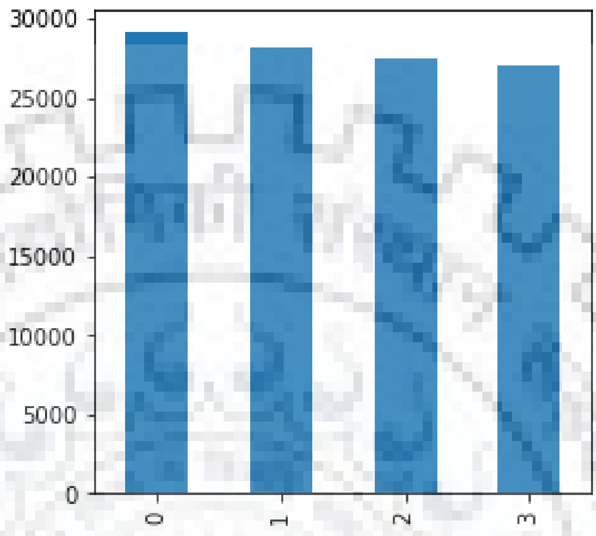


Fig. 6.4: Tweet distribution after manual annotation. Here *0* represents the *pre-disaster*, *1* represents *during-disaster*, *2* represents the *post-disaster*, and *3* represents the *irrelevant* category of SM messages.

Label	# of Tweets
0	29120
1	28114
2	27454
3	27058

Tab. 6.3: Tweet Distribution Statistics After Manual Annotation.

6.3.3 Data Pre-processing

The collected real-world dataset is raw, unstructured, and meaningless. Therefore, data needs to be pre-processed in order to ensure the reliability of the knowledge discovery from the data. We pre-process the text following the below-mentioned steps:

- *Removing dispensable features*: The fetched data contains 20 features mentioned in Table: 5.1. The research study requires only the 'text' and 'manually annotated label' features. Hence, all the features except 'text' and 'manual annotation label' are eliminated.
- *Retweet Removal*: Retweets repeat the old information. They have nothing new to offer. Therefore, the tweets starting with 'RT' are eliminated.
- *Removal of URLs and Hyperlinks*: The research study is limited to textual content. Hence, all the URLs and hyperlinks are eliminated.

- *Removal of User Mentions, Hashtags, Special Characters, and White Spaces*: User handles reveal user information. Hence, we ensure privacy by eliminating user handles. User handles with '@,"#,' punctuation characters, and white spaces are removed from the dataset.

- *Removing StopWords*: Stopwords are words that do not add much meaning to the sentence. 'A,"an,"the,' etc. are some of the stopwords. NLTK stopwords list is eliminated. In this research study, 179 stopwords are eliminated.

- *Split Camel Case and Case-folding*: Camel case strings are those in which spaces do not separate the words, and every capital letter signifies the beginning of a new word. They are split, following the words being folded into lower case.

- *Tokenization*: This step chops the tweets into tokens, i.e., words.

6.3.4 Model Training

In the present study, we develop a DL-based hybrid model composed of multiple layers of the cascade. The model consists of an LSTM layer, which receives the word embeddings for each token in the tweet as inputs. The LSTM layer generates a new encoding for the original output. Subsequently, the LSTM layer's outcome is pooled to a smaller dimension. The output of the pooling layer is fed into the convolution layer, where it extracts the local features. Dropout layers are used to overcome the overfitting issues. The dense layer ultimately outputs the tweet as *pre-, during-, post-disaster* or *irrelevant* to the disaster. The pseudocode of the proposed model is summarized in Algorithm 2. Figure: 6.5 elucidates the proposed model, encapsulating the layers.

Algorithm 2: Pseudo-code for Proposed Hybrid Model

Input: Training dataset, Validation dataset, Manual label of messages

Output: Label Prediction of SM message

Begin
For number of training iterations **do**
Construct sentence matrix using embedding layer
Employ LSTM to learn the contextual features of text sequences
Reduce the dimensionality by maxpooling layer
Apply ReLU activation function *f(x)*= max(0,x) to convolution layer
Add dropout to prevent overfitting
Feed the comprehensive representations into the softmax classifier to get the class labels
Update parameters of the model using the loss function with the Adam method
The performance metrics (Accuracy, P, R, F-score, AU-ROC, and AU-PR-Curve) are calculated using equations in Table: 5.6
End

We randomly divide the dataset into 60% training, 20% validation, and 20% testing datasets (Alam et al., 2018; Tam et al., 2021). The model is trained using a training dataset. In order to find the best hyperparameters for our model, we perform a random grid search on the validation dataset.

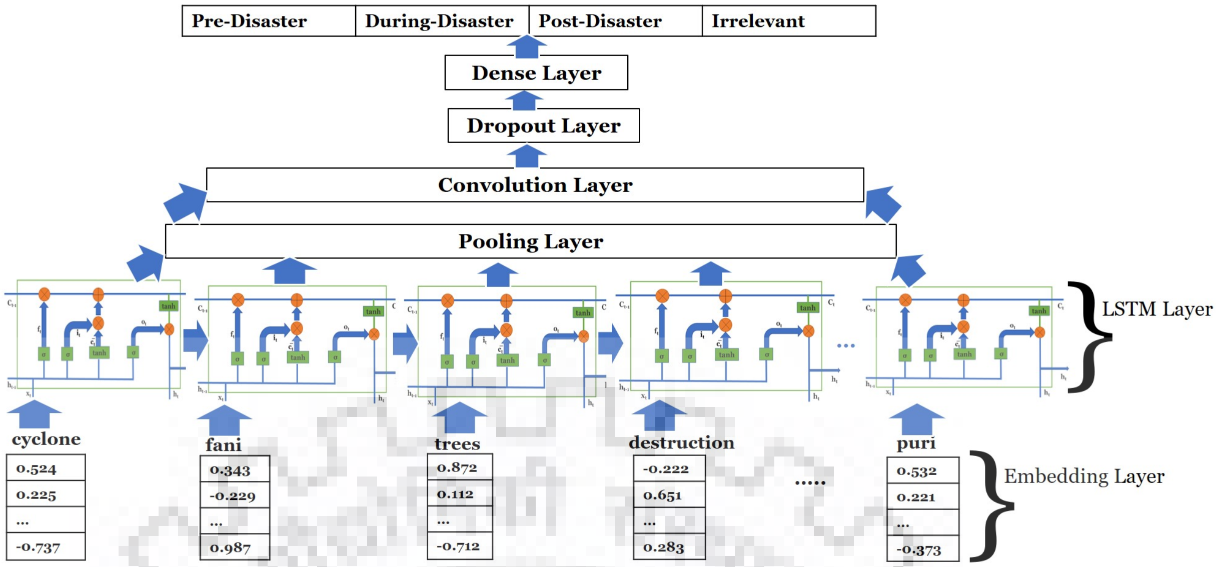


Fig. 6.5: Architecture of the proposed hybrid model for SM message classification into *pre-, during, post-disaster*, and *disaster-irrelevant*. The model consists of embedding, LSTM, pooling, CNN, dropout, and dense layers.

The hyperparameters used in the research study are showcased in Table: 6.4. Fig. 6.6 presents the proposed model generated by the Anaconda Integrated Development Environment.

6.3.5 Evaluation

In order to evaluate the performance of the proposed approach, the focus is not only on one performance metric; instead, multiple metrics are considered. Accuracy, P, R, F1-score are the performance metrics. Furthermore, Area under ROC and Area under PR-Curve are deployed to measure the completeness and robustness of the framework. The performance metrics used in this research study are defined in Table: 5.6.

6.3.6 Results and Analysis

In this subsection, we conduct experiments on the cyclone Fani dataset. The experimental results are then compared against state-of-the-art and existing research studies.

Comparison with Other models

The experiments have been conducted using the proposed model, CNN, and LSTM models. The implemented CNN and LSTM models generated by Anaconda Integrated Development Environment are displayed in Figures: 6.7a, and 6.7b, respectively. The model performance results are described in Table: 6.5.

ROC curve for our proposed model, implemented on cyclone Fani dataset, is illustrated in Figure: 6.8a. Likewise, the ROC curve for cyclone Fani dataset using CNN and LSTM architecture is shown in Figure: 6.8b and 6.8c, respectively, taking TPR and FPR. We apply two approaches apt

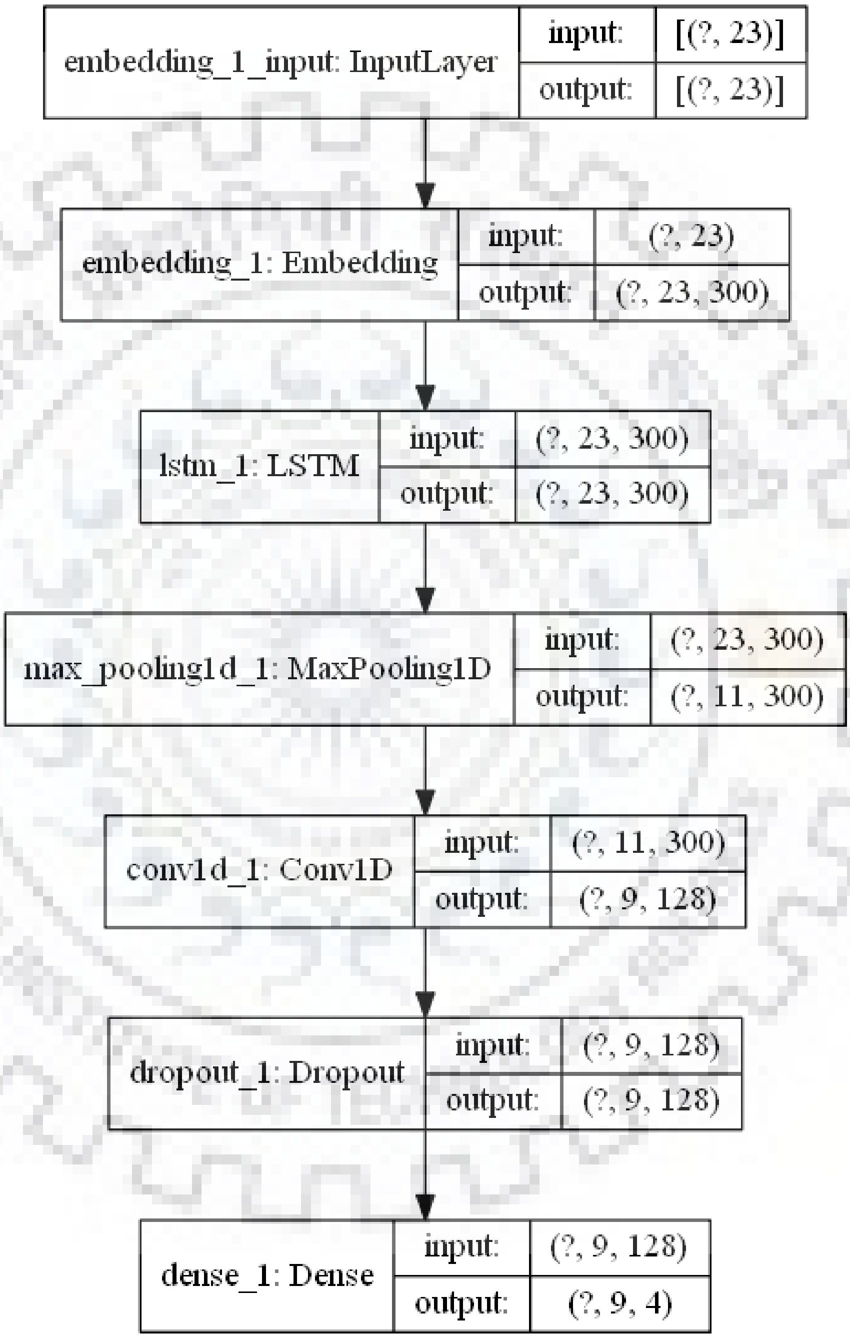
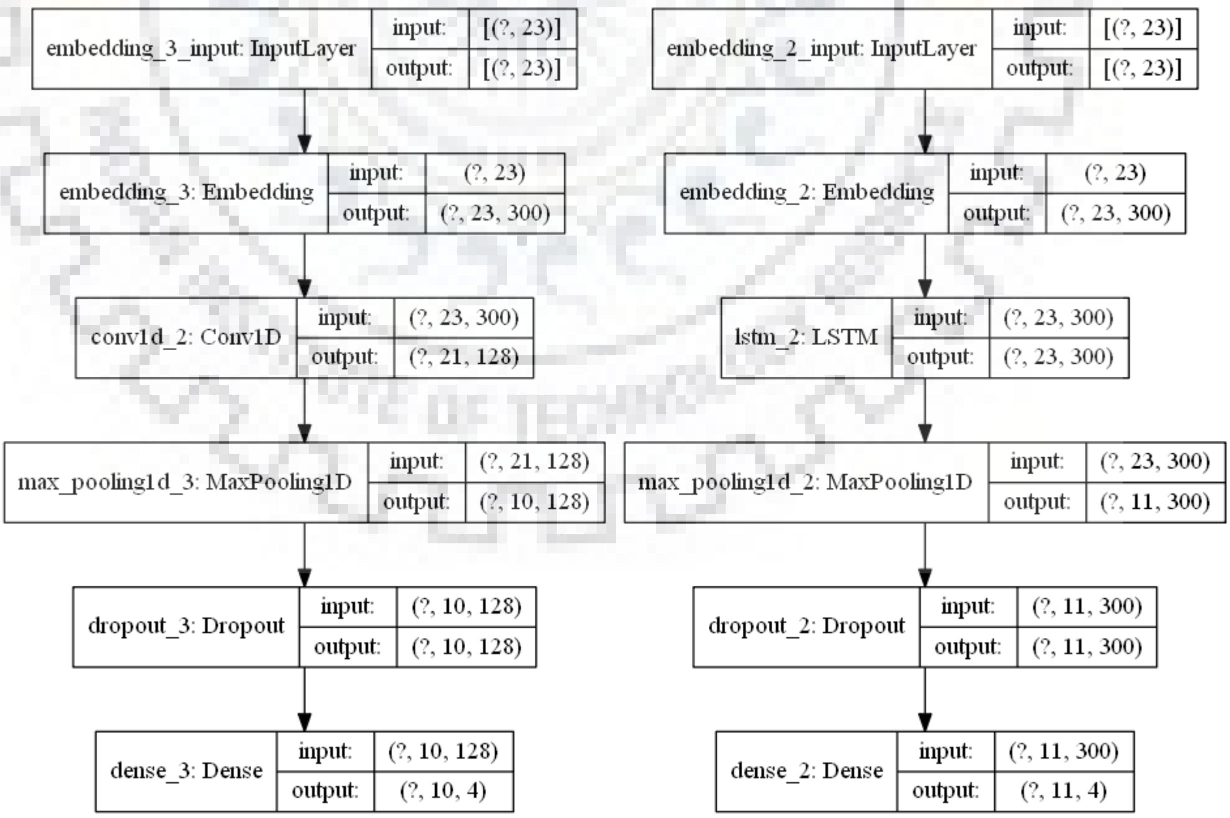


Fig. 6.6: Implementation of the proposed model on cyclone Fani dataset, generated by Anaconda platform.

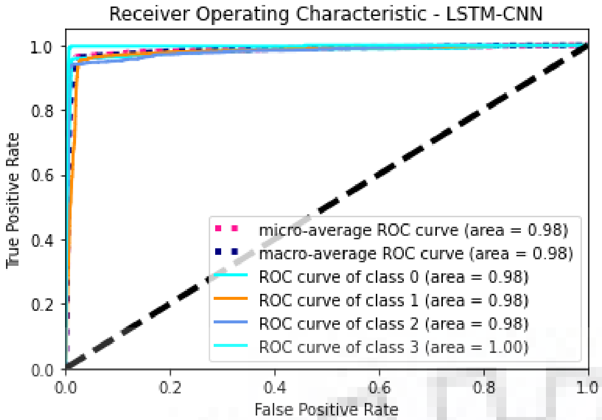
Hyperparameter	Definition	Value
Epochs	Iteration count	25
Embedding Dimension	Size of vector used to represent each of the word embedding	300
Maximum Sequence length	Maximum tweet-length	23
Activation Function	Calculates the weighted sum of its input, adds a bias and then decides whether to activate the neuron or not	ReLU
Dropout Rate	Some hidden layer neurons are discarded with the 40% probability	0.4
Optimizer	The method used to update the weights in order to reduce the error	Adam(1e-3)
Batch Size	Data is grouped into batches prior to feeding it into DL model	1024
l2 regularizer	A regularizer to apply penalties on layer parameters during optimization to prevent overfitting	0.001
Loss Function	Function to assess model prediction	Categorical CrossEntropy

Tab. 6.4: Definition of technical parameters and their values used in the proposed approach.

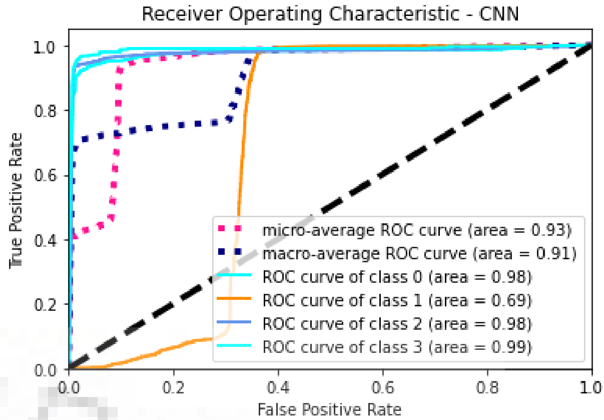


(a) CNN *(b)* LSTM

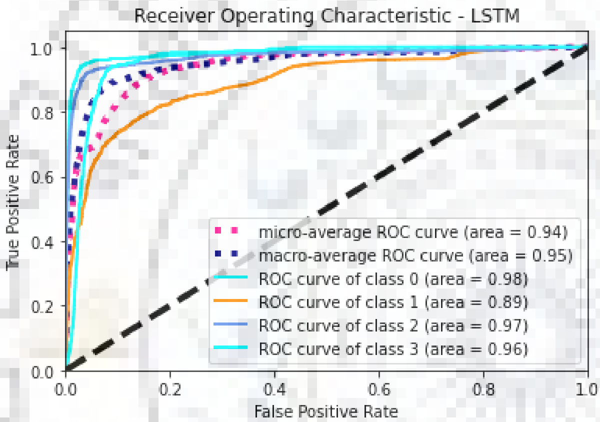
Fig. 6.7: Implemented CNN, and LSTM models, generated by Anaconda Integrated Development Environment.



(a) Macro- and micro-averaged ROC curves with iStage model. The curves almost overlap, and the AUCs are almost identical with 0.98 and 0.98 values, respectively.



(b) Macro- and micro-averaged ROC curves with CNN model. The curves almost overlap, and the AUCs are almost identical with 0.91 and 0.93 values, respectively.



(c) Macro- and micro-averaged ROC curves with LSTM model. The curves almost overlap, and the AUCs are almost identical with 0.95 and 0.94 values, respectively.

Fig. 6.8: ROC curves of all models. The results illustrate the higher AU-ROC value of the proposed model than other models.

for an all-inclusive performance in multi-class problems, i.e., *micro-average* and *macro-average*. In the *macro-average* approach, the ROC curves for the classes (4 in our case) are averaged regardless of their frequencies. In the *micro-averaged* method, the ROC curve is calculated afresh by weighing ROC curves by the relative frequencies of the respective classes and then averaged. In our research study, the AUC score of the proposed model is 0.98 for *macro-* as well as *micro-average*. AUC for the ROC curve using CNN is 0.93 for *micro-average* and 0.91 for *macro-average*. Likewise, AUC for the ROC curve considering the LSTM model is 0.95 and 0.94 for *macro-average* and *micro-average*, respectively. The higher the AUC score, the better is the model performance. Hence, our proposed model outperforms both CNN and LSTM models, and it acts an a contribution fulfilling the research objective of comparing different DL models in terms of performance.

The PR-Curve in Figure: 6.9 illustrates the performance comparison of different models. The AU-PR-Curve of the proposed approach, i.e., 6.9a, is 0.96, which is larger than the others, i.e.,

0.93 for LSTM, and 0.94 for CNN, shown in 6.9b, and 6.9c, respectively.
The accuracy vs. epochs curves is plotted in Figure: 6.10. The proposed approach achieves a higher accuracy score (0.96), illustrated in Figure: 6.10a, than the other models (0.94 for CNN, shown in Figure: 6.10b, and 0.94 for LSTM, displayed in Figure: 6.10c.

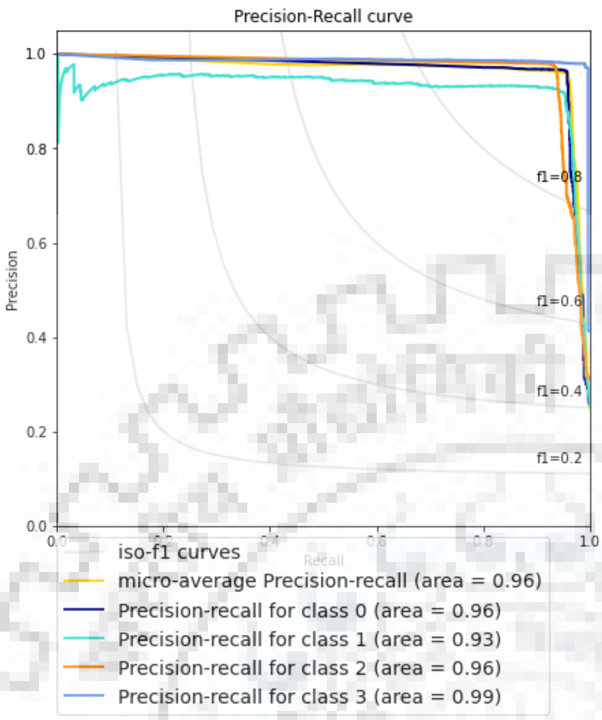
A higher P-value of the proposed approach (0.96) than the CNN (0.94) and LSTM (0.94) corresponds to column II, Table: 6.5. A higher value of R and F1-score using the proposed approach (0.96) compared to other models' R and F1-score values (0.94 for CNN and LSTM) are exhibited in columns III and IV of Table: 6.5 respectively. The comparison of performance metrics is illustrated pictorially in Figure: 6.11.
It can be seen that the proposed model contributes significantly towards the performance. The experimental results re-enforce our belief that incorporating the proposed approach increases the efficiency of iStage.

Approach	Accuracy	Precision	Recall	F1-score	AUC	PR Score
Proposed Model	0.96	0.96	0.96	0.96	0.98	0.96
CNN	0.94	0.94	0.94	0.94	0.93	0.94
LSTM	0.94	0.94	0.94	0.94	0.94	0.93

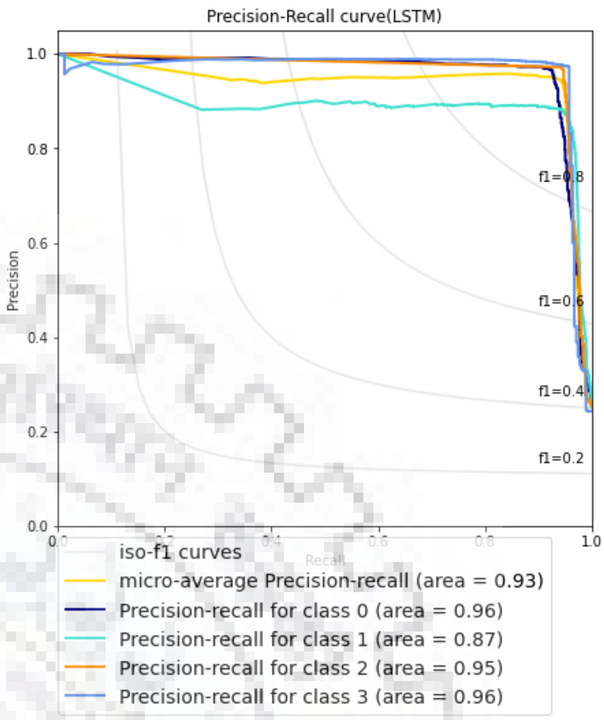
Tab. 6.5: Comparison results for the proposed model against other classifiers

Comparison with Existing Studies

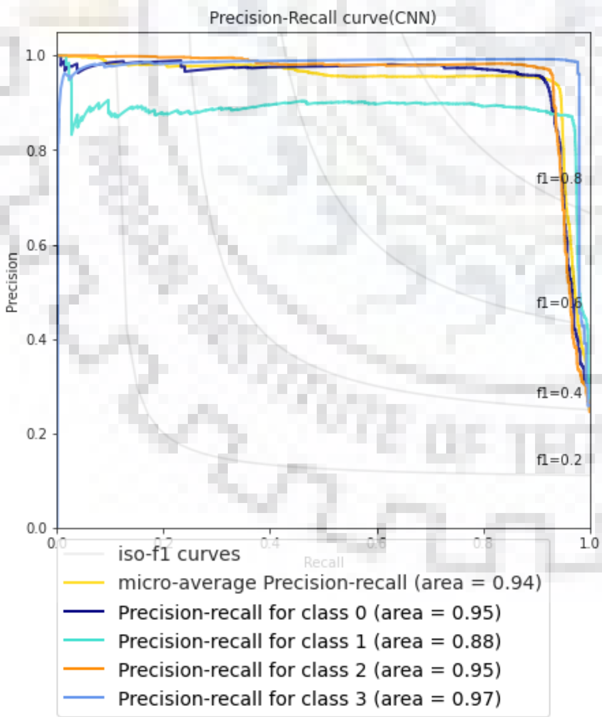
The research study depicts that the proposed hybrid architecture is highly effective in determining the stage of the disaster from the SM message. Although comparing our results with existing works is not straightforward due to differences in the used datasets and experimental setup; our study differs due to its methodological contribution. In addition, compared to other previous studies, our study is based on DL methods. Studies suggesting different approaches are reviewed in Table: 3.4 in terms of accuracy, P, R, F-score, and AU-ROC. Accordingly, it is seen that the proposed method is comparable with previous studies in terms of chosen evaluation metrics.
The most significant advantage of the iStage is the hybrid architecture which makes full use of CNN and LSTM models. To further evaluate the performance of the iStage, we compare the proposed work with the (Chowdhury et al., 2013). A comparison of iStage with existing literature is presented in Table: 6.6. iStage has higher P-value, i.e., 0.96, whereas (Chowdhury et al., 2013) has a 0.69 P-value, i.e., a significant improvement of 39.13%. iStage has higher R value than the reference article (0.96 vs. 0.93), i.e., an improvement of 3.22%. In addition, the F-score value is higher for iStage, i,e, 0.96, whereas the reference article has 0.71, a substantial improvement of 35.21%. Hence, it can be said that the iStage performs better than the existing literature study in determining the stage of the disaster from the SM message. The improvement %age is calculated using the equation: 6.1.



(a) PR-Curve of the proposed model, with a micro-averaged area of 0.96

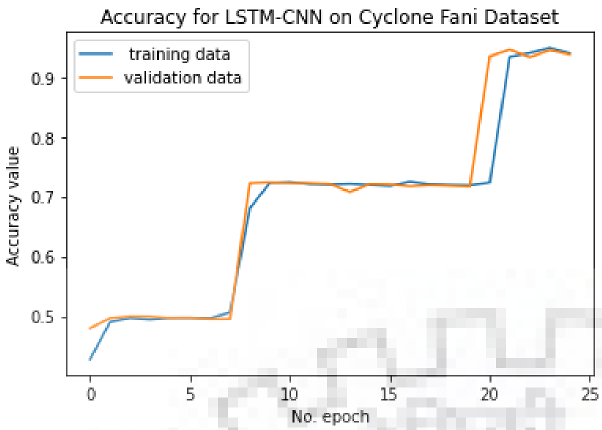


(b) PR-Curve of the LSTM model, with a micro-averaged area of 0.93

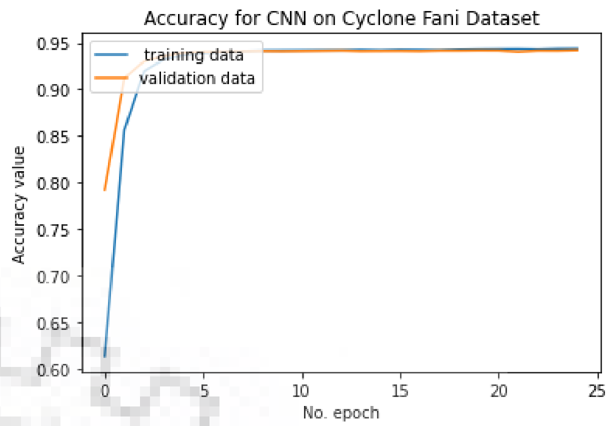


(c) PR-Curve of the CNN model, with a micro-averaged area of 0.94

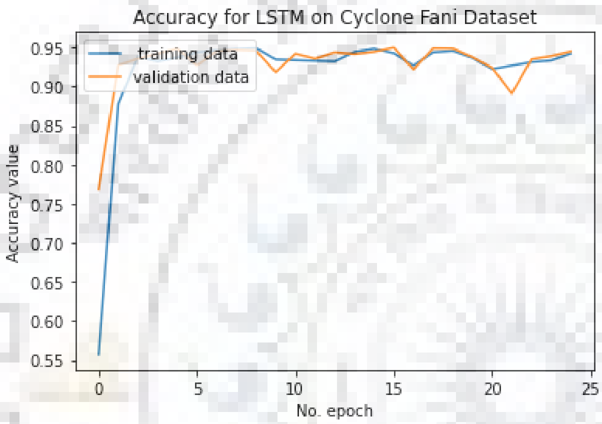
Fig. 6.9: PR-Curve of the models used in the research study. The results illustrate the area of the proposed model is higher than other models.



(a) Accuracy



(b) 0.94



(c)

Fig. 6.10: Accuracy vs. Epoch plots of all models.

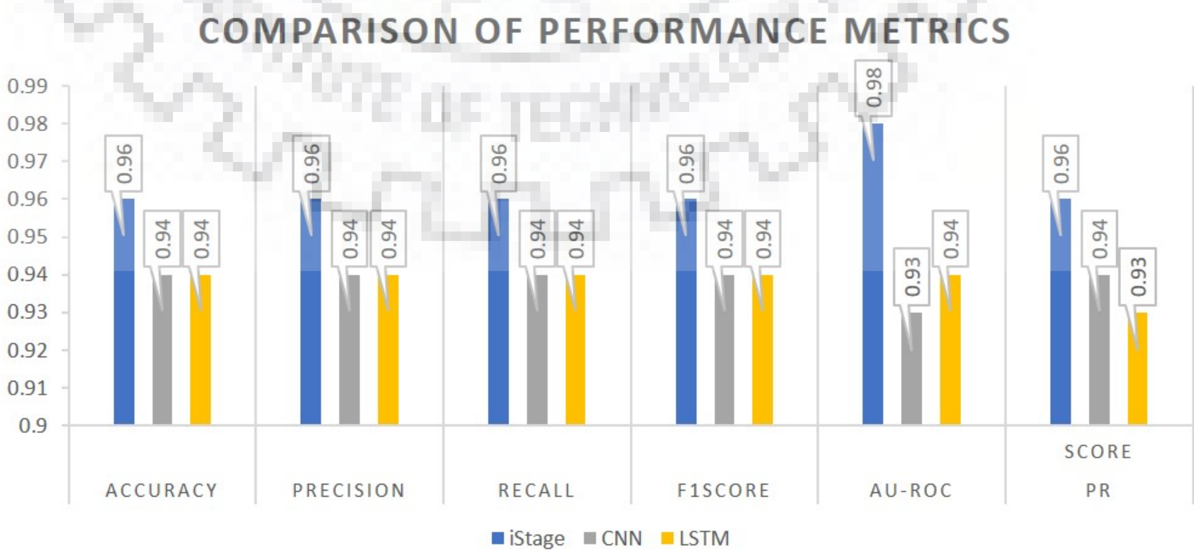


Fig. 6.11: Comparison of Performance Metrics.

Article	Method Used	Accuracy	Precision	Recall	F-score	AU-ROC	AU-PRC
Tweet4Act Chowdhury et al. (2013)	Dictionary-based period classifier; SVM; DT; RF; MaxEnt	-	0.69	0.93	0.71	-	-
iStage	LSTM-CNN	0.9598	0.96	0.96	0.96	0.98	0.96
Improvement (in %age)	-	-	39.13%	3.22%	35.21%	-	-

Tab. 6.6: Comparison of iStage with Existing Literature.

$$Improvement(\%age) = \frac{Proposed\ Max\ Value(\%age) - Other\ techniques\ Value(\%age)}{Other\ techniques\ Value\ (\%age)} \tag{6.1}$$

6.3.7 Decision-Making

The superiority of the iStage is verified by comparing multiple metrics and previous studies; thus, we adopt the trained proposed model to determine the stage of the disaster from an SM message to support decision-making. As mentioned in section: 6.3.4, data is divided into *training, validation, and testing* data. The *testing* data is used to test the proposed model with stage determination. After the training process, the trained model is saved and called when the testing process comes into play. Table: 6.7 depicts the predictions of the cyclone Fani dataset. It provides the tweet number and text and provides the predicted stage label. To be precise, the predicted label of a tweet is *pre, during, post, or irrelevant*. The correct classification of the tweet is colored blue, and the incorrect prediction of the tweet is colored red. This is done to distinguish tweets with accurate and inaccurate predictions visually.

Accurate predictions of *pre-disaster* in Table: 6.7 include tweet no.18796 (row3), 55430 (row 6), where the user mentions *impending disaster*, and *cyclone to make landfall in Odisha today*. Similarly, *during-disaster* tweets with tweet no. 9707 (row 1) states that the *cyclonefani is presently near Kharagpur, Egra, Tamluk.* Moreover, *it has already entered Bengal* signifies the *during-disaster* event. *Post-disaster* tweets with tweet no. 68442 (row 7), indicating that the *cyclone is gone*. *Irrelevant to disaster* tweets with tweet no. 101797 (row 9) contains no information related to cyclone Fani; instead, it mentions political campaign. Tweet no. 82054 (row 2), 16217 (row 5), and 9095 (row 8) present the misclassifications.

Tweet No	Tweet	Manual Annotation	Predicted Stage
9707	#CycloneFani is presently near Kharagpur, Egra, Tamluk. It has already entered #Bengal \n\n#CycloneFaniUpdates\xe2\x80\xa6 https://t.co/nRMHfSktR2’	1	1
82054	@ArmaanMalik22 Praying for all who have been affected by #CycloneFani \xf0\x9f\x98\x94\n And all others be safe	2	3
18796	b’Have cancelled my rallies for the next 48 hours because of what could be an impending disaster #CycloneFani . We ar\xe2\x80\xa6 https://t.co/sRBgt0fTNk’,,,,,	0	0
47019	b’”India set to evacuate a million people as Cyclone Fani on path to hit the eastern coastline” #CycloneFani https://t.co/1nGe3uM1uA’	1	1
16217	#CycloneFani likely to affect the weather pattern in some parts of #Nepal today.’	1	3
55430	b’Cyclone Fani to make a landfall in Odisha today, 800,000 being evacuated\nhttps://t.co/KqDi0BPnZ8\n#Cyclonefani #Odisha #Puri’	0	0
68442	Cyclone is gone. #Odisha#CycloneFani https://t.co/wDeLxyIwHT’	2	2
9095	#CycloneFani likely to affect the weather pattern in some parts of #Nepal today.’	1	2
101797	b”\xf0\x9f\xa7\xa1\n That’s great #RahulGandhi ji. 15 more days left for the campaign. Please work harder!\n\nWe need 400+ seats for\xe2\x80\xa6 https://t.co/Az9y06UcFk”	3	3

Tab. 6.7: Tweets with the predicted stage of the Cyclone Fani disaster. 0 represents *Pre*, 1 illustrates *During*, 2 represents *Post*, and 3 represents *Irrelevant*. The blue color represents the correct predictions, and red color signifies the incorrect predictions.

To support the real-world scenario of DM, we experiment iStage on different disastrous events, i.e., cyclone Amphan, cyclone Titli, and cyclone Nisarga, as explained in section: 6.3.1. Table: 6.8 showcases the prediction of different disastrous events. The table displays the tweet number of the respective disaster dataset and a predicted label. The *pre-disaster* tweets are labeled as *0* with the color blue, *during-disaster* tweets are labeled as *1* with a color green, *post-disaster* tweets with labeled as *2* with color brown, and *irrelevant* tweets are labeled as *3* with red color.

Tweet no. 900 (row 1), 913 (row 2), and 928 (row 3) in Table: 6.8 present tweets of cyclone Nisarga. Tweet no. 900 mentions the cyclone *advancing towards* Indian states, predicting the tweet to be in the *pre-disaster* stage. Tweet no. 913 states that *Mumbai struggles cyclone*, making it a *during-disaster* message. Tweet no. 928 mentions *cyclone making landfall to the south....*, classifying the message into *during-disaster* category.
Tweet no. 17 [(row 4) of Table: 6.8] of cyclone Amphan dataset is classified into an *irrelevant* category as it has no relevance to cyclone Amphan. Tweet no. 221 (row 5), and 8211(row 6)

Tweet No	Tweet	Predicted Stage
900	b' RT @maritime: Cyclone **#Nisarga** being formed in Arabian Sea. Advancing towards coastal Karnataka, Goa, Maharashtra and Gujarat	**0**
913	b' RT @TPE_connect: Mumbai struggles severity of cyclone Nisarga **#CycloneNisarga**	1
928	b' RT @metofficestorms Satellite and radar imagery show the eye of Cyclone Nisarga is making landfall to the south of Mumbai in India **#CycloneNisarga**	**1**
17	@DGFS_HGs_CD: Handed over the letter of Chairman Kolkata Municipal Corporation**#CyloneAmphan**	**3**
221	@FirhardHakim: Bengal continues to battle the twin crises of #Covid19 and **#CyloneAmphan**	**1**
8211	@AjayChairman: As if the #Covid19 destructions are not enough, Locust attack in Western India and **#CycloneAmphan** hits Eastern India	**1**
96278	b'RT @CMO_Odisha: CM @Naveen_Odisha has sanctioned 102 crore as first installment of gratuitous relief for the **#CycloneTitli** affected Gajapa	**2**
96268	b'RT @ncbn: **#CycloneTitli** has left many of the lives in devastation in the northern coastal regions of Andhra Pradesh. Join hands with Govern	**2**

Tab. 6.8: Predictions on Cyclone Amphan, Cyclone Nisarga, Cyclone Titli, and Covid19 datasets. 0 represents *Pre-Disaster messages* in blue color, 1 represents *During-Disaster messages* in green color, 2 represents *Post-Disaster messages* in brown color, and 3 represents *Irrelevant* messages in red color.

are predicted as *during-disaster* tweets. Tweet no. 96278 (row 7) and 96268 (row 8) are from cyclone Titli datasets, classified into a *post-disaster* category. The predicted labels signify that the framework works very well on sundry disasters. Regarding the neural networks, iStage performs better than LSTM and CNN, making it more amicable to complex structures.

To investigate the reusability of iStage for disaster events, other than cyclones, we determine the stage of disaster on the Covid-19 dataset. Table: 6.9 represents the predicted stage label of the SM message of Covid-19. Tweet no. 1 (row 1) of Table: 6.9 is from the Covid-19 dataset. The message is accurately classified into the *during-disaster* category. Similarly, tweet no. 44 (row 2, Table: 6.9) is correctly classified into a *during-disaster* category, mentioning the rising cases of the disaster. Tweet no. 13584 (last row, Table: 6.9) is aptly predicted as an *irrelevant* category. The results demonstrate that iStage works well on non-cyclonic events as well.

Tweet No	Tweet	Predicted Period
1	b'RT @PunjabGovtIndia: To further augment its **#COVID** management and care strategy, CM @capt_amarinder Singh led #PunjabGovernment has constituted committee	1
44	b"#COVID19: #Ahmedabad's **#COVID-19** cases rise by 187 to 22,262, five more deaths take toll to 1,496: Health department https://t.co/0mYXRaVgsR"	1
13584	b'Where is the official demarcation of metro Melbourne???? **#COVID-19'**	3

Tab. 6.9: Predictions on Covid19 dataset. 0 represents *Pre-Disaster messages* in blue color, 1 represents *During-Disaster messages* in green color, 2 represents *Post-Disaster messages* in brown color, and 3 represents *Irrelevant* messages in red color.

6.4 Conclusion

The focus of this study is to determine the stage of the disaster from an SM message or tweet. The objective of this study is to determine whether the tweet is *pre-disaster, during-disaster, post-disaster, or irrelevant*. In this study, we propose a DL-based framework for the aforesaid purpose. We sequentially complete the said task by (1) gathering cyclonic and pandemic data from Twitter, (2) manually annotating the tweet message as *pre, during, post, or irrelevant*, (3) pre-processing the raw and unstructured data, (4) development of hybrid DL model, (5) evaluation using performance metrics, (6) comparison against state-of-the-art and existing research studies, (7) predictions on different datasets.

In the empirical section, we conduct extensive experiments against different neural network models and validate the effectiveness of iStage. To demonstrate the technical potency of iStage, we determine the stage of cyclonic disasters like cyclone Titli, cyclone Amphan, and cyclone Nisarga. In addition, to scrutinize the proposed model on other catastrophic events, we examine iStage for disastrous Covid-19 event, and we see that iStage is effective.

One advantage of iStage is that the results are visualizable enough to get valuable insights for making decisions at an appropriate time in practical applications. Furthermore, iStage is flexible enough to be scaled to other disastrous events. Finally, the stage determination of the disaster may provide insights that are valuable for effective DM. In addition, the promising results demonstrated in the study hold the potential for novel applications in DM and can open new opportunities in the study of SM usage for DM in real-life settings.

7. A NOVEL WEB-BASED SMART DISASTER MANAGEMENT SYSTEM FOR DETERMINING THE NATURE OF A SOCIAL MEDIA MESSAGE FOR DECISION-MAKING USING DEEP LEARNING - CASE STUDY OF COVID-19

7.1 Introduction

The automatic identification of information in DM is one of the most compelling and necessitu-ous problems across the globe as countries across the world have witnessed a significant increase in the intensity of disasters. With the spiraling coronavirus disease, i.e., Covid-19, it has become indispensable to extract and disseminate accurate as well as timely information. The World Health Organisation (WHO) declared the coronavirus disease outbreak as pandemic on March 11, 2020.

The situation worsened day by day, and therefore, speedy and on-time information retrieval is imperative. Social distancing, lockdowns, travel bans, self-quarantines, and business closures have forced people to glue to SM more than ever before (Zhang et al., 2020a). SM's real-time data production capability makes data enormous and diverse. However, only a small fraction of content is meaningful and relevant. SM data can serve as a valuable channel for seeking help, offering assistance, situational awareness, general opinions, and coordinating activities in disaster. At a system level, understanding the needs, availabilities, situational updates, public opinions enhances the planning and rescue operations, improving disaster resilience.

Timely determination in DM is vital, but it still is challenging. DL algorithms are an important part of DM systems using SM data (Caragea et al., 2016a; Huang et al., 2020a; Nguyen et al., 2019). Regarding the application of SM data for DM, research suggests that lack of tools for managing SM data during a disaster makes it difficult for disaster professionals to not understand how SM data can be helpful for the public. To fill this gap in the literature, this paper explores disaster data using DL techniques to determine the nature of an SM message.In addition, the existing KM systems lack the required features. Hence, the objective is born out of the gap by developing DisDSS. The chapter is important for dissertation as it fulfils the criteria of DIKW model and presents the knowledge derived out of raw and unstructured SM data. Our key original contributions are:

1. Developing a new SM-based Covid-19 dataset with the label of nature of message from April 22, 2021, to May 05, 2021, with 1,03,839 tweets in total.

2. Propose a new fusion model for determining the nature of a disaster-related SM message by integrating the structure of CNN and BiLSTM together.

3. The proposed fusion model is benchmarked against other state-of-the-art models and experimental results of previous research studies.

4. Demonstrate the technical efficacy by determining the nature of SM message on the Covid-19 disaster dataset.

5. Propose a web-based interface to visualize the structured information, where policy-makers, decision-takers can utilize this information for efficient management of catastrophic events.

The rest of the chapter is organized as follows. The methodology is explicated in section: 7.2. Section: 7.3 briefs about the evaluation metrics. Section: 7.4 is about the experimental results and analysis followed on a covid-19 dataset. Section: 7.5 is about the practical application of the proposed model. Section: 7.6 details the user experience of DisDSS. Lastly, we conclude the chapter in section: 7.7.

7.2 Materials and Methods

This section proposes DisDSS: a smart web-based DM system that can be explored to determine the nature of a SM message from the Twitter dataset. DisDSS is a web-based DM system, a visual application designed to support needs and availability-based information, situational information, and general views regarding the Covid-19 disaster. The general architecture of DisDSS is illustrated in Figure: 7.1. The following subsections explain the methodology in detail.

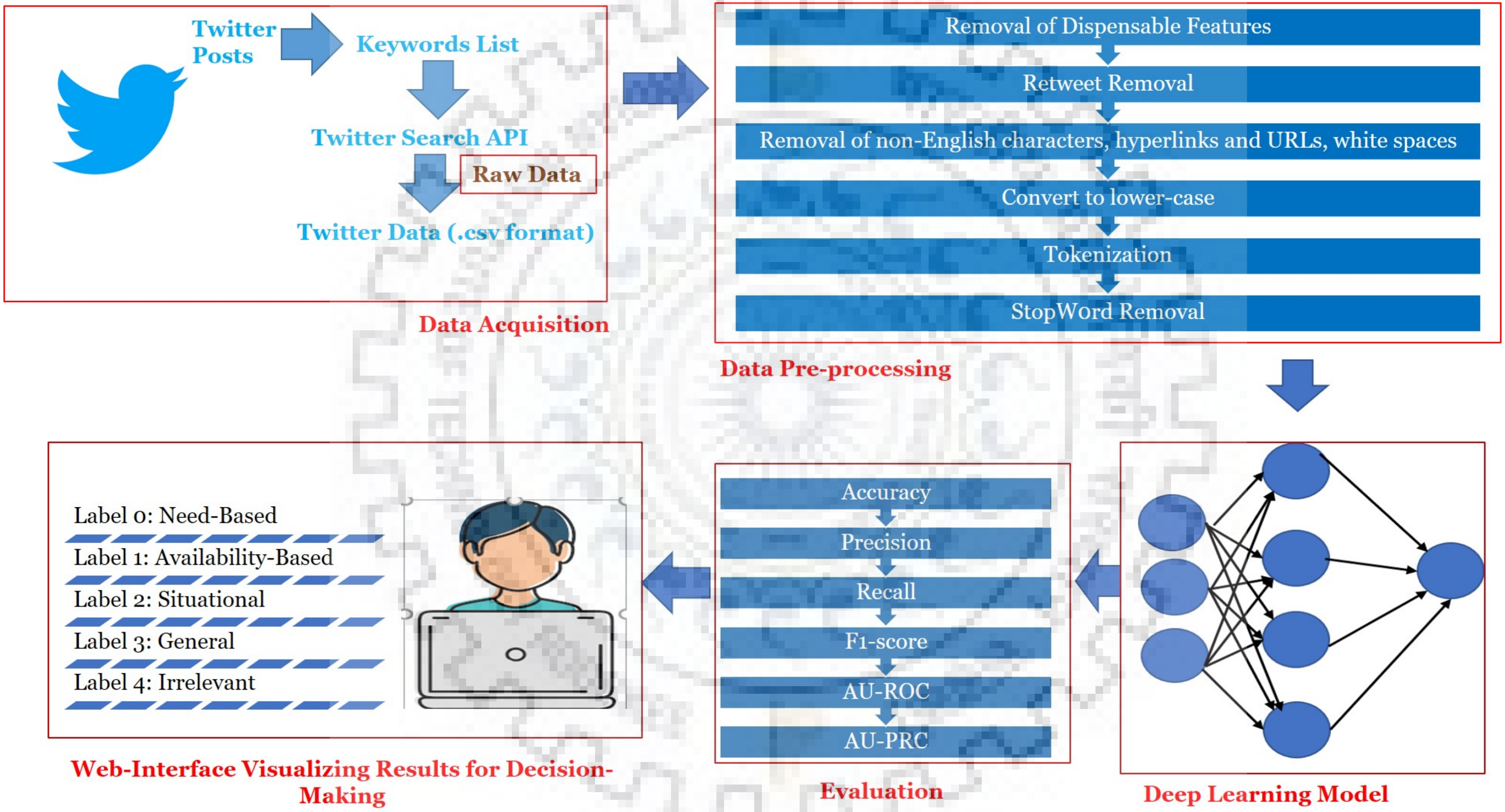


Fig. 7.1: The structure of DisDSS: a smart disaster management system for nature of social media message determination.

7.2.1 Data Acquisition

The proposed DisDSS considers the Twitter platform to extract data for SM message determination. The data is fetched using the Twitter search application programming interface via Python language with keyword #Covid. The data contains numerous features like created_at, text, lang, id, user object. Further, the user object collected features are screen_name, userid, id_str, name, location, url, description, verified, followers_count, friends_count, listed_count, favorites_count, statuses_count, created_at, profile_image_url. We gathered tweets from April 22, 2021, to May 05, 2021, with 1,43,340 tweets in total.

7.2.2 Manual Annotation

In order to annotate the data, we de-duplicate the dataset. Only unique messages are considered for annotation, i.e., 1,03,839 messages. We label the dataset into five categories: need-based, availability-based, situational-based, general, and irrelevant. The manually annotated dataset acts as first contribution mentioned in section 7.1. We show the annotation scheme with the class label below:

0. *Need-Based SM messages*: messages those contain information about the need or unattainable resources, such as ICU beds, oxygen cylinders, vaccination, food, water, electricity, etc.

1. *Availability-Based SM messages*: messages that inform about the availability of certain resources. This class comprehends the *potentially available* as well as *actually-available* resources.

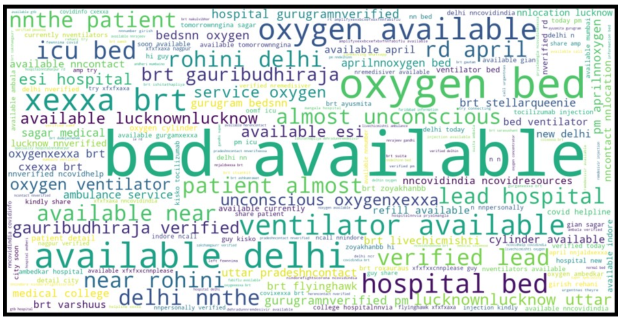
2. *Situational SM Messages*: tweets that provide information related to the number of cases, deaths, injuries, etc.

3. *General SM Messages*: the class embodies tweets mentioning prayers, cautions, advice, emotional support, criticizing, etc.

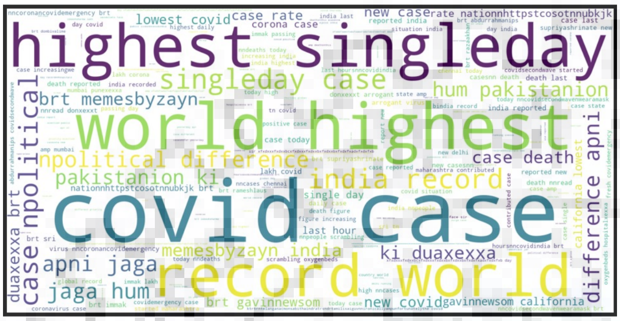
4. *Irrelevant SM Messages*: tweets that contain no information related to the disaster.

The word cloud is presented for each human-annotated class in Figure: 7.2, and the message distribution in Figure: 7.3.

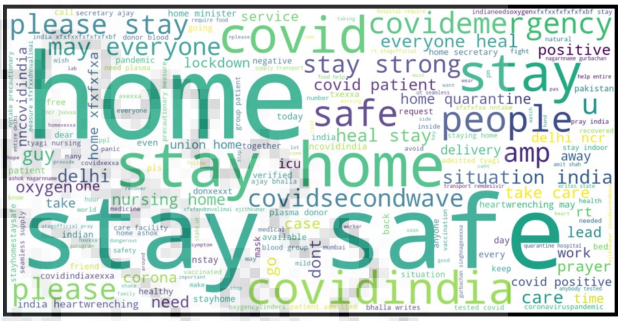
(a) Need-based



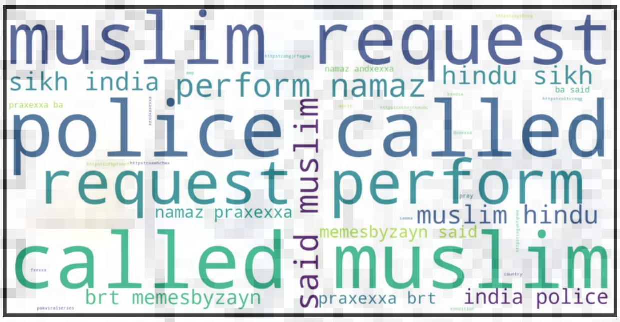
(b) Available-based



(c) Situational-based



(d) General



(e) Irrelevant

Fig. 7.2: Word Clouds of Human Annotated Classes.

Fig. 7.3: Covid dataset message distribution after manual annotation.

7.2.3 Data Pre-processing

The collected dataset is raw, noisy, unstructured, and informal. The next step in the development of DisDSS is to pre-process the data. To do so, the raw data is preprocessed by removing retweets eliminating dispensable features. All the features are banished except tweet no, text, and manual label. The hyperlinks and URLs are removed. Non - English characters along with additional white spaces are removed. The tweets are folded into lower case and tokenized into smaller units called tokens. Further, English language stop words are removed.

7.2.4 Proposed Model

We randomly split the dataset into three subsets: *training, validation, and testing* in the ratio of 60:20:20, which is a commonly used split ratio (Alam et al., 2018; Tam et al., 2021). The *training* data is used to train the model, and the *validation* dataset is used to tune the hyperparameters and provide an unbiased evaluation to select the best model. The *testing* dataset is used for testing and prediction purposes. Figure: 7.4 depicts the distribution of the covid dataset and workflow of *training*, *validation*, and *testing datasets*. It is worth noting that we save the trained model from the validation stage and call the saved model for the testing stage.

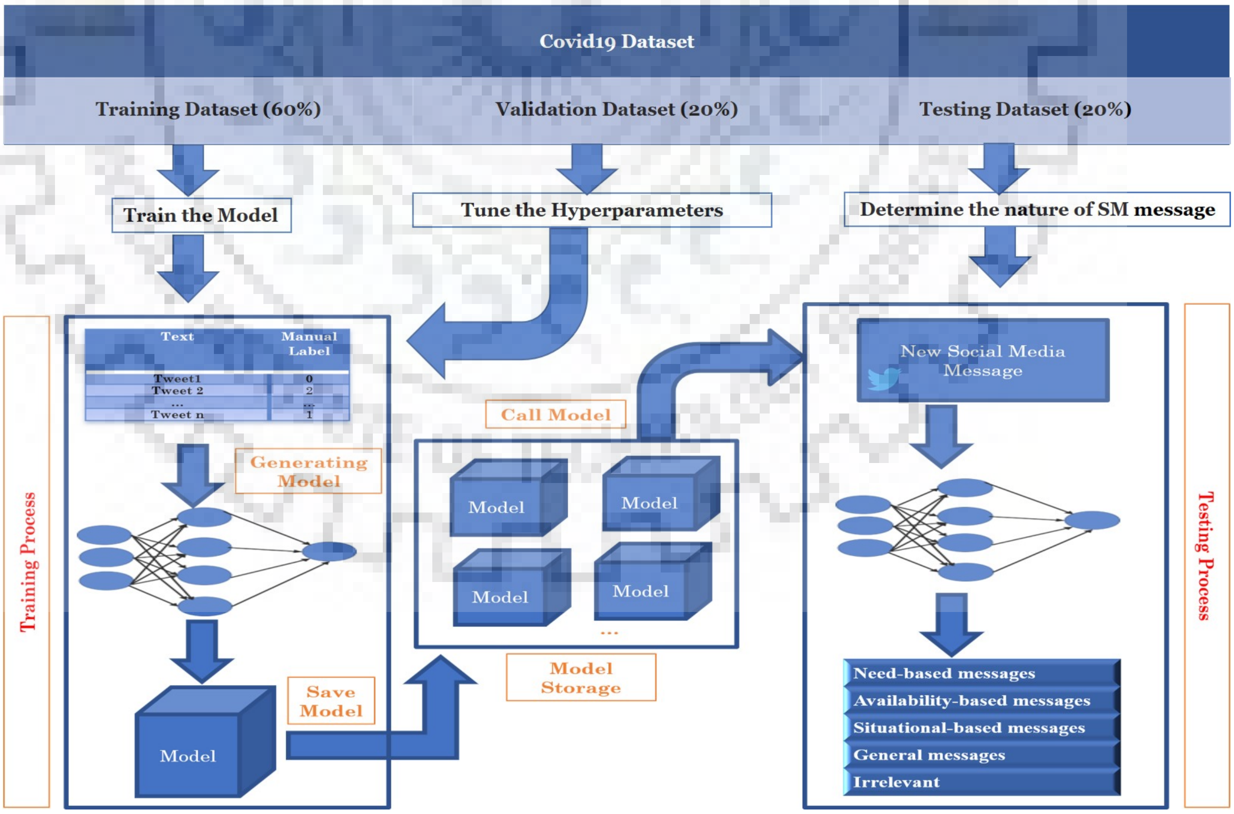


Fig. 7.4: Trained model is saved and called upon for testing process, which is further used for determination of the SM message as *need-based message*, *availability-based message*, *situational-based message*, *general message*, or *irrelevant*.

In this chapter, we develop a DL-based hybrid model composed of multiple layers of the cascade. Section 3.6 describes the literature of hybrid model usage and its superior perfomance than existing models. Figure: 7.5 elucidates the proposed model, encapsulating the layers. The model consists of a CNN layer, which receives the word embeddings for each token in the tweet as inputs. Subsequently, the convolution layer's output is pooled to a smaller dimension. The outcome of the pooling layer is fed into the BiLSTM layer, where it extracts the context information of words. Dropout layers are used to overcome the overfitting issues. The dense layer ultimately outputs the tweet as *need-based, availability-based, situational-based, general-based*, or *irrelevant* messages using a softmax function. The pseudo-code for the proposed model is summarized in Algorithm 3.

Integrating CNN and BiLSTM layers, not only reaps the benefits of CNN extracting local features, but also considers the advantages of BiLSTM in contextual information of text sequences. BiLSTM is used as it uses both forward LSTM and backward LSTM. LSTM considers only the past information, ignoring the future information. Moreover, not all the parts of the SM message are equally relevant, but LSTM cannot recognize the different relevance between each part of the message. Our proposed model is relatively complex using CNN and BiLSTM layers. From a linguistic point of view, BiLSTM considers the meaning of the words in context and overcomes the drawback that LSTM does not consider the information afterwords (Liu et al., 2020). The parameter settings for the proposed model are shown in Table: 7.1.
The working environment of the proposed approach is as follows:

- *Hardware Configuration*: Intel(R) Core(TM) i5-8250 8th Generation, 8GB RAM, 512 GB SSD, and NVIDIA GeForce MX150.
- *Software Configuration*: Microsoft Windows 10, 64-Bit, Python 3.8.3.

Python, a high-level, general-purpose programming language, is used to interact with DL libraries as application program interfaces (APIs). The experiments are carried out using open source libraries such as NumPy, Pandas, TensorFlow, and Keras (DL Framework).

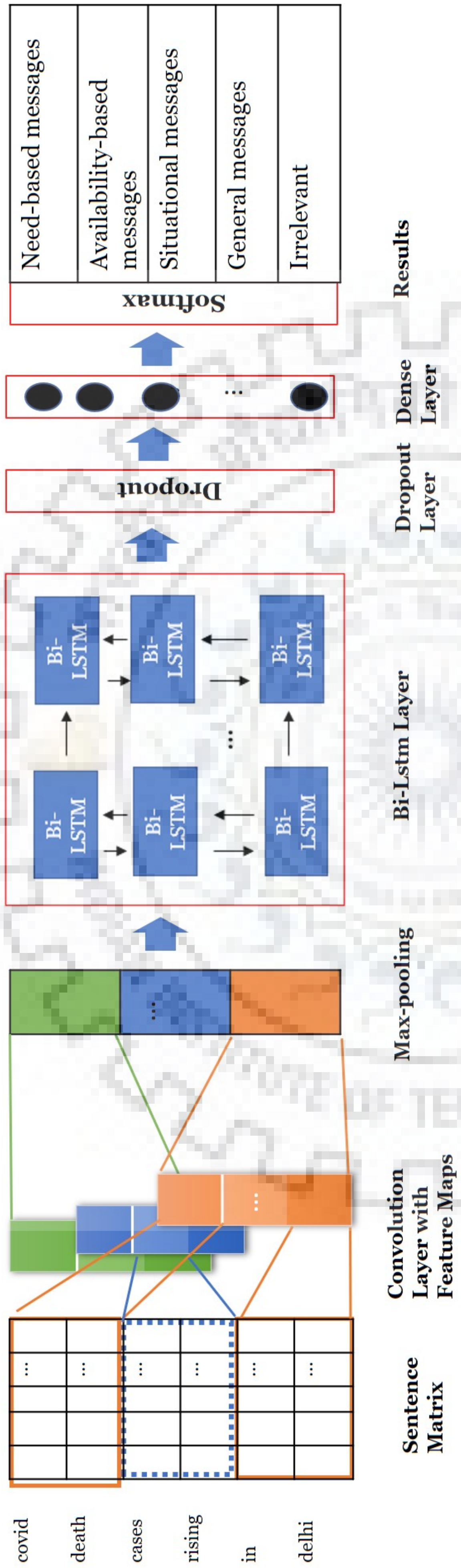


Fig. 7.5: Proposed model encapsulating the layers.

Algorithm 3: Pseudo-code for Proposed Model

Input: Training dataset, Validation dataset, Manual Label of Messages

Output: Label Prediction of SM message.

Begin
For number of training iterations **do**
Construct sentence matrix using embedding layer to convert each token into an integer
Apply ReLU activation function *f(x)*= max(0,x) to convolution layer
Reduce the dimensionality by maxpooling layer
Employ BiLSTM to obtain the preceding contextual features and the succeeding contextual features
Add dropout to prevent overfitting
Feed the comprehensive context representations into the softmax classifier to get the class labels
Update parameters of the model using the loss function with the Adam method
The performance metrics (Accuracy, P, R, F-score, AU-ROC, and AU-PR-Curve) are calculated using Equations in Table: 7.2
End

Hyperparameter	**Definition**	**Value**
Epochs	Iteration count	25
Embedding Dimension	Size of vector used to represent each of the word embedding	300
Maximum Sequence length	Maximum tweet-length	23
Activation Function	Calculates the weighted sum of its input, adds a bias and then decides whether to activate the neuron or not	ReLU
Dropout Rate	Some hidden layer neurons are discarded with the 40% probability during the training process to reduce the dependence on some local features in each iteration.	0.4
Optimizer	The method used to update the weights, in order to reduce the error	Adam(1e-3)
Batch Size	Data are grouped into batches prior to feeding it into DL model	1024
Loss Function	Function to assess model prediction	Categorical CrossEntropy

Tab. 7.1: The key parameters (hyperparameters) for the proposed DL model.

7.3 Evaluation

In order to evaluate the performance of our model, the focus is not only on one performance metric; instead, multiple metrics are considered. Accuracy, P, R, F1-score are the performance metrics. Furthermore, Area under ROC and Area under PR-Curve are deployed to measure the completeness and robustness of the framework. The performance metrics used in this research study are defined in Table: 7.2.

Even though there is no universally accepted cutoff point for aforementioned performance measures like accuracy, precision, recall, and F-score, it is widely acknowledged that the higher the value of metrics, the higher the performance. The best and worst values of accuracy, P, R, F-score, AU-ROC, and AU-PR-Curve, are 1, 0 (Luo and Xu, 2021).

Name	**Description**
Accuracy =(TP+TN)/(TP+TN+FP+FN)	%age of correct predictions by the classifier.
P=TP/(TP+FP)	A measure of exactness(%age of predicted relevant tweets that are actually relevant)
R=TP/(TP+FN)=TPR	A measure of completeness(%age of relevant tweets labeled as such)
F-measure =$(a^2+1)P*R/(a^2(P+R)$	Highly informative measure, considering harmonic mean of P and R.
AU-ROC	The area under ROC is a metric to measure accuracy through an area under the ROC.
AU-PRC	The area under PR-Curve is a measure based on PR-Curve, i.e., a plot of P(y-axis) and the R(x-axis).

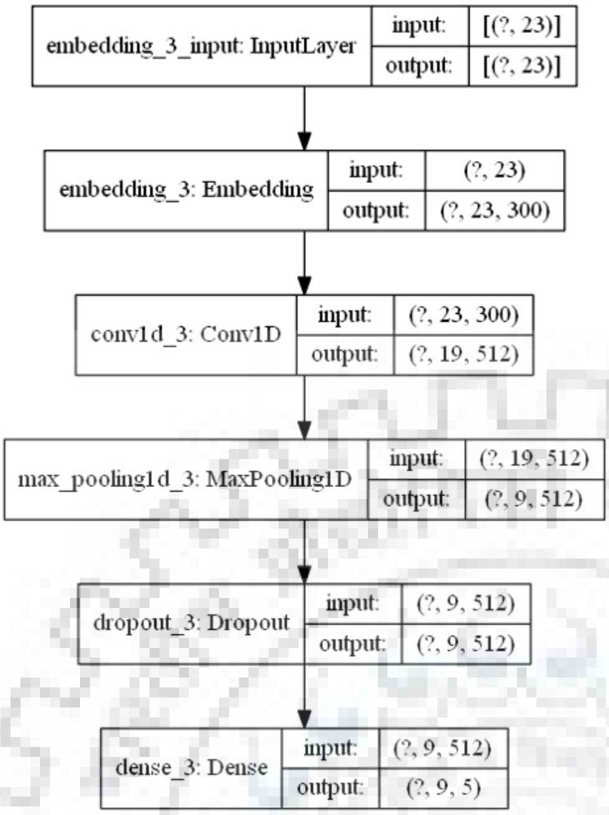
Tab. 7.2: Measures of performance for the proposed study.

7.4 Results

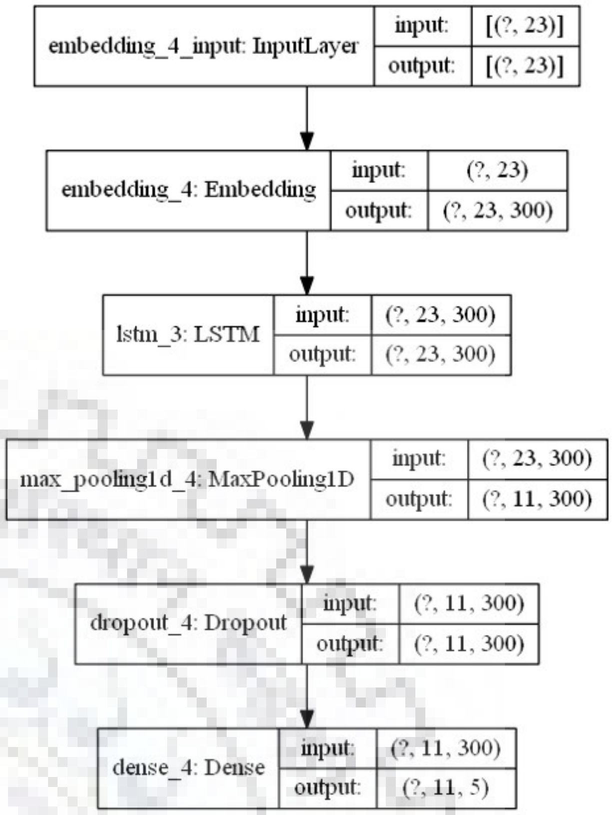
The subsections that follow explore the results of the conducted experiments, considering the third contribution mentioned in section 7.1. First, section: 7.4.1 compares the results of the proposed approach with the state-of-the-art. Next, section: 7.4.2 compares the results of our proposed model with the existing research studies.

7.4.1 Comparison with State-of-the-Art

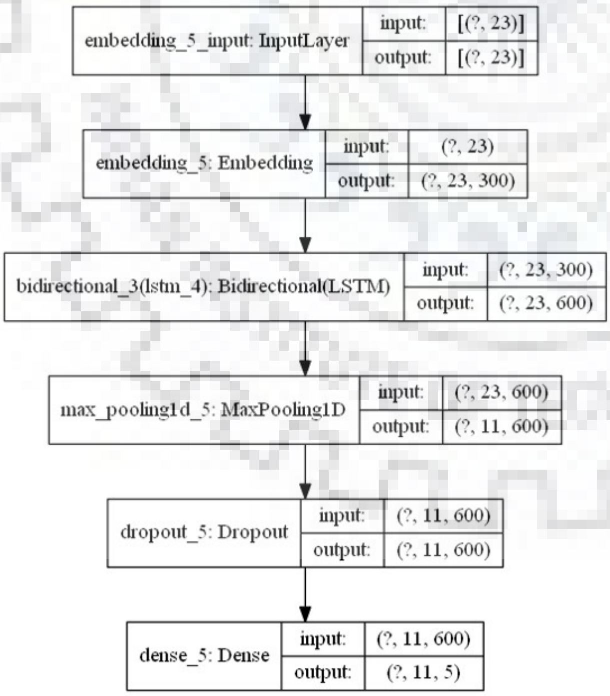
In this subsection, we train DL models - CNN, LSTM, and BiLSTM. The comprehensive descriptions of CNN, LSTM, and BiLSTM are narrated in Kim (2014), Hochreiter (1997), and Goodfellow et al. (2016), respectively. Correspondingly, their general architectures are displayed in Figure: 7.6a, 7.6b, and 7.6c. The architecture of the proposed model is showcased in 7.6d. The architectures are generated through Anaconda Integrated Development Environment. To further evaluate the performance, we compare our proposed approach with CNN, LSTM, and BiLSTM. Table: 7.3 shows the performance results of the conducted experiments in tabular format, and Figure: 7.7 depicts the results pictorially.



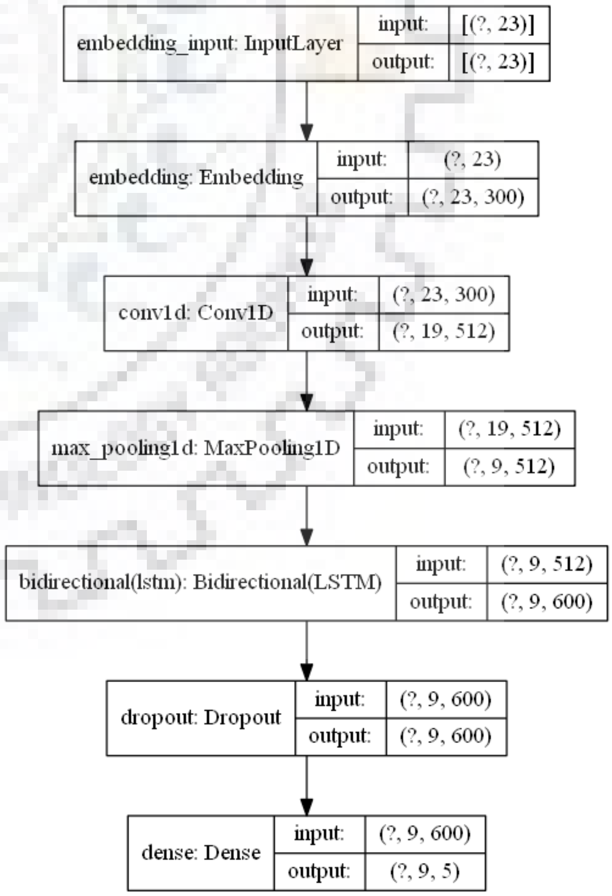
(a) CNN Architecture



(b) LSTM Architecture



(c) Bi-LSTM Architecture



(d) Proposed Model

Fig. 7.6: Architectures of four DL models.

Approach	Accuracy	Precision	Recall	F1-score	AUC	PR Score
Proposed Model	0.91	0.90	0.90	0.90	0.97	0.91
CNN	0.87	0.87	0.86	0.86	0.96	0.89
LSTM	0.87	0.88	0.88	0.88	0.96	0.89
Bi-LSTM	0.90	0.90	0.89	0.89	0.96	0.89

Tab. 7.3: Comparison results for the proposed model against other classifiers.

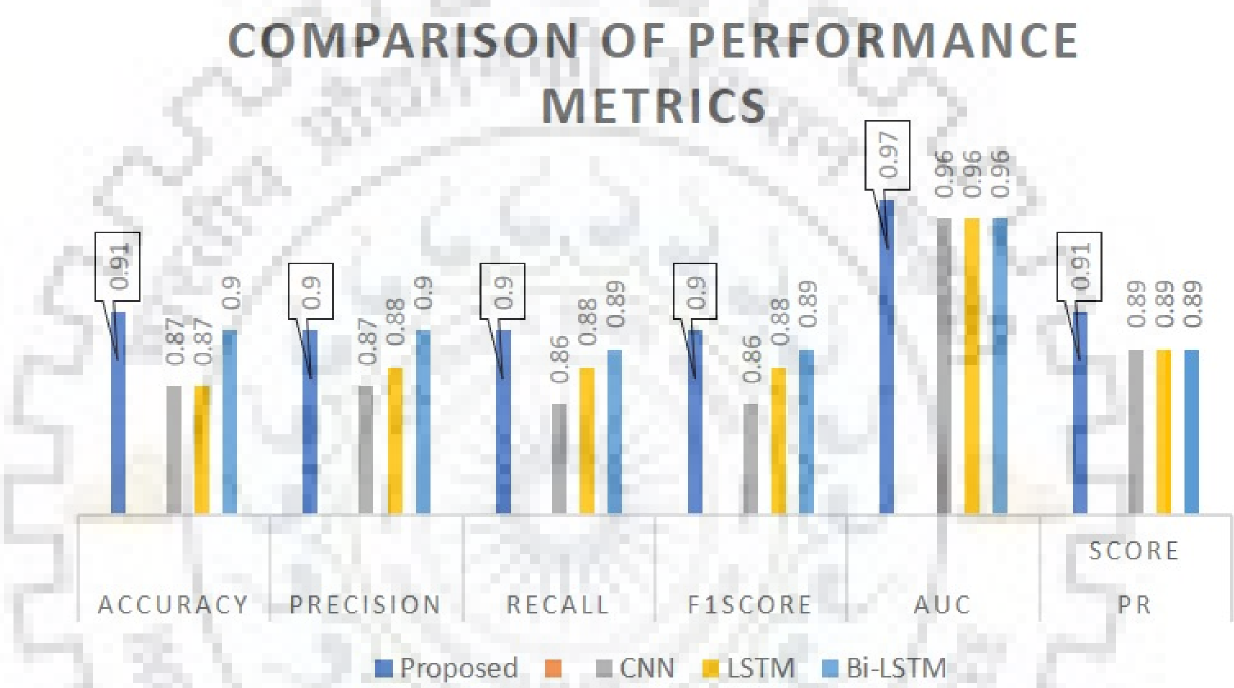
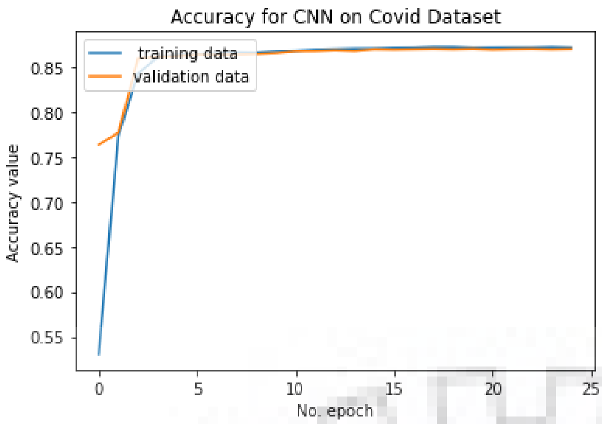


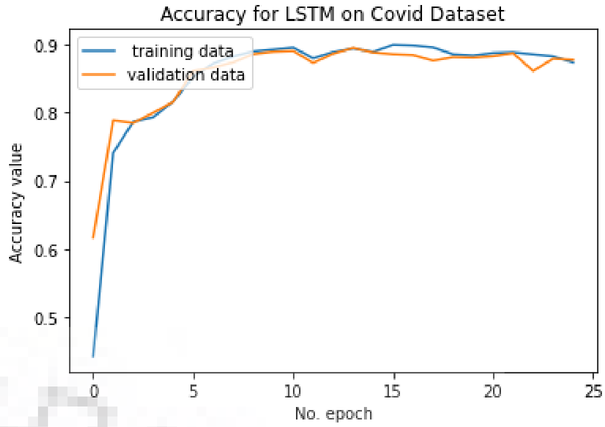
Fig. 7.7: Comparison of Performance Metrics.

Accuracy vs. epochs plots is illustrated in Figure: 7.8. The X-axis represents the number of epochs, and the y-aixs represent the learning curve with the scaling (0,1). We can observe that the accuracy of proposed model as 0.91, i.e., 91% in Figure: 7.8d, which is higher than 0.87 of CNN [Figure: 7.8a], 0.87 of LSTM [Figure: 7.8b], and 0.90 of BiLSTM [Figure: 7.8c]. In the proposed model, it reaches an accuracy higher than 90% in both training and validation sets. The training accuracy curve and validation accuracy curve in Figure: 7.8d levels and stabilizes, indicating no significant overfitting problem.

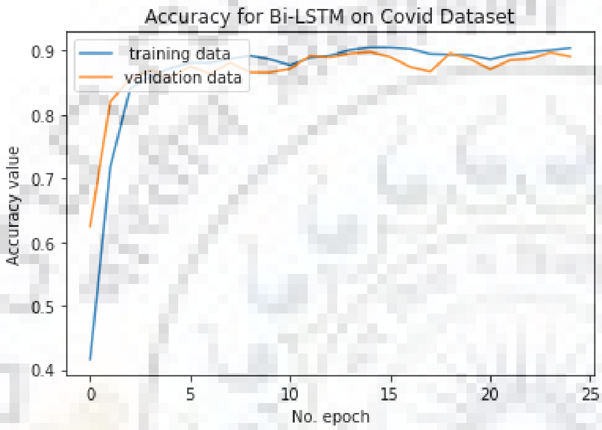
As can be seen from column III, Table: 7.3, the P of the proposed model is 0.90, CNN has 0.87, LSTM has 0.88, and BiLSTM has 0.90. A higher value of R of proposed approach (0.90) compared to other models (0.86 of CNN, 0.88 of LSTM, and 0.89 of BiLSTM) is showcased in column IV, Table: 7.3. In column V, Table: 7.3, the F-score is displayed. The F-score value of proposed approach(0.90) is higher than all the models(0.86 for CNN, 0.88 for LSTM, and 0.89 for BiLSTM).



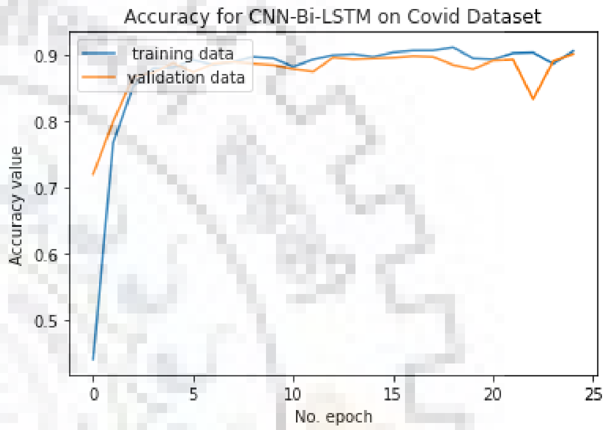
(a) CNN



(b) LSTM



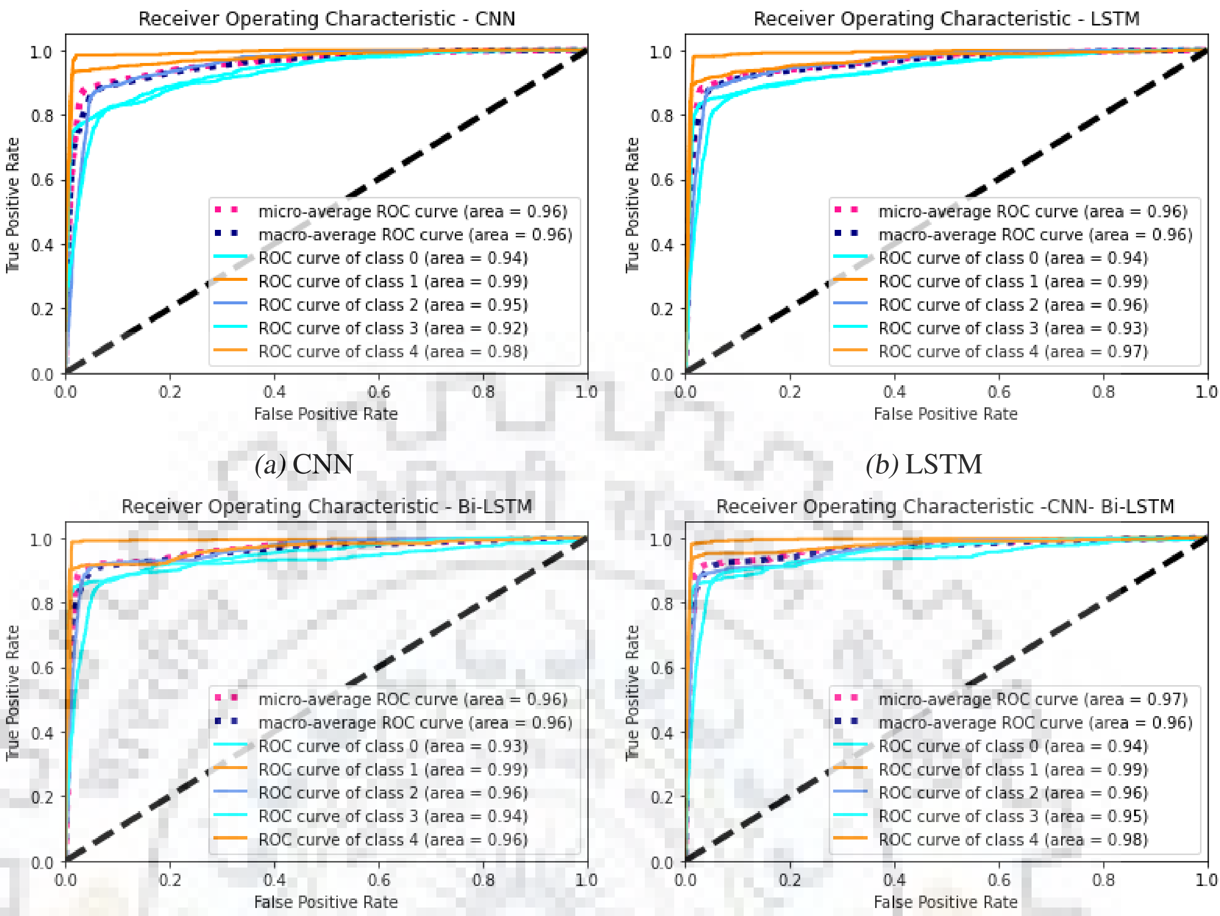
(c) Bi-LSTM



(d) Proposed Model

Fig. 7.8: Accuracy vs. Epochs plots.

We calculate ROC and micro-averaged the performance. ROC curve using CNN, LSTM, and BiLSTM architecture is shown in Figure: 7.9a, 7.9b, and 7.9c, respectively, considering TPR and FPR. Likewise, the ROC curve for our proposed model, implemented on a covid dataset, is illustrated in Figure: 7.9d. Our proposed model performs best amongst all the models. The area under ROC metric of the proposed model is 0.97, which is larger than that of the others (0.96, 0.96, 0.96).



(a) CNN

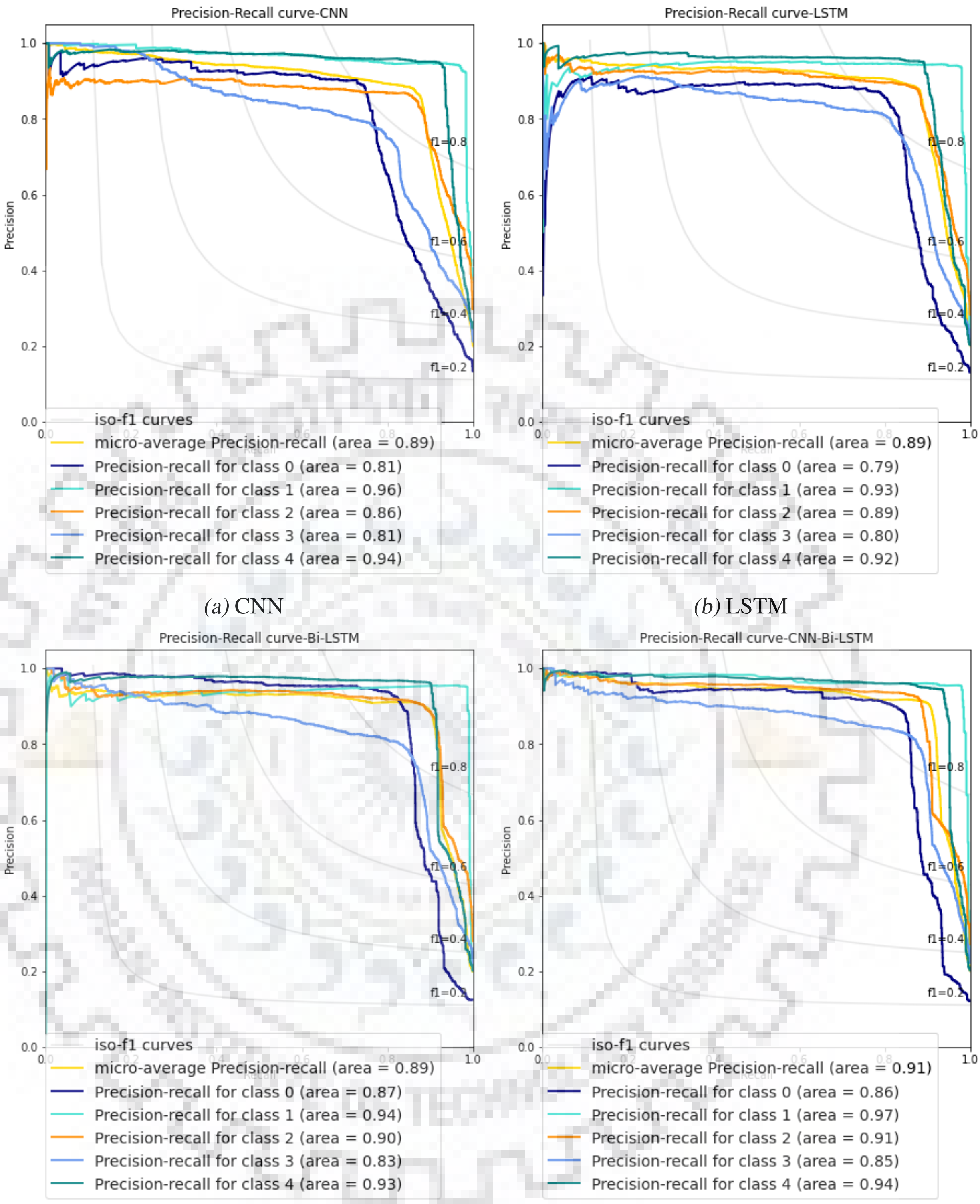
(b) LSTM

(c) Bi-LSTM

(d) Proposed Model

Fig. 7.9: ROC plots of DL models.

Figure: 7.10 plots the PR-Curve of all DL models. The AU-PR-Curve of the proposed model is 0.91 [Figure: 7.10d], which is higher than CNN [Figure: 7.10a], LSTM [Figure: 7.10b], and BiLSTM [Figure: 7.10c]. We can interpret that the AU-PR-Curve of proposed model is 2.24% higher than all the other models. Hence, considering all the evaluation metrics, we can observe that the proposed approach is better than CNN, LSTM, and BiLSTM.



(a) CNN *(b)* LSTM

(c) BiLSTM *(d)* Proposed Model

Fig. 7.10: PR-Curve of all DL models.

7.4.2 Comparison with Existing Studies

The performance of DisDSS is compared with the existing research studies. Similar to other previous studies, our study is based on DL methods. However, it differs due to its methodological contribution. Studies suggesting different methods previously made using DL methods are reviewed in Table: 3.5 in terms of accuracy, P, R, F-score. Accordingly, it is seen that the proposed

method is comparable with previous studies in terms of accuracy, P, R, and F-score.
As can be seen from Table: 3.5, many studies, including CNN, LSTM, and BiLSTM models, have been carried out so far. The most significant advantage of the proposed architecture is the hybrid architecture utilizing the best features of CNN and BiLSTM layers. However, the proposed approach in this study is different from previous ones.
To further evaluate the performance of DisDSS, we compare the proposed work in terms of accuracy, P, R, F1-score, AU-ROC, and AU-PR-Curve with state-of-the-art methods that are applied on SM-based covid datasets, as presented in Table: 7.4.
As shown in Table: 7.4, many studies, including CNN and LSTM. The results demonstrate that the DisDSS significantly outperforms the existing research studies, considering all the evaluation metrics.

[redacted]	**Method**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**AU-ROC**	**AU-PR-Curve**
Naseem et al. (2021)	CNN; BiLSTM; SVM; DT; NB; and RF	0.869	-	-	-	-	-
Behl et al. (2021)	MLP; CNN; and LR	0.83	-	-	-	-	-
Chakraborty et al. (2020)	Multinomial NB; SVM; LR; RF; Adaboost; SVM	0.81	-	-	-	-	-
Jelodar et al. (2020)	LSTM; LDA	0.81	-	-	-	-	-
Abdelminaam et al. (2021)	Modified-LSTM; and Modified GRU	0.83	0.84	0.83	0.83	-	-
Proposed Approach	CNN; LSTM; Bi-LSTM; CNN- Bi-LSTM	**0.91**	**0.90**	**0.90**	**0.90**	**0.97**	**0.91**

Tab. 7.4: Comparison with existing studies.

7.5 DisDSS System Architecture

This section contributes in developing DisDSS: the web interface for visualization of experimental results. DisDSS obtains the resultant data from the proposed model, stored in csv format. The main aim of the DisDSS is to make the system easily accessible and comprehensive for use by disaster professionals. The DisDSS is integrated by merging two modules: (i) Front-end: to build the basic layout of the system, and (ii) Back-end: to store the disaster-related database and enable a graphical user interface. Detailed descriptions of these two modules are provided in the following subsections.

7.5.1 Front-end: creating the layout of the web platform

We recognize that the scientists and disaster professionals with knowledge in DM but limited computer programming skills are potential users of this interface. We use Hyper-Text Markup Language (HTML) for these users, which is compatible with most recent web browsers on computers, laptops, etc. HTML is a standard mark-up language used to describe the structure and provide a design to display various documents in the web browser. With the recent merging technologies like Cascading Style Sheet (CSS), the visualization of web browsers has augmented substantially. The embedding of HTML, images, various forms, etc., is done into a web browser considering the user's needs. The most updated HTML 5 is used to produce a lightweight, user-friendly interface. The URL for the browser is http://127.0.0.1:5000.

7.5.2 Back-end: setting up data storage and enabling graphical user interface

The back-end takes care of the storage of the Covid-19 database and, in addition, controls the graphical user interface. This is brought about by *comma-separated values* and Python. CSV is a widespread data format that is increasingly used nowadays. All core model functions are written in Python language. Python offers several advantages over other languages as it is open-source, accessible, and lightweight. Many leading organizations use Python for web development these days. Flask framework, a lightweight python framework, is used for web development.

More specifically, the tool consists of four main pages:

1. <<Home Page>>: The first page, which runs, is a homepage. The page is to select the concerned country's data. The user chooses the country using the drop-down list [Figure: 7.11].

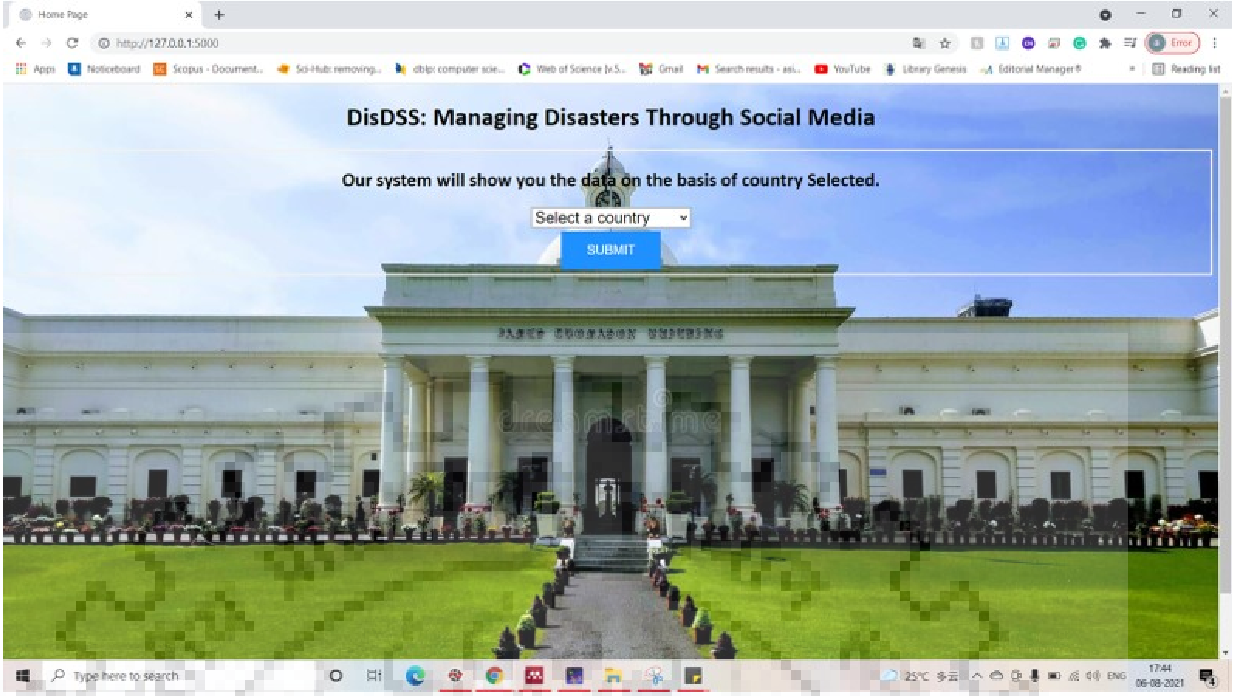


Fig. 7.11: Graphical User Interface for Country Selection.

2. <<State and Metropolitan City Selection>>: the States of India are drop-down upon India's selection as a country [Figure: 7.12].

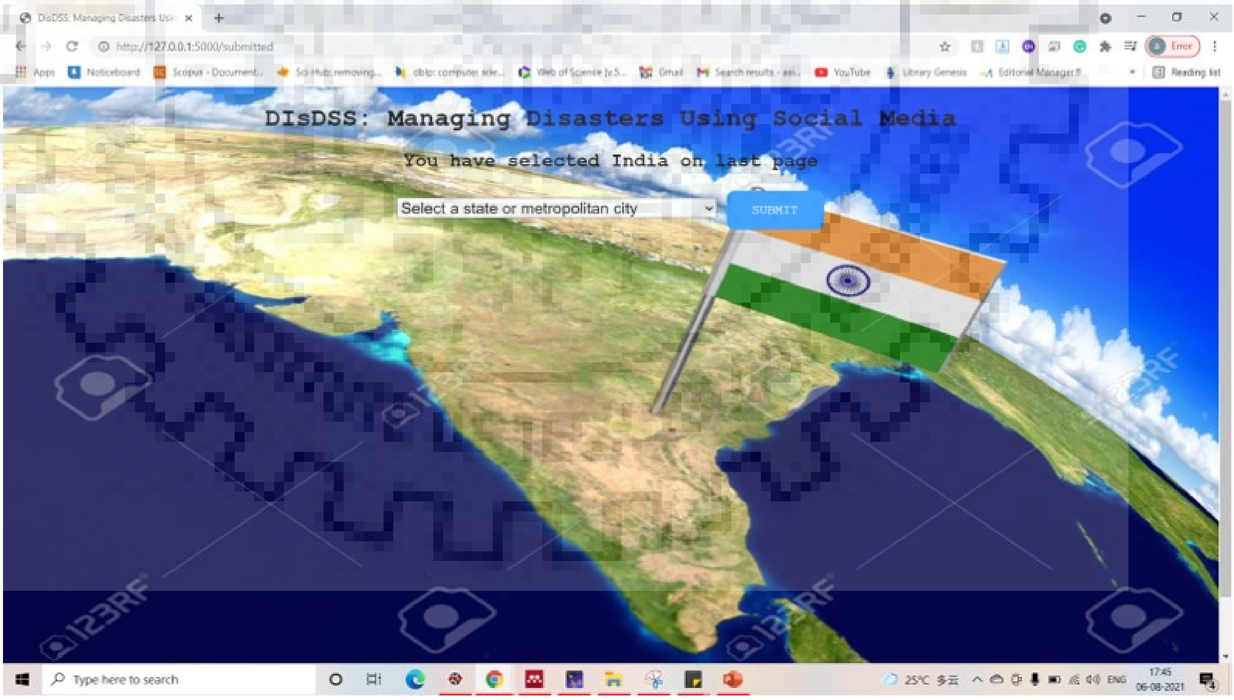


Fig. 7.12: Graphical User Interface for State/ Metropolitan Selection.

3. << Type of Data Selection>>: The user can select the type of data he wants to see, i.e., *need-based, availability-based, situational-based, general, or irrelevant*. To visualize the results, the user can directly click on the respective button[Figure: 7.13].

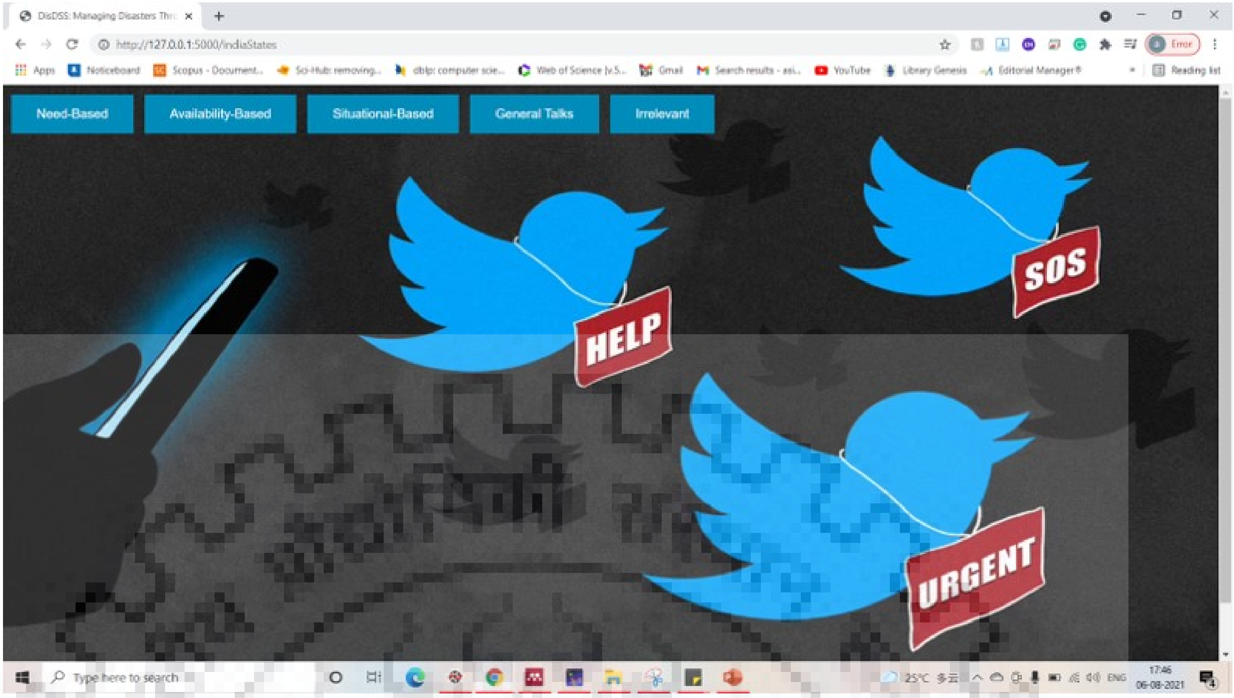


Fig. 7.13: Graphical User Interface showcasing *need-based*, *availability-based*, *situational-based*, *general* or *irrelevant messages*.

4. <<Data Visualisation>>: DisDSS showcases the results in the tabular format. The table provides the tweet no., tweet text, name of the tweeting user, and location [Figure: 7.14].

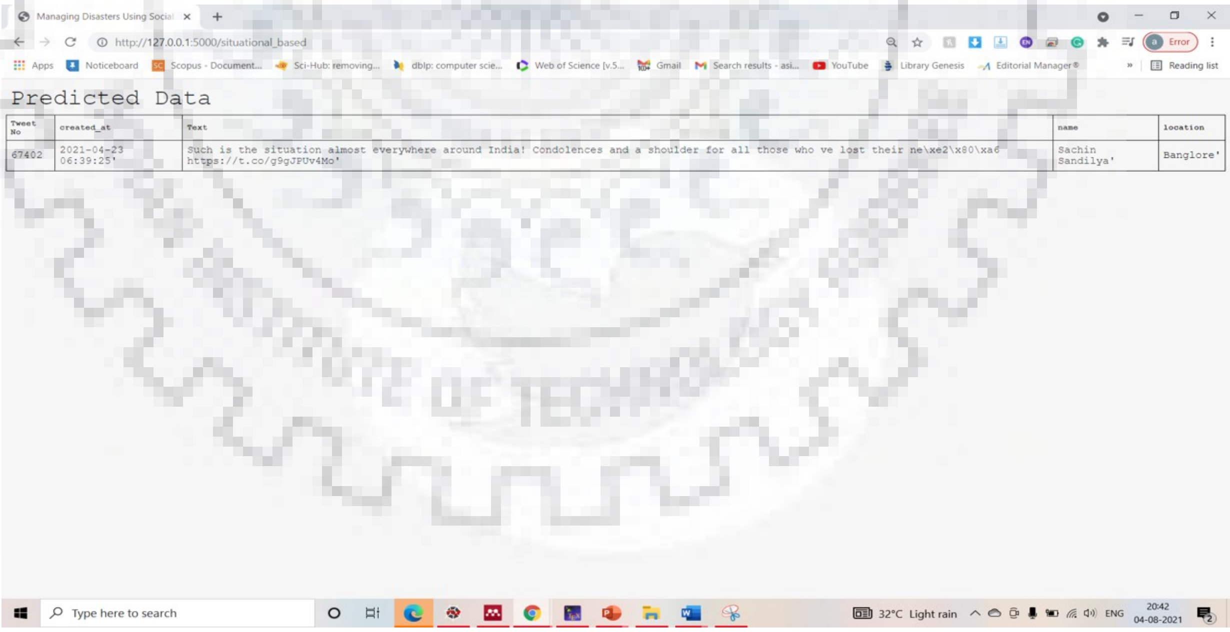


Fig. 7.14: Graphic user interface for result visualization.

7.6 User Experience

This section provides usage scenarios that demonstrate DisDSS's effectiveness and usability.

7.6.1 Usage Scenario 1

The user is interested in identifying people who need help during Covid-19 in the Delhi area. Therefore, the user selects India in the country drop-down, subsequently selecting Delhi in State/Metropolitan drop-down. After applying the Delhi filter, the user clicks the need-based button to see the need-based SM messages. Figure: 7.15 illustrates the output of the Delhi region's need-based SM messages.

The user explores the filtered messages in Figure: 7.15 and finds messages that say (tweet no. 96500, row 1) *urgently in need of an icu bed in Delhi, Ghaziabad, Noida, Gurgaon*. Tweet no. 94548, row 2 states that *Name: Rajnish Jain Area: Rani Bagh Oxygen Level 90 Age: 38 Urgently Bed Required in Delhi*. Similarly, tweet no. 97517, penultimate row in Figure: 7.15 mentions the *need for oxygen cylinders in Sar Gangaram Hospital Delhi, with more than 500 patients admitted*. After further finding the tweet messages in Figure: 7.15, the user finds misclassifications as well. Tweet no. 95885, row 4 states *CharuPragya Anyone can please help with a source for a bed in medanta #Lucknow ??This is urgent*. The SM message is asking for help in Lucknow, and not Delhi.

96500	2021-04-23 03:52:02'	Urgently in need of an icu bed in Delhi, Ghaziabad, Noida Gurgaon\n#COVIDEmergency2021 #CovidResources #COVIDSecondWaveInIndia #Covid'	Swati Bhatia'
94548	2021-04-22 17:58:30'	Name : Rajnish Jain\nArea : Rani Bagh\nOxygen Level < 90\nAge : 38\nUrgently Bed Required in Delhi \nPlease contact \n+9\xe2\x80\xa6 https://t.co/5OAQqRuzxW'	\xe0\xa4\xb9\xe0\xa4\xb0\xe0\xa4\xbf \xe0\xa4\xaa\xe0\xa4\xb0\xe0\xa4\xb0\xe
100556	2021-04-22 16:48:18'	Cn we tlk abt hw our cntry is dire need of O2 nt only delhi bt in ech n evry state its d tym dat the gov need to ge\xe2\x80\xa6 https://t.co/22r9k1PJTv'	Arpita jain'
95885	2021-04-23 07:05:33'	@CharuPragya \n\nAnyone can please help with a source for a bed in medanta #Lucknow ??\nThis is urgent for one of my a\xe2\x80\xa6 https://t.co/1P8s0LCYMl'	Divya'
98174	2021-04-22 17:58:30'	Name : Rajnish Jain\nArea : Rani Bagh\nOxygen Level < 90\nAge : 38\nUrgently Bed Required in Delhi \nPlease contact \n+9\xe2\x80\xa6 https://t.co/5OAQqRuzxW'	\xe0\xa4\xb9\xe0\xa4\xb0\xe0\xa4\xbf \xe0\xa4\xaa\xe0\xa4\xb0\xe0\xa4\xb0\xe
97107	2021-04-22 13:18:36'	India needs oxygen as well as a PM.\n#CovidIndia'	Mr Natwarlal'
98053	2021-04-22 13:18:41'	Why states like Delhi can't mandate hospital above 1000 beds capacity to build in house oxygen production capacity?\xe2\x80\xa6 https://t.co/iw83D4VRrD"	Mayank Sharma'
97517	2021-04-23 03:21:50'	Sar ganga ram hospital in Delhi hospital needs oxygen cyclinders urgently as they have more than 500 Corona patient\xe2\x80\xa6 https://t.co/q2jVlZ456S'	Sandhya\xf0\x9f\x90\xbc(ayush \xf0\x9f\xa5\x80) Rubina,\xe2\x9c\:
103125	2021-05-03 12:10:31'	Any leads for ventilator beds availability in Chennai hospitals? Would be of great help, thank you! \n#CovidHelp\xe2\x80\xa6 https://t.co/pYJIua4riQ'	Supriya Ramesh'

Fig. 7.15: Results displaying need-based messages in Delhi. The results contain tweet number, date and time of tweet, text of the message, name of the user.

7.6.2 Usage Scenario 2

To demonstrate the diverse functionality of the DisDSS, we apply a filter at availability-based button for India. Figure: 7.16 showcases the availability-based SM messages of India. Tweet no. 80869, row 1, states the *verified oxygen availability in Delhi with mobile number*. Furthermore, tweet no. 92172 (row 2), 84313 (row 3), 84241 (row 6), 72954 (row 7) mention the details about *icu beds availability* at Indore, and Lucknow. The misclassifications are depicted in a tweet no. 9163 (row 4), and 11138 (row 8).

Predicted Data

Tweet No	created_at	Text
80869	2021-04-23 04:04:54'	VERIFIED \n#oxygen will be available in #delhi TODAY ,i.e., 23rd April at 11:00 am, call +91 99994 45542.\xe2\x80\xa6 https://t.co/WEcJnpyUD3'
92172	2021-04-22 16:57:31'	VERIFIED BY OOMF ICU BEDS AVAILABLE IN INDORE \nCall - 9111158240 \n\n#INDORE \n#IndoreFightsCorona \n#CovidIndia\xe2\x80\xa6 https://t.co/l2uJnpaYGj'
84313	2021-04-23 05:14:06'	ICU BEDS AVAILABLE -\n\nContact no. - 9340819001\n\nName - Roshan patidar\n\nAdd - #maheshwar #khargon #indore \n\nVarifica\xe2\x80\xa6 https://t.co/jDeKqNNWUS'
9163	2021-04-23 04:22:43'	Come on yrr \xf0\x9f\x99\x8f @AdityaDubey2003\n@PuneetVuneet @iPriyankaBhatt\n\nPlz help for faridabad \xf0\x9f\x99\x8f \nRt \xe2\x99\xbb\xef\xb8\x8f share max guys\xe2\x80\xa6 https://t.co/QxmZb9iUg0'
77688	2021-05-05 13:41:27'	If you know in which hospital beds and oxygen are available tag us with full details and contact number. (2/2)\xe2\x80\xa6 https://t.co/VFKj1fto71'
84241	2021-04-23 06:28:16'	Oxygen cylinders in lucknow. Available from tomorrow. #COVIDEmergency2021 #Lucknow #Verified #COVID19India https://t.co/EEIsC1pQYZ'
72954	2021-04-22 22:19:26'	Oxygen Available :\n\nContact:9889556666 \n\nLocation :Lucknow \n\n#Verified 08:05\n#CovidHelp #CovidIndia\xe2\x80\xa6 https://t.co/rvIkQGF892'
11138	2021-04-22 17:31:31'	@NicolaCareem @yogital #CovidIndia \nAny authentic source from where I can get one Remdivir. In Delhi or NCR. \nNeed urgently \xf0\x9f\x86\x98'
	2021-04-	

Fig. 7.16: Results displaying availability-based messages in Delhi. The results contain tweet number, date and time of tweet, text of the message, name of the user.

7.7 Conclusion

The chapter seeks to develop DisDSS, a novel platform composed of a web browser using DL to help determine the nature of SM messages and help disaster professionals quickly respond to disasters. Our system attempts to reduce the gap between decision-makers and SM usage for DM. The idea is to expedite their use and reinforce the decision-making process with accurate and reliable knowledge to put together correct actions to mitigate the effects of a disaster.

We develop a new covid dataset with a manual label of *need-based*, *availability-based*, *situational-based*, *general*, and *irrelevant SM messages*. In the current study, we propose a hybrid model considering the contextual information of SM messages. It is composed of CNN and BiLSTM layers. In the empirical section, we conduct extensive experiments against different neural network models and existing research studies to validate the effectiveness of the proposed model. To demonstrate the technical potency of the proposed model, we developed the web interface to visualize the experimental results of the SM messages. The results show that the model significantly contributes to effective decision-making at catastrophic times.

Our DisDSS is developed with user convenience in mind, influencing our design, computational evaluation, and implementation choices. Our user-centered model and web interface contribute to both DL and DM communities. We bridge the two fields by demonstrating how models can be trained, evaluated, and used to facilitate need-based, availability-based mes¬sages as well as situational awareness for real-life, practical use.

8. CONCLUSION, LIMITATIONS, AND FUTURE SCOPE

This dissertation focuses on the KM of SM data for DM, which helps broaden our human understanding of knowledge that is encoded in the text through QDA and DL techniques. The dissertation consider DIKW model of KM, considering its evolution with the rapidly changing future. It helps in better organization and management of intellectual assets. To this end, the work is presented addressing the four research gaps identified from the literature review, one by one.

The first objective focuses on understanding the role of SM for DM. It explores the purposes of SM usage in emergency situations. The study investigates the challenges and enablers of SM usage in disaster management considering the Indian context by conducting FGD, following grounded theory approach. The objective works out complex information that requires unpacking the hidden barriers and enablers. It utilizes Atlas.ti software for results visualization.

The second objective aims to unravel the hidden treasures of information from the tsunami of data available on SM. The SM data is transformed into knowledge by proposing a framework, iRelevancy, to identify the relevancy of the SM message to the disaster, considering cyclonic and pandemic data. The proposed DL model comprises of LSTM and CNN layers. The objective contributes methodologically by proposing a hybrid model utilizing the best features of both CNN and LSTM. The experiments outperform the state-of-the-art techniques and existing studies, considering multiple performance metrics. The considered evaluation metrucs are accuracy, precision, recall, f-score, area under ROC, and area under PR-Curve. The results are promising on different disaster datasets.

The third objective offers a DL-based hybrid framework iStage, to determine the stage of the DMC from SM message. The proposed DL model consists of LSTM and CNN layers, utiziling the combined synergy of both models. It considers different cyclonic and non-cyclonic datasets to determine the stage of disaster from SM message. The experimental results outperform the existing techniques and studies. It considers six evaluation metrics on cyclonic and pandemic datasets. The study transmutes the raw and unstructured SM data into a structured format, i.e., knowledge, so that timely information helps in timely action at the right place.

The fourth objective develops DisDSS: an easy-to-use web interface, to categorize the SM messages into need-based, availability-based, situational-based, general, and irrelevant messages. The research proposes a DL-based model, consisting of CNN and BiLSTM layers. The research makes demonstration on Covid-19 disaster. Although there are several ways of categorizing information, a web-based platform is identified as the most suitable means due to its easy availability and flexibility. A web interface is developed so that all the stakeholders can understand it for effective decision-making.

This dissertation contributes to the existing knowledge of the usage of SM data for DM. The thesis explores the SM through the lens of DL techniques. The dissertation provides contextual and methodological contribution by developing a web-based interface using deep learning algorithms, specifically for DM domain. It bridges a significant gap between people and disaster professionals. The outcome of this dissertation can be used by policymakers to devise and implement more insightful and potentially practical policies for DM.

The study has a limitation that deserves to be analyzed and commented on. Although the participants are working in an academic environment of DM, they may not entirely represent all people in DM. The research study has a limitation corresponding to transferability.

In terms of future research directions, it would be of real interest to have disaster managers onboard and a deeper understanding of the challenges and enablers of SM in catastrophic situations. Future studies should consider that disaster operations vary from disaster to disaster, and all disasters are not of equal magnitude and intensity. Therefore, future researchers might consider a specific disaster type to understand the subject thoroughly.

Even though the results obtained are more than encouraging concerning the applicability and efficacy of the DL frameworks, there are some limitations and research directions that can be considered for future research.

First, the dissertation considers only Twitter platform. In the future, other SM platforms data will be considered in order to obtain an enormous amount of data. Second, the research study uses manual labeling of data, which is highly time-consuming that further affects the timely information sharing. The plan is to apply automatic labeling methods to avoid this limitation in the future. Thirdly, it is reasonable to mention that the dynamic nature of the Twitter platform, compels the framework to be dynamic in nature. And the author would leave that problem for the future.

Bibliography

Abbasimehr, H. and Paki, R. (2020). Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos, Solitons and Fractals*, 142.

Abdelminaam, D. S., Ismail, F. H., Taha, M., Taha, A., Houssein, E. H., and Nabil, A. (2021). CoAID-DEEP: An Optimized Intelligent Framework for Automated Detecting COVID-19 Misleading Information on Twitter. *IEEE Access*, 9(December 2019):27840–27867.

Abedin, B., Babar, A., and Abbasi, A. (2014). Characterization of the Use of Social Media in Natural Disasters: A Systematic Review. *2014 IEEE Fourth International Conference on Big Data and Cloud Computing*, (2004):449–454.

Abel, F., Hauff, C., Hoube, G. J., Tao, K., and Stronkman, R. (2012a). Semantics + Filtering + Search = Twitcident exploring information in social web streams. *HT'12 - Proceedings of 23rd ACM Conference on Hypertext and Social Media*, pages 285–294.

Abel, F., Hauff, C., Houben, G. J., Tao, K., and Stronkman, R. (2012b). Twitcident: Fighting fire with information from Social Web streams. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion*, pages 305–308.

Agarwal, A. and Toshniwal, D. (2020). Identifying Leadership Characteristics from Social Media Data during Natural Hazards using Personality Traits. *Scientific Reports*, 10(1):1–15.

Ahmed, A. (2011). Using Social Media in Disaster Management. *Proceedings of the 32th International Conference on Information Systems*.

Ajao, O., Bhowmik, D., and Zargari, S. (2018). Fake News Identification on Twitter with Hybrid CNN and RNN Models. *Proceedings of the International Conference on Social Media & Society*, number July.

Akter, S. and Wamba, S. F. (2017). Big data and disaster management: a systematic review and agenda for future research. *Annals of Operations Research*, pages 1–21.

Al-Saggaf, Y. and Simmons, P. (2015). Social media in Saudi Arabia: Exploring its use during two natural disasters. *Technological Forecasting and Social Change*, 95:3–15.

Alam, F., Ofli, F., and Imran, M. (2018). Processing Social Media Images by Combining Human and Machine Computing during Crises. *International Journal of Human-Computer Interaction*, 34(4):311–327.

Alampay, E. A., Asuncion, X. V., and delos Santos, M. (2018). Management of Social Media for Disaster Risk Reduction and Mitigation in Philippine Local Government Units. *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance - ICEGOV '18*, pages 183–190.

Alhussein, M., Aurangzeb, K., and Member, S. (2020). Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting. *IEEE Access*, 8.

Amin, M. S. and Ahn, H. (2021). Earthquake disaster avoidance learning system using deep learning. *Cognitive Systems Research*, 66:221–235.

Anderson, K. M., Aydin, A. A., Barrenechea, M., Cardenas, A., Hakeem, M., and Jambi, S. (2015). Design challenges/solutions for environments supporting the analysis of social media data in crisis informatics research. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, pages 163–172.

Anikeeva, O., Steenkamp, M., and Arbon, P. (2015). The future of social media use during emergencies in Australia: Insights from the 2014 Australian and New Zealand Disaster and Emergency Management Conference social media workshop. *Australian Journal of Emergency Management*, 30(1):22–26.

Anjali Marar, A. S. (2020). Cyclone Nisarga Explained: How big is the threat on west coast?

Anson, S., Watson, H., Wadhwa, K., and Metz, K. (2017). Analysing social media data for disaster preparedness: Understanding the opportunities and barriers faced by humanitarian actors. *International Journal of Disaster Risk Reduction*, 21(November 2016):131–139.

Aven, T. (2013). A conceptual framework for linking risk and the elements of the data-information-knowledge-wisdom (DIKW) hierarchy. *Reliability Engineering and System Safety*, 111:30–36.

Badam, S. K., Kisselburgh, L., and Ramani, K. (2017). Integrating Visual Analytics Support for Grounded Theory Practice in Qualitative Text Analysis. *Eurographics Conference on Visualization (EuroVis) 2017*, volume 36.

Bandera, C., Keshtkar, F., Bartolacci, M. R., Neerudu, S., and Passerini, K. (2017). Knowledge management and the entrepreneur: Insights from Ikujiro Nonaka's Dynamic Knowledge Creation model (SECI). *International Journal of Innovation Studies*, 1(3):163–174.

Basheer Ahammed, K. K. and Pandey, A. C. (2021). Characterization and impact assessment of super cyclonic storm AMPHAN in the Indian subcontinent through space borne observations. *Ocean and Coastal Management*, 205(January).

Basu, M., Ghosh, S., Jana, A., Bandyopadhyay, S., and Singh, R. (2017). Resource mapping during a natural disaster: A case study on the 2015 Nepal earthquake. *International Journal of Disaster Risk Reduction*, 24(May):24–31.

Basu, M., Shandilya, A., Khosla, P., Ghosh, K., and Ghosh, S. (2019). Extracting Resource Needs and Availabilities From Microblogs for Aiding Post-Disaster Relief Operations. *IEEE Transactions on Computational Social Systems*, 6(3):1–15.

Behl, S., Rao, A., Aggarwal, S., Chadha, S., and Pannu, H. S. (2021). Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. *International Journal of Disaster Risk Reduction*, 55(January).

Bertot, J. C., Jaeger, P. T., and Hansen, D. (2012). The impact of polices on government social media usage: Issues, challenges, and recommendations. *Government Information Quarterly*, 29(1):30–40.

Bhushan, N., Vu, M., Teal, R., Carda-Auten, J., Ward, D., and Erinosho, T. (2017). Assessing Challenges in Low-Income Families to Inform a Life Skills–Based Obesity Intervention. *Health Promotion Practice*, pages 1–10.

Burel, G. and Alani, H. (2018). Crisis event extraction service (CREES) – Automatic detection and classification of crisis-related content on social media. *Proceedings of the International ISCRAM Conference*, 2018-May(May):597–608.

Caballero-Anthony, M., Cook, A. D., and Chen, C. (2021). Knowledge management and humanitarian organisations in the Asia-Pacific: Practices, challenges, and future pathways. *International Journal of Disaster Risk Reduction*, 53(November 2020).

Caragea, C., Silvescu, A., and Tapia, A. (2016a). Identifying informative messages in disaster events using convolutional neural networks. *ISCRAM 2016 Conference Proceedings – 13th International Conference on Information Systems for Crisis Response and Management*, number May.

Caragea, C., Silvescu, A., and Tapia, A. H. (2016b). Identifying informative messages in disaster events using Convolutional Neural Networks. *Proceedings of the International ISCRAM Conference*, (May).

Centre for Research on the Epidemiology of Disasters and United Nations Office for Disaster Risk Reduction (2021). 2020: The Non-COVID Year in Disasters. Technical report.

Chakraborty, K., Bhatia, S., Bhattacharyya, S., Platos, J., and Bag, R. (2020). Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers — A study to show how popularity is affecting accuracy in social media. *Applied Soft Computing Journal*, 97:106754.

Chaudhuri, N. and Bose, I. (2020). Exploring the role of deep neural networks for post-disaster decision support. *Decision Support Systems*, 130(July 2019):113234.

Chen, J., Chen, H., Hu, D., Pan, J. Z., and Zhou, Y. (2015). Smog disaster forecasting using social web data and physical sensor data. *Big Data (Big Data), 2015 IEEE International Conference on*, pages 991–998.

Chen, R., Wang, X., Zhang, W., Zhu, X., Li, A., and Yang, C. (2019). A hybrid CNN-LSTM model for typhoon formation forecasting. *GeoInformatica*, 23(3):375–396.

Chikaraishi, M., Garg, P., Varghese, V., Yoshizoe, K., Urata, J., Shiomi, Y., and Watanabe, R. (2020). On the possibility of short-term traffic prediction during disaster with machine learning approaches: An exploratory analysis. *Transport Policy*, 98(June):91–104.

Chowdhury, R., Chowdhury, S. R., and Castillo, C. (2013). Tweet4act : Using Incident-Specific Profiles for Classifying Crisis-Related Messages. *Proceedings of the 10th International ISCRAM Conference*, number May, pages 834–839.

Comunello, F. and Mulargia, S. (2017). A # cultural _ change is needed . Social media use in emergency communication by Italian local level institutions. *Proceedings of the 14th ISCRAM Conference*, pages 512–521.

Dash, B. and Walia, A. (2020). Role of multi-purpose cyclone shelters in India: Last mile or neighbourhood evacuation. *Tropical Cyclone Research and Review*, 9(4):206–217.

Devaraj, A., Murthy, D., and Dontula, A. (2020). Machine-learning methods for identifying social media-based requests for urgent help during hurricanes. *International Journal of Disaster Risk Reduction*, 51:101757.

Do, H. H., Prasad, P. W., Maag, A., and Alsadoon, A. (2019). Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review. *Expert Systems with Applications*, 118:272–299.

Domdouzis, K., Akhgar, B., Andrews, S., Gibson, H., and Hirsch, L. (2016). A social media and crowdsourcing data mining system for crime prevention during and post-crisis situations. *Journal of Systems and Information Technology*, 18(4):364–382.

Dorasamy, M., Raman, M., and Kaliannan, M. (2013). Knowledge management systems in support of disasters management: A two decade review. *Technological Forecasting and Social Change*, 80(9):1834–1853.

Dutt, R., Basu, M., Ghosh, K., and Ghosh, S. (2019). Utilizing microblogs for assisting post-disaster relief operations via matching resource needs and availabilities. *Information Processing and Management*, 56(5):1680–1697.

Eeuwijk, P. V. and Angehrn, Z. (2017). How to.... conduct a focus group discussion (FGD): Methodological manual by Peter van Eeuwijk and Zuzanna Angehrn. Technical Report April.

Fanchiotti, M., Dash, J., Tompkins, E. L., and Hutton, C. W. (2020). The 1999 super cyclone in Odisha, India: A systematic review of documented losses. *International Journal of Disaster Risk Reduction*, 51(September):75–78.

Feldman, D., Contreras, S., Karlin, B., Basolo, V., Matthew, R., Sanders, B., Houston, D., Cheung, W., Goodrich, K., Reyes, A., Serrano, K., Schubert, J., and Luke, A. (2016). Communicating flood risk: Looking back and forward at traditional and social media outlets. *International Journal of Disaster Risk Reduction*, 15:43–51.

Feng, Y. and Sester, M. (2018). Extraction of pluvial flood relevant volunteered geographic information (VGI) by deep learning from user generated texts and photos. *ISPRS International Journal of Geo-Information*, 7(2).

Friese, S. (2016). Computer-Assisted Grounded Theory Analysis With ATLAS.ti. *ATLAS.ti User Conference 2015 – qualitative data analysis and beyond*, pages 1–22.

Gao, D.-w., Zhu, Y.-s., Zhang, J.-f., He, Y.-k., Yan, K., and Yan, B.-r. (2021). A novel MP-LSTM method for ship trajectory prediction based on AIS data. *Ocean Engineering*, 228.

Gaspar, R., Yan, Z., and Domingos, S. (2019). Extreme natural and man-made events and human adaptive responses mediated by information and communication technologies' use: A systematic literature review. *Technological Forecasting and Social Change*, 145(January):125–135.

Giannakas, F., Troussas, C., Voyiatzis, I., and Sgouropoulou, C. (2021). A deep learning classification framework for early prediction of team-based academic performance. *Applied Soft Computing*, 106.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. `http://www.deeplearningbook.org`.

Goswami, R., Roy, K., Dutta, S., Ray, K., Sarkar, S., Brahmachari, K., Kr, M., Mainuddin, M., Banerjee, H., Timsina, J., and Majumdar, K. (2021). Multi-faceted impact and outcome of COVID-19 on smallholder agricultural systems : Integrating qualitative research and fuzzy cognitive mapping to explore resilient strategies. *Agricultural Systems*, 189(December 2020):103051.

Goswami, S., Chakraborty, S., Ghosh, S., Chakrabarti, A., and Chakraborty, B. (2016). A review on application of data mining techniques to combat natural disasters. *Ain Shams Engineering Journal*, 80(1).

Gray, B. J., Weal, M. J., and Martin, D. (2017). Social Media and Disasters. *International Journal of Information Systems for Crisis Response and Management*, 8(4):41–55.

Gunessee, S., Subramanian, N., Roscoe, S., and Ramanathan, J. (2017). The social preferences of local citizens and spontaneous volunteerism during disaster relief operations*. *International Journal of Production Research*, 7543:pp.1–16.

Habdank, M., Rodehutskors, N., and Koch, R. (2017). Relevancy Assessment of Tweets using Supervised Learning Techniques Mining emergency related Tweets for automated relevancy classification. *2017 4th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*.

Hagen, L., Scharf, R., Neely, S., and Keller, T. (2018). Government social media communications during zika health crisis. *Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age - dgo '18*, pages 1–10.

Han, J., Kamber, M., and Pei, J. (2011). *Data Mining Concepts and Techniques*.

Hiltz, S. R. and Kushma, J. (2014). Use of Social Media by U . S . Public Sector Emergency Managers : Barriers and Wish Lists. *Proceedings of the 11th International ISCRAM Conference*, pages 602–611.

Hochreiter, S. (1997). Long Short-Term Memory. 1780:1735–1780.

Hofmann, M., Betke, H., and Sackmann, S. (2015). "Process-oriented disaster response management: a structured literature review". *Business Process Management Journal*, 21(5):966–987.

Huang, C.-M., Chan, E., and Hyder, A. A. (2016). Web 2.0 and Internet Social Networking: A New Tool for Disaster Management? Lessons from Taiwan. *BMC Medical Informatics and Decision Making*, pages 191–200.

Huang, X., Li, Z., Wang, C., and Ning, H. (2020a). Identifying disaster related social media for rapid response: a visual-textual fused CNN architecture. *International Journal of Digital Earth*, 13(9):1017–1039.

Huang, X., Li, Z., Wang, C., and Ning, H. (2020b). Identifying disaster related social media for rapid response: a visual-textual fused CNN architecture. *International Journal of Digital Earth*, 13(9):1017–1039.

Imran, M., Castillo, C., Lucas, J., Meier, P., and Vieweg, S. (2014). AIDR: Artificial intelligence for disaster response. In *WWW 2014 Companion - Proceedings of the 23rd International Conference on World Wide Web*, number October, pages 159–162.

Inan, D. I., Beydoun, G., and Pradhan, B. (2018). Developing a decision support system for Disaster Management: Case study of an Indonesia volcano eruption. *International Journal of Disaster Risk Reduction*, 31(July):711–721.

Jamali, M., Nejat, A., Ghosh, S., Jin, F., and Cao, G. (2019). Social media data and post-disaster recovery. *International Journal of Information Management*, 44(September 2018):25–37.

Jelodar, H., Wang, Y., Orji, R., and Huang, S. (2020). Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach. *IEEE Journal of Biomedical and Health Informatics*, 24(10):2733–2742.

Kaewkitipong, L., Chen, C. C., and Ractham, P. (2016). A community-based approach to sharing knowledge before, during, and after crisis events: A case study from Thailand. *Computers in Human Behavior*, 54:653–666.

Kanagarajoo, M. V. (2018). A Framework for Social Media Use in Project Management.

Kanjo, E., Younis, E. M., and Ang, C. S. (2019). Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. *Information Fusion*, 49(August 2018):46–56.

Kavanaugh, A. L., Fox, E. A., Sheetz, S. D., Yang, S., Li, L. T., Shoemaker, D. J., Natsev, A., and Xie, L. (2012). Social media use by government: From the routine to the critical. *Government Information Quarterly*, 29(4):480–491.

Kavota, J. K., Kamdjoug, J. R. K., and Wamba, S. F. (2020). Social media and disaster management: Case of the north and south Kivu regions in the Democratic Republic of the Congo. *International Journal of Information Management*, 52(August 2019):102068.

Kibanov, M., Stumme, G., Amin, I., and Gun, J. (2017). Mining social media to inform peatland fire and haze disaster management. *Social Network Analysis and Mining*, pages 1–19.

Kim, J. and Hastak, M. (2018). Social network analysis : Characteristics of online social networks after a disaster. *International Journal of Information Management*, 38(August 2017):86–96.

Kim, K., Jung, K., and Chilton, K. (2016). Strategies of social media use in disaster management: Lessons in resilience from Seoul, South Korea. *International Journal of Emergency Services*, 5(2):110–125.

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751.

Kirac, E. and Milburn, A. B. (2018). A general framework for assessing the value of social data for disaster response logistics planning. *European Journal of Operational Research*, 269(2):486–500.

Koenig, M. E. (2011). *Knowledge Management in Theory and Practice (2nd ed.)*, volume 62.

Kundu, S. (2018). Classification of Short-Texts Generated During Disasters : A Deep Neural Network Based Approach. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 790–793.

Kusumastuti, R. D., Arviansyah, A., Nurmala, N., and Wibowo, S. S. (2021). Knowledge management and natural disaster preparedness: A systematic literature review and a case study of East Lombok, Indonesia. *International Journal of Disaster Risk Reduction*, 58(December 2020).

Lambert, C. E. (2020). Earthquake Country : A Qualitative Analysis of Risk Communication via Facebook Earthquake Country : A Qualitative Analysis of Risk Communication. *Environmental Communication*, 0(0):1–14.

Li, H., Guevara, N., Herndon, N., Caragea, D., Neppalli, K., Caragea, C., Squicciarini, A., and Tapia, A. H. (2015). Twitter Mining for Disaster Response: A Domain Adaptation Approach. *12th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2015)*, pages 1–7.

Li, M. and Gao, F. (2003). Why Nonaka highlights tacit knowledge: A critical review. *Journal of Knowledge Management*, 7(4):6–14.

Li, T., Hua, M., and Wu, X. U. (2020). A Hybrid CNN-LSTM Model for Forecasting. *IEEE Access*, pages 26933–26940.

Li, Y., Chai, S., Ma, Z., and Wang, G. (2021). A Hybrid Deep Learning Framework for Long-Term Traffic Flow Prediction. *IEEE Access*, 9:11264–11271.

Liang, W., Yao, J., Chen, A., Lv, Q., Zanin, M., Liu, J., Wong, S. S., Li, Y., Lu, J., Liang, H., Chen, G., Guo, H., Guo, J., Zhou, R., Ou, L., Zhou, N., Chen, H., Yang, F., Han, X., Huan, W., Tang, W., Guan, W., Chen, Z., Zhao, Y., Sang, L., Xu, Y., Wang, W., Li, S., Lu, L., Zhang, N., Zhong, N., Huang, J., and He, J. (2020). Early triage of critically ill COVID-19 patients using deep learning. *Nature Communications*, 11(1):1–7.

Lieneck, C., Heinemann, K., Patel, J., Huynh, H., Leafblad, A., Moreno, E., and Wingfield, C. (2022). Facilitators and Barriers of COVID-19 Vaccine Promotion on Social Media in the United States: A Systematic Review. *Healthcare (Switzerland)*, 10(2).

Linders, D. (2012). From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, 29(4):446–454.

Link, D., Hellingrath, B., and Ling, J. (2016). A human-is-the-loop approach for semi-automated content moderation. *Proceedings of the ISCRAM 2016 Conference*, (May).

Liu, G. and Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337:325–338.

Liu, Z.-x., Zhang, D.-g., Luo, G.-z., Lian, M., and Liu, B. (2020). A new method of emotional analysis based on CNN–BiLSTM hybrid neural network. *Cluster Computing*, 23(4):2901–2913.

Luna, S. and Pennock, M. (2015). Social media in emergency management advances, challenges and future directions. *2015 Annual IEEE Systems Conference (SysCon) Proceedings*, pages 792–797.

Luo, Y. and Xu, X. (2021). Comparative study of deep learning models for analyzing online restaurant reviews in the era of the COVID-19 pandemic. *International Journal of Hospitality Management*, 94(September 2020).

McKenna, B., Myers, M. D., and Newman, M. (2017). Social media in qualitative research: Challenges and recommendations. *Information and Organization*, 27(2):87–99.

Meneghello, J., Thompson, N., Lee, K., Wong, K. W., and Abu-Salih, B. (2020). Unlocking social media and user generated content as a data source for knowledge management. *International Journal of Knowledge Management*, 16(1):101–122.

Mishra, D. R., Kumar, A., Muduli, P. R., Acharyya, T., Acharya, P., Singh, S., and Rastogi, G. (2021). Landfall season is critical to the impact of a cyclone on a monsoon-regulated tropical coastal lagoon. *Science of the Total Environment*, 770:145235.

Mols, A. and Pridmore, J. (2019). The Blurring of Boundaries in WhatsApp Neighbourhood Crime Prevention Groups in The Netherlands. *Surveillance & Society*, 17(4):272–287.

Morrow, N., Mock, N., APapendieck, A., and Kocmich, N. (2011). Independent Evaluation of the Ushahidi Haiti Project.

Muhammad, K., Ahmad, J., Lv, Z., Bellavista, P., Yang, P., and Baik, S. W. (2019). Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(7):1419–1434.

Naseem, U., Razzak, I., Khushi, M., Eklund, P. W., and Kim, J. (2021). COVIDSenti: A Large-Scale Benchmark Twitter Data Set for COVID-19 Sentiment Analysis. *IEEE Transactions on Computational Social Systems*, 8(4):976–988.

National Disaster Management Authority, M. o. H. A. (2019). National Disaster Management Plan. Technical Report November.

Neppalli, V. K., Medeiros, M. C., Caragea, C., Caragea, D., Tapia, A. H., and Halse, S. (2016). Retweetability analysis and prediction during Hurricane sandy. *Proceedings of the International ISCRAM Conference*, number May.

Nguyen, D. T., Joty, S., Imran, M., Sajjad, H., and Mitra, P. (2016). Applications of Online Deep Learning for Crisis Response Using Social Media Information. *25th Conference of Information and Knowledge Management (CIKM)*.

Nguyen, L., Yang, Z., Li, J., Pan, Z., Cao, G., and Jin, F. (2019). Forecasting People's Needs in Hurricane Events from Social Network. *IEEE Transactions on Big Data*, pages 1–1.

Ofli, F., Alam, F., and Imran, M. (2020). Analysis of Social Media Data using Multimodal Deep Learning for Disaster Response. *17th ISCRAM Conference*, number May, Blacksburg, VA, USA.

Oktari, R. S., Munadi, K., Idroes, R., and Sofyan, H. (2020). Knowledge management practices in disaster management: Systematic review. *International Journal of Disaster Risk Reduction*, 51(January):101881.

Panagiotopoulos, P., Barnett, J., Bigdeli, A. Z., and Sams, S. (2016). Social media in emergency management: Twitter as a tool for communicating risks to the public. *Technological Forecasting and Social Change*, 111:86–96.

Park, C. H. and Johnston, E. W. (2017). A framework for analyzing digital volunteer contributions in emergent crisis response efforts. *New Media and Society*, 19(8):1308–1327.

Pereira, J., Monteiro, J., Silva, J., Estima, J., and Martins, B. (2020). Assessing flood severity from crowdsourced social media photos with deep neural networks. *Multimedia Tools and Applications*.

Petersen, L., Fallou, L., Havarneanu, G., Reilly, P., Serafinelli, E., and Bossu, R. (2018). November 2015 Paris Terrorist Attacks and Social Media Use: Preliminary Findings from Authorities, Critical Infrastructure Operators and Journalists. *15th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2018)*, pages 629–638.

Petmezas, G., Haris, K., Stefanopoulos, L., and Kilintzis, V. (2021). Biomedical Signal Processing and Control Automated Atrial Fibrillation Detection using a Hybrid CNN-LSTM Network on Imbalanced ECG Datasets. *Biomedical Signal Processing and Control*, 63(August 2020):102194.

Plotnick, L. and Hiltz, S. R. (2018). Software Innovations to Support the Use of Social Media by Emergency Managers. *International Journal of Human-Computer Interaction*, 34(4):367–381.

Powell, T. H. (2007). A critical review of Nonaka's SECI Framework. *16th EDAMBA Summer Academy*, (July):1–15.

Pretorius, J. G., Mathews, M. J., Maré, P., Kleingeld, M., and van Rensburg, J. (2019). Implementing a DIKW model on a deep mine cooling system. *International Journal of Mining Science and Technology*, 29(2):319–326.

Qu, Z., Li, B., Wang, X., Yin, S., and Zheng, S. (2018). An Efficient Recommendation Framework on Social Media Platforms Based on Deep Learning. *Proceedings - 2018 IEEE International Conference on Big Data and Smart Computing, BigComp 2018*, pages 599–602.

Rambaree, K. (2012). Three Methods Of Qualitative Data Analysis Using ATLAS . ti : ' A Posse Ad Esse '. *ATLAS.ti User Conference 2013*, pages 1–15.

Ranjan, S. and Gupta, B. B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing Journal*, 100:106983.

Reuter, C., Ludwig, T., Kaufhold, M. A., and Spielhofer, T. (2016). Emergency services attitudes towards social media: A quantitative and qualitative survey across Europe. *International Journal of Human Computer Studies*, 95:96–111.

Rexiline Ragini, J., Rubesh Anand, P. M., and Bhaskar, V. (2018). Mining crisis information: A strategic approach for detection of people at risk through social media analysis. *International Journal of Disaster Risk Reduction*, 27(August 2017):556–566.

Rhem, A. J. (2021). AI ethics and its impact on knowledge management. *AI and Ethics*, 1(1):33–37.

Richter, I. L. (2006). A critique of Nonaka ' s SECI model Key words Introduction and research objectives Used methods and research results.

Sarker, M. N. I., Peng, Y., Yiran, C., and Shouse, R. C. (2020). Disaster resilience through big data: Way to environmental sustainability. *International Journal of Disaster Risk Reduction*, 51(March).

Scientific Software Development GmbH (2013). Atlas.ti 7.1 User Guide and Reference. pages 1–469.

Simon, T., Goldberg, A., and Adini, B. (2015). Socializing in emergencies - A review of the use of social media in emergency situations. *International Journal of Information Management*, 35(5):609–619.

Simon Kemp (2021). Digital 2021: Global Overview Report — DataReportal – Global Digital Insights.

Singh, J. P., Dwivedi, Y. K., Rana, N. P., Kumar, A., and Kapoor, K. K. (2017). Event classification and location prediction from tweets during disasters. *Annals of Operations Research*, pages 1–21.

Sit, M. A., Koylu, C., and Demir, I. (2019). Identifying disaster-related tweets and their semantic , spatial and temporal context using deep learning , natural language processing and spatial analysis : a case study of Hurricane Irma. *International Journal of Digital Earth*, 8947.

Snyder, L. S., Lin, Y.-S., Karimzadeh, M., Goldwasser, D., and Ebert, D. S. (2019). Interactive Learning for Identifying Relevant Tweets to Support Real-time Situational Awareness. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1–1.

Society, I. R. C. (2019). Odisha FANI cyclone Assessment Report. Technical report.

Squicciarini, A., Tapia, A., and Stehle, S. (2017). Sentiment analysis during Hurricane Sandy in emergency response. *International Journal of Disaster Risk Reduction*, 21(May 2016):213–222.

Subramani, S., Wang, H. U. A., Vu, H. U. Y. Q., Li, G., and Member, S. (2018). Domestic Violence Crisis Identification From Facebook Posts Based on Deep Learning. *IEEE Access*, 6:54075–54085.

Tagliacozzo, S. (2018). Government Agency Communication during Postdisaster Reconstruction: Insights from the Christchurch Earthquakes Recovery. *Natural Hazards Review*, 19(2):04018001.

Takahashi, B., Tandoc, E. C., and Carmichael, C. (2015). Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines. *Computers in Human Behavior*, 50:392–398.

Tam, S., Said, R. B., and Tanriöver, Ö. (2021). A ConvBiLSTM Deep Learning Model-Based Approach for Twitter Sentiment Classification. *IEEE Access*, 9:41283–41293.

Tang, Z., Zhang, L., Xu, F., and Vo, H. (2015). Examining the role of social media in California's drought risk management in 2014. *Natural Hazards*, 79(1):171–193.

Tarasconi, F., Farina, M., Mazzei, A., and Bosca, A. (2017). The role of unstructured data in real-time disaster-related social media monitoring. *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, 2018-January:3769–3778.

Tian, H., Cen Zheng, H., and Chen, S. C. (2018). Sequential Deep Learning for Disaster-Related Video Classification. In *Proceedings - IEEE 1st Conference on Multimedia Information Processing and Retrieval, MIPR 2018*, pages 106–111. IEEE.

Truelove, V., Freeman, J., and Davey, J. (2019). "I Snapchat and Drive!" A mixed methods approach examining snapchat use while driving and deterrent perceptions among young adults. *Accident Analysis and Prevention*, 131(August 2018):146–156.

Varsamopoulos, S., Bertels, K., and Almudever, C. G. (2018). Designing neural network based decoders for surface codes Accelerated BWA-MEM View project hartes View project Designing neural network based decoders for surface codes. *arXiv:1811.12456 [quant-ph]*, (November):1–12.

Verma, S., Vieweg, S., Corvey, W. J., Palen, L., Martin, J. H., Palmer, M., Schram, A., and Anderson, K. M. (2011). Natural Language Processing to the Rescue? Extracting "Situational Awareness" Tweets During Mass Emergency. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, (January):385–392.

Wang, S. H., Nayak, D. R., Guttery, D. S., Zhang, X., and Zhang, Y. D. (2021). COVID-19 classification by CCSHNet with deep fusion using transfer learning and discriminant correlation analysis. *Information Fusion*, 68(October 2020):131–148.

Williams, D. (2014). Models, Metaphors and Symbols for Information and Knowledge Systems. *Journal of Entrepreneurship, Management and Innovation*, 10(1):79–107.

Wilson, T., Stanek, S., Spiro, E., and Starbird, K. (2017). Language limitations in rumor research? Comparing French and english tweets sent during the 2015 Paris attacks. *Proceedings of the International ISCRAM Conference*, pages 546–553.

Wladdimiro, D., Gonzalez-Cantergiani, P., Hidalgo, N., and Rosas, E. (2016). Disaster management platform to support real-Time analytics. *Proceedings of the 2016 3rd International Conference on Information and Communication Technologies for Disaster Management, ICT-DM 2016*.

Xia, H., An, W., Li, J., and Justin, Z. (2020). Outlier knowledge management for extreme public health events : Understanding public opinions about COVID-19 based on microblog data. *Socio-Economic Planning Sciences*, (July):100941.

Yabe, T. and Ukkusuri, S. V. (2019). Integrating information from heterogeneous networks on social media to predict post-disaster returning behavior. *Journal of Computational Science*, 32:12–20.

Yang, J., Yu, M., Qin, H., Lu, M., and Yang, C. (2019). A twitter data credibility framework—Hurricane Harvey as a use case. *ISPRS International Journal of Geo-Information*, 8(3).

Yasin Kabir, M., Gruzdev, S., and Madria, S. (2020). STIMULATE: A System for Real-Time Information Acquisition and Learning for Disaster Management. *Proceedings - IEEE International Conference on Mobile Data Management*, 2020-June(Mdm):186–193.

Yates, D. and Paquette, S. (2011). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *International Journal of Information Management*, 31(1):6–13.

Yu, M., Huang, Q., Qin, H., Scheele, C., and Yang, C. (2019). Deep learning for real-time social media text classification for situation awareness – using Hurricanes Sandy , Harvey , and Irma as case studies. *International Journal of Digital Earth*, 0(0):1–18.

Yuan, F., Li, M., Liu, R., Zhai, W., and Qi, B. (2021). Social media for enhanced understanding of disaster resilience during Hurricane Florence. *International Journal of Information Management*, 57(December 2020):tt.

Zeroual, A., Harrou, F., Dairi, A., and Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons and Fractals*, 140.

Zhai, W., Peng, Z.-r., and Yuan, F. (2020). International Journal of Disaster Risk Reduction Examine the effects of neighborhood equity on disaster situational awareness : Harness machine learning and geotagged Twitter data. *International Journal of Disaster Risk Reduction*, 48(December 2019):101611.

Zhang, D., Zhou, L., and Lim, J. (2020a). From Networking to Mitigation: The Role of Social Media and Analytics in Combating the COVID-19 Pandemic. *Information Systems Management*, 37(4):318–326.

Zhang, X. C., Gong, J. G., and Xuan, F. Z. (2021). A deep learning based life prediction method for components under creep, fatigue and creep-fatigue conditions. *International Journal of Fatigue*, 148(October 2020).

Zhang, Y., Yan, B., and Aasma, M. (2020b). A novel deep learning framework : Prediction and analysis of financial time series using CEEMD and LSTM. *Expert Systems With Applications*, 159:113609.

Zheng, L. (2013). Social media in Chinese government: Drivers, challenges and capabilities. *Government Information Quarterly*, 30(4):369–376.

Zhu, J., Chen, H., and Ye, W. (2020). A Hybrid CNN – LSTM Network for the Classification of Human Activities Based on Micro-Doppler Radar. *IEEE Access*, 8:24713–24720.

Zuheros, C., Martínez-Cámara, E., Herrera-Viedma, E., and Herrera, F. (2021). Sentiment Analysis based Multi-Person Multi-criteria Decision Making methodology using natural language processing and deep learning for smarter decision aid. Case study of restaurant choice using TripAdvisor reviews. *Information Fusion*, 68(October 2020):22–36.