Blockchain-Based Event Detection and Trust Verification Using Natural Language Processing and Machine Learning

ZEINAB SHAHBAZI AND YUNG-CHEOL BYUN

Department of Computer Engineering, Institute of Information Science Technology, Jeju National University, Jeju-si 63243, South Korea

Corresponding author: Yung-Cheol Byun (ycb@jejunu.ac.kr)

ABSTRACT Information sharing is one of the huge topics in social media platform regarding the daily news related to events or disasters happens in nature or its human-made. The automatic urgent need identification and sharing posts and information delivery with a short response are essential tasks in this area. The key goal of this research is developing a solution for management of disasters and emergency response using social media platforms as a core component. This process focuses on text analysis techniques to improve the process of authorities in terms of emergency response and filter the information using the automatically gathered information to support the relief efforts. Specifically, we used state-of-art Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP) based on supervised and unsupervised learning using social media datasets to extract real-time content related to the emergency events to comfort the fast response in a critical situation. Similarly, the blockchain framework used in this process for trust verification of the detected events and eliminating the single authority on the system. The main reason of using the integrated system is to improve the system security and transparency to avoid sharing the wrong information related to an event in social media.

INDEX TERMS Event detection, machine learning, blockchain, natural language processing, deep learning.

I. INTRODUCTION

Disasters are part of the daily news in social media during the past few years. There is various type of disasters such as earthquake, flood, typhoon, pandemics of diseases and similarly, human-made disasters, e.g., incidents of terrorism and industrial accidents [1]–[3]. The number of social media networks and their activity increasing with a high-speed day by day and daily information sharing and user-generated contents is passing hand by hand between millions of internet users [4]. The user-generated content mainly focuses on the daily events and news, which are the current discussed topics in the real world. Internet platforms consider a powerful communication environment between people for information exchanging in a large variety of daily events.

The use of social networking and information sharing in an emergency type of events and dangerous disasters is the research challenge for event detection and tracking it in the early stage. Recently, the extensive connection and increase of social media platforms give the opportunity for the management of crises based on crowd-sourcing. One of the famous tools of crowd-sourcing is Ushahidi [5], which visualize the reports of crowd-sourced it's a perfect example for improving the awareness of various social networks. There are various ways to share information in recent developments, e.g., national security agencies, media outlets, civil defense, etc. The social media potentiality caught the attention through the crisis for higher management quality.

The capability of the limited generalization reason is the level of micro-blogging, which is a changeable topic in terms of abbreviations, informal language, limitation of characters, etc. The recent novel approach proposed by Kruspe *et al.* [6] regarding the Twitter detection based on clustering method and event detection proposed by Fedoryszak *et al.* [7] based on full Twitter firehose demonstrating the contextual

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna Dulizia .

information value by aggregating and sentiment the micro-blog messages.

In terms of security and information privacy in event detection system, there are some requirements that are necessary to follow as authentication, correctness and integrity, privacy, efficiency and non-repudiation. The authentication, verifies the identity of messages through network. The correctness and integrity, checks the data transmission and modification. The privacy, checks the right identity linking process to source of data. Efficiency presents the real-time processing follow the above conditions and non-repudiation clears that during the process the sender can't reject any request. The advantages of this system comparing with the other existing studies is finding the real-news based on the shared information in the social media. This process contains the extraction of the user information regarding the shared post, user location and the geotag. More specifically, the blockchain framework designed to improve the security and transparency of system to avoid sharing wrong information.

Figure 1 presents the overview architecture of crisis event detection in term of supervised and unsupervised learning based on collected data source. The process considering three factors of location, time and content of the shared information in social media and blockchain environment. The data source is related to shared information, posts and comments in social media which is analysed based on twitter features, Geo-tags, NLP features, explicit the time of posts, bag-of-words and custom features.

A. HIGHLIGHTS AND PROBLEM STATEMENT

In this research, we develop the blockchain-based framework using cloud computing and big data techniques for event detection during crisis. To enable data sources, we applied machine learning techniques, gathering and process the information to provide useful insights and decisions in Disaster Risk Management (DRM). The proposed system contains the prediction of hazards, risk assessments, risk mitigation, and clearance. The performed task in this system are:

- **First:** capturing multilingual and multi-modal data dynamically from real-time social media frameworks.
- **Second:** Translating the contents based on the language specific model into unified ontology.
- **Third:** Applying the Artificial Intelligence (AI) and Machine Learning (ML) techniques for language-based intelligent inference.
- **Fourth:** Applying blockchain framework the reliable and secure platform for event detection.
- **Last:** Clarifying results and information of the related stakeholders in an interactive dashboard.

The main concept of this research is to extract the right information which is sharing in social media during crisis by using the recent technologies which gives us the trustworthy and secure information to avoid fake contents and fake users.

The rest of the process is arranged as follows: Section 2 presents the brief literature review of the recent techniques and supports the activities related to crisis and

TABLE 1. Social media usage in cycle of crisis preparation and response.

Social Media Crisis Usage	Phase of Crisis
Disaster warning receive and provide	Preparation
Event detection	Preparation
Learning the shared news information	Preparation-Response
News consuming	Preparation-Response
Actionable information and response to news	Preparation-Response
Expressing the concern of emotions	Response
Spread of awareness	Response

their short-comes and benefits. Section 3 presents the detailed methodology and data collection process and the solutions for end-to-end crisis management and response. Section 4 presents the implementation of the developed framework based on the machine learning techniques and social media post mapping, and we conclude this research in the conclusion section.

II. RELATED WORK

Recently, IoT and blockchain using machine leaning has brought an immense revolution in various walks of life by converging the physical and digital world together, especially in the area of healthcare [8]–[13], navigation [14]–[16], security [17]–[21], cloud computing [22], and smart grid systems [23]. In this section, various social media platforms discussed which extract the information related to crisis for supporting the activities related to disasters. Hui-Jia Li *et al.* [24] proposed the optimization algorithm based on dynamical clustering to get more accurate and fast configuration for the system of electronic commerce. In another approach of this author [25], based on applying the optimization algorithm, they tried to solve the problem of efficient community detection identification. In [26], Hui-Jia Li *et al.* proposed the solution for the problem of epidemic spreading by applying the dynamic approach on signed network.

A. MANAGEMENT CYCLE OF DISASTERS

Disaster events are normally divided into four main parts of preparation, mitigation, response, and recovery. The preparation and mitigation transpire before the effects of disaster and the other two phases after the disaster. The preparation used for reducing the action which is taken the impact of an event. Those events about to happen are taken as a signal from social media and start the preparation process. The taken response from the emergency action through the disaster causes the direct aftermath based on the event of a disaster. Social media is used for the proportionality of active emergency during the response phase, which is conspicuous in ML systems to extract the posts with useful contents related to disasters. Table 1 presents the social media functions for the crisis cycle.
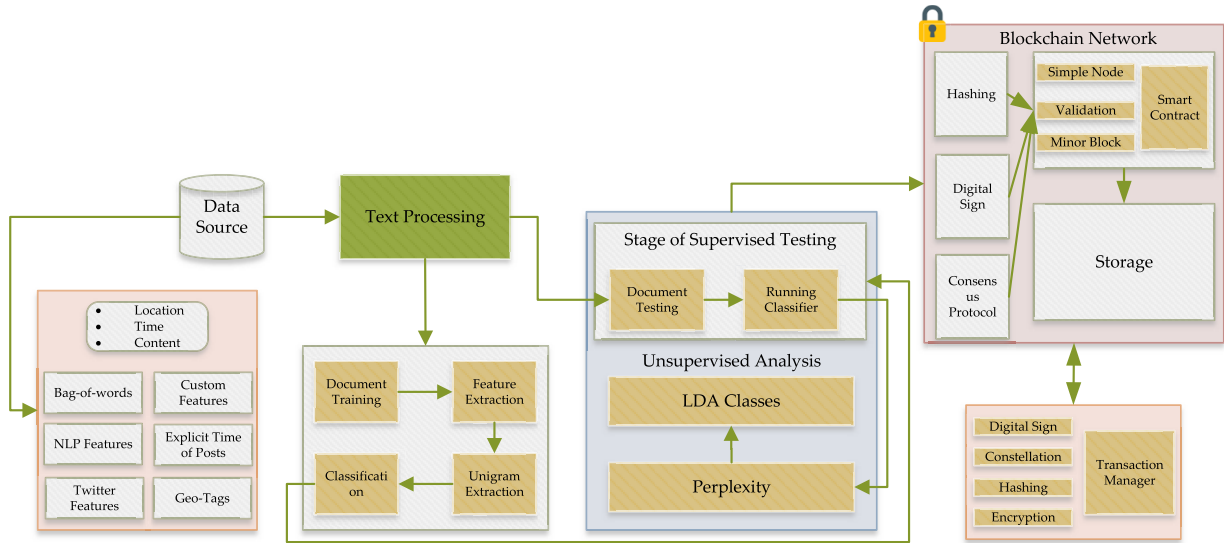
FIGURE 1. Overview of supervised and unsupervised analysis in crisis event detection based on blockchain.

B. SOCIAL MEDIA ACTIONABLE INFORMATION RELATED TO DISASTERS

Social media shared information contains actionable information in terms of sharing the available contents for coordination and decision making [27]. During disasters or some serious topics, the information exchange is a lot between users, but not all the information is right and useful. There are some messages, e.g., advice and caution, utilities, affected people, donations, and needs which are appropriate for support and coordination [28].

- **Advises and caution:** This type of content and posts are giving warning information of the upcoming disaster and some tips for a serious situation
- **Utilities:** The information and contents related to infrastructure damage is also part of the event detection category, which covers the shared news in this area.
- **Donations and needs:** Sharing the contents regarding people and society which are in need in terms of food or medicine, etc., to aware the other people about this problem which this type of posts are the most famous for content sharing between users in social media.
- **Affected people:** Sharing the contents related to those people who are in a trap.

Most of the discussed contents are related to a visual display of crisis-related information in social media based on the thematic, temporal, and spatial aspects for awareness of the situation. The main elements show the various computations between capabilities e.g content extraction regarding special criteria and using Natural Language Processing (NLP) techniques, applying Named Entity Recognition (NER) and other concepts. Some of the social media platforms point to making actionable reports for the relief activity and supporting disaster response. To do this, creating a report requires tagging the pre-defined categories in cloud-source. Similarly, there is a lack of related documents to extract the information for

creating a report for a possible response. Table 7 presents the comparison of the recent event detection approaches based on social media contents.

C. BLOCKCHAIN CONSENSUS MECHANISM IN EVENT DETECTION

Generating the values of the information detected from events contains the need of blockchain framework. This system gives the secure and distributed records for further process. In addition, the blockchain decentralized nature, gives the management trust and consensus mechanism and avoid the authentication problem [35]–[38]. The PoW approach of blockchain depends on the power of computing and evaluating the value of hash comparing with current target. The PoS approach needs to make the block for holding the stake. The one who has the higher number of records has more chance to find the next block. Finally, the PoA approach, depends on the stake identity and the procedure of block depends on trusted nodes which join to the network. This process is also fixed in permissioned blockchain. Ashutosh *et al.* [39] proposed a challenges and opportunities of the 5G-enabled based on integration of blockchain and artificial intelligence. In this process they focus on solving the problem of network scalability based on distributed network of blockchain. Dionysis *et al.* [40] proposed the 5G based assets trading in blockchain network. In this process they used mobile data for the blockchain framework and gave the ability to users for trading, sharing and consume the assets of mobile edge network.

III. PROPOSED EVENT DETECTION SYSTEM BASED ON CRISIS MANAGEMENT AND SOCIAL MEDIA ANALYTICS

The main focus of this system is creating a cloud-based environment for the management of the crisis in social media using social media analytics. The key point of the developed

TABLE 2. Comparison of related researches to event detection.

Author	Proposed Approach	Advantages
Yuwei et al. (2021) [29]	Novel knowledge preserving incremental heterogeneous graph neural network	Extracting the unseen information from shared contents in social media
Habib et al. (2021) [30]	Novel dynamic crowd video modeling for event detection	Improving the performance system based on limited multi-frame data
Heinrich et al. (2021) [31]	Sound event detection based on triple threshold	Improve the performance of segmented localized event level
Lida et al. (2021) [32]	Social media emergency event detection based on similarity between contents	Clustering the textual contents regarding the 3w attributes.
Akash et al. [33] (2021)	Blockchain-based event-driven for the decentralized market place	Effecting on the specifications of event based on the flexible requirement.
Ebtesam et al. [34] (2021)	Blockchain-based event detection from road traffic social sensing	Extracting the validate events without need of prior knowledge from twitter media.
Proposed Approach	Blockchain-based event detection for the shared information related to crisis	Improving the performance of system in term of security and transparency and predictive analysis of textual dataset.

environment is the augmentation of available sensor-based Disaster Risk Management (DRM) with the capability of social media to keep the human sensor in public. This process activates the authority of the related disaster management for integrating and internet-based data access based on applying semantic analysis for action generating and content responses. The collected results can be used to monitor the related emergency and management of disasters, early warning, risk mitigation, and assessments. Figure 2 presents the main architecture of event detection from social media contents. This architecture has four main components: event identification, automatic reasoning, incident monitoring and blockchain. The event identification uses real-time data from social networks. Automatic reasoning extracts the information and knowledge from accessible data using intelligent techniques. Incident monitoring, processes the knowledge-based professional emergency using the sensory interfaces and blockchain framework analyse the security and transparency of system and similarly the proof-of-authority for having the secure and stable system based on trust. Each component presented in detail below.

A. EVENT DETECTION USING NATURAL LANGUAGE PROCESSING

This section presents the main components of the NLP technique used in this process in detail. NLP is one of the famous approaches in terms of knowledge discovery from textual information. Social media information mostly is in terms of posts and tweets that share contents that are happening in the real world about the events happening worldwide. The three main components that are used in this system are defined as below.

1) REPRESENTATION AND IDENTIFICATION OF EVENTS

Event detection can be triggered automatically and manually based on the operator. Data crawling suppose to have some parameters. The requirements of a location-based crawler are social media configuration of the network, window size, and pre-defined area. The location coordinate provides the information related to a location using Google API. There is a need to define specific search terms or pre-defined terms in the database to search the keywords. Based on this process, the crawler searches for the match contents and shared posts with the goal of multi-language contents detection. The language translation service used Google and Microsoft API to translate the contents based on the target language and save them into a knowledge-based special keywords database to reach the defined goal. After setting all the requirements, the system starts to crawl the contents from social media platforms. Every source contains news, posts, images, text, video, location, etc. The crawled information transformed into an appropriate format for further pre-processing and applying semantic analysis. Equation 1, 2 present the number of events that appear in shared content [41]. w is the representation of terms that appear in document t and a is representing the unobserved variable class.

$$X(w, t) = 1 + log_{10} \tag{1}$$

$$X(t, w) = X(t) \sum_a X(w|a)X(a|t) \tag{2}$$

2) AUTOMATIC REASONING

Transforming data in a suitable format and saving it into a database to apply sentiment analysis is the first step of the automatic reasoning process. The automatic reasoning module goes through topic extraction, classification, sentiment analysis, video, and image analysis and finally extracts similar contacts in terms of posts, topics, etc. Content classification cause mapping the information into pre-defined categories. Social media content is changing continuously, and this aspect is not a practical process to explain a disaster. The ontology of disaster will be ready to explain mapping the extracted metadata from social media.

3) VISUALIZATION AND INCIDENT MONITORING

The incident monitoring based on the automatic process and visualizing the crisis from social media shared information required the web-based interface. This interface is
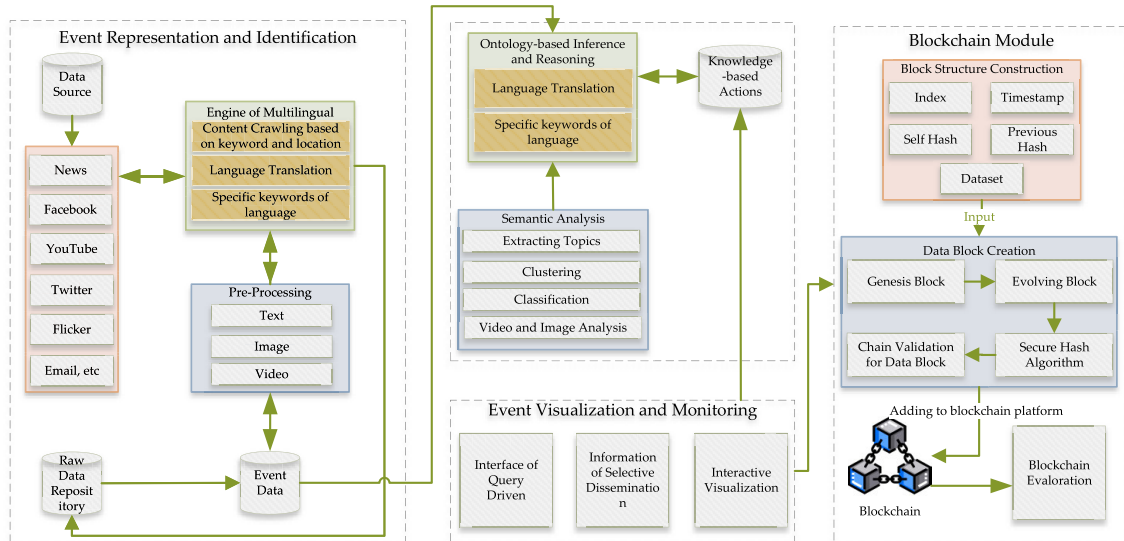
FIGURE 2. Social media analysis architecture for event detection and management of crisis.

divided into information selection and dissemination, interactive visualization and navigation, and the query interface.

- **Information Selection and Dissemination:** Sharing the true information with the real people in social media. Identifying the true contents based on defining sufficient filters in the system.
- **Interactive Visualization:** Developing a general dashboard related to extraction of incidents. Visualizing map, time plot, graphs to show the differences and relationship between events and various incidents.
- **Navigation and Query Interface:** Information filtering based on the detail of event and incident. Able to provide the extra details regarding the disaster and type of damage and location.

B. EVENT DETECTION USING BLOCKCHAIN

In order to achieve to the consensus mechanism, the distributed operation in term of high resilience and tamper-proof, blockchain platform has the power of identification of public keys. The government service in real world needs the identification of government issues which web applications regarding to social media and private email addresses can step forward it. In blockchain platform, the identities and public addresses extracted by using the identification purposes. Smart contracts, using the decentralized manners for running applications based on DLT Virtual Machine using the DTL platform that the user can send message to the network. Figure 3 shows the relationship between the persistent storage and smart contract regarding the event detection in the proposed system. The system designed based on definite and modular aspects to be effective on separating the event detection and event aggregation together. The reports in system can be occasionally and as it is anticipated, regarding the human mobility and observation differences and responding time, the reports of aggregated events only send for the module of

detection that brings the required time for the phase of event detection. The blockchain framework contains two phase of event detection and event aggregation in the proposed system to provide the strong detection process and similarly, identify users and protecting the system. The highest straightforward records regarding the user reports save into blockchain and process later and requires to discarded and disclose the user data.

C. EVENT DETECTION USING DEEP LEARNING

Users of social media are interested in posting their situational information which can be related to the disaster happening around them and the effects of responses for making the better options for decision making. During this process there is importance of posts classification into various categories related to humanitarian for having the efficient processing. After data classification, the dataset become more instructive for applying the specific responses. Various works done by applying the deep learning models such as Convolutional Neural Network (CNN) [42], Gated Recurrent Unit [43] and Long Short Term Memory [44] in term of classification of important contents in critical time period. The main element which make the performance of this systems weak is the input embedding. Most of the existing studies tried to encode the textual data using the package of pre-trained embedding but lots of packages of pre-trained embedding have the fixed parameter and are unidirectional so it will not work for various categories of disasters without doing the process of tuning.

IV. PREDICTIVE ANALYSIS BASED ON EVENT DETECTION

In this section, the predictive analysis on the event detection process is applied to improve the system's performance and check the feed-backs of the process. Similarly, the available
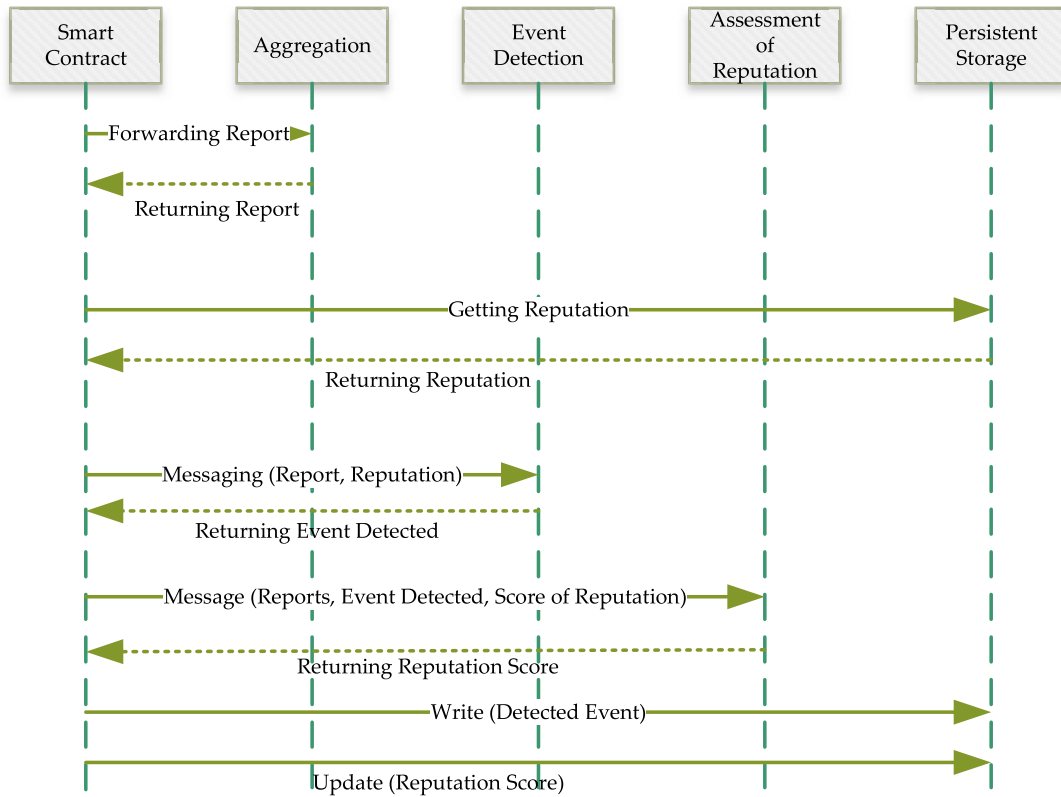
FIGURE 3. Relationship between persistent storage and smart contract in event detection.

dataset for this process was analyzed from every possible aspect.

A. PREDICTION MODEL LEARNING

The predictive model applied in this process is classified into different modules: learning module and prediction algorithm. Normally, the historical data in the prediction model is used for the training set and finding the relationship between words and hidden patterns among the input and output parameters. In the next step, the output of the user input data for the training model is predicted. The prediction model performance depends on some conditions. The training data and input data application scenarios are the same, but non of the prognosis algorithms are not enough for dynamic training of input states. Therefore, we presented the prediction model learning in Figure 4. In this process, to improve the prediction model accuracy, we use the learning module for tuning the prediction algorithm. The presented system monitors the prediction algorithm performance and similarly it depends on the external parameters that are part of learning module. After exploring the external factors and outputs of the prediction model, the learning module has the ability of updating prediction algorithm tunable parameters and to improve the performance it replace the train model to prediction algorithm when it observe the environmental tiggers.

The applied algorithm improves the performance and accuracy of the system based on tuning using a learning module.
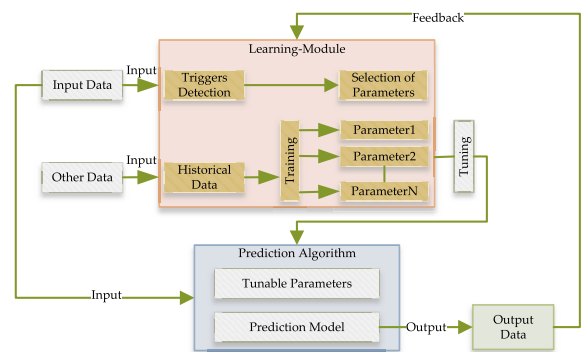

FIGURE 4. Event detection conceptual view for predictive model learning.

The learning module checks the performance of the system continuously based on getting feed-back as output.

V. RESULTS AND IMPLEMENTATION OF THE PROPOSED EVENT DETECTION

This section presents the results of the applied deep learning algorithm of the contents collected from Twitter shared information and platform and analyzed the proposed approach compared with other existing works in this area.

A. DEVELOPMENT ENVIRONMENT AND EXPERIMENTAL EVALUATION

The development environment of implementing the proposed event detection system summarized in Table 3. In total, there

TABLE 3. Development environment.

Module	Component	Description
Predictive Analysis Model	Operating System	Microsoft Windows 10
	CPU	Intel(R) Core(TM) i7-8700@3.20GHz
	Main Memory	16GB RAM
	Core Programming Language	Python
	IDE	PyCharm Professional 2020
	ML Algorithm	Deep Learning
Web Application	Operating System	Windows 10
	Browser	Chrome, Firefox, IE
	Programming Language	HTML, JavaScript
	Library/Framework	Notify.js
Blockchain Framework	Operating System	Ubuntu Linux 18.04 LTS
	Docker Engine	Version 18.06.1-ce
	Docker Composer	Version 1.13.0
	IDE	Composer Playground
	Programming Language	Node.js

are six main component during processing this system as an operating system, which is Microsoft Windows 10, CPU that is Intel(R) Core(TM) i7-8700 @3.20GHz. Used memory in this system is 16GB RAM. The core programming language is python with the IDE of PyCharm Professional 2020 and deep learning model.

B. AVAILABLE EVENT DETECTION DATASET AND ANALYSIS

Social media data collection during a crisis is one of the important aspects of knowledge-based systems for developing a system based on user's needs. Based on the collected records, Twitter contents are the most focused and available information. Table 4 contains the list of available datasets from Twitter contents during crisis and disasters event. There are three definitions for this dataset: type of data, number of tweets, and events.

Figure 5 shows the process of data model for event detection in term of reporting the event, sources and reputation.

Figure 6 represents the comparison of the three dataset categories based on applying seven machine learning algorithms and comparing them with the used deep learning model in this system. The algorithms are Naive Bayes, K-Nearest Neighbour, Support Vector Machine, Logistic Regression, XGBoost, and Deep Learning. Dataset categories are related to COVID19 dataset. As it shown the presented approach is performing good in every data category comparing with other algorithms.

1) CLASSIFICATION OF TWEETS

Social media posts and contents classification into humanitarian categories is important to capture the events and areas. During this process, nine categories were defined for labeling almost 2000 tweets as summarized in Table 5. The defined model train the 90% of the collected shared posts and 10% for testing set. Table 5 shows the number of uneven humanitarian categories of the labeled tweets. The irrelevant category presents the contents which are not related to the mentioned

TABLE 4. Crisis-related dataset records.

| Data	No. Tweets	Events
GeoCov19 [45]	524.000.000 From May 2020	COVID19
covid19_twitter [46]	400.000.000 Mid June 2020	COVID19
Kaggle covid19 [47]	1.500.000 1st May 2020	COVID19
Appen Disaster Response Messages [48]	30.000	Earthquake, Flood, Hurricane
TREC-IS 2019A [49]	30.000	Bombing, Earthquake, Flood, Typhoon, Wildfire, Shooting
Disaster Tweet Corpus 2020 [50]	150.000	Earthquake, Flood, Biological, Hurricane, Wildfires, ...
Florence [51]	600.000	Hurricane
Epic [52]	25.000	Hurricane
Crisis MMD [53]	16.000	Hurricane, Hurricane, Flood, Typhoone, ...
CrisisLexT26 [54]	26.000	Earthquake, Flood, Wildfire, Meteor, Typhoon, ...
Events 2012 [55]	150.000	Disasters and accidents
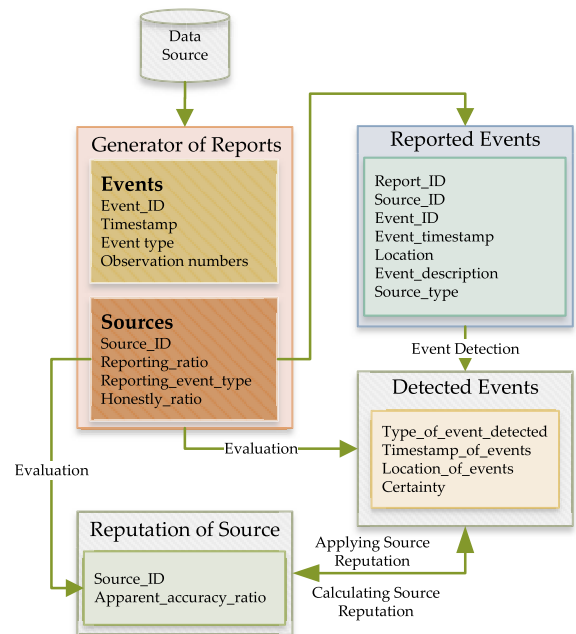
FIGURE 5. Event detection reports and reputation source assessment in the data model.

topic. The authorities category contains the contents related to the government rules. The impact contents are related to the reports which people are getting affected by this disease. The reports present the information of the number of death records and the number of affected cases. The prevention
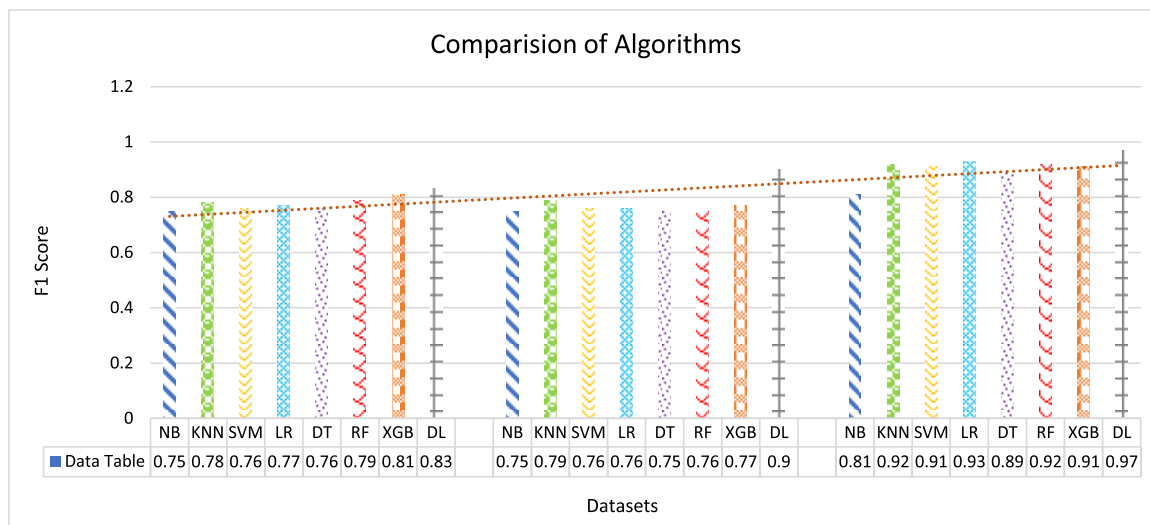
FIGURE 6. Comparative analysis between datasets.

TABLE 5. Train and test labeled dataset detail information.

#	Labeled Tweets	Tested Tweets
Irrelevant	60	3544
Authorities	71	258
Impact	440	744
Reports	250	3350
Prevention	167	2525
Sign\Symptoms	186	1640
Treatment	139	1458
Transmission	83	253
Other information	160	3544
Total	1556	17316

TABLE 6. Confusion matrix for the evaluated event detection.

#	Actual Positive	Actual Negative	Actual Positive	Actual Negative
Predicted Positive	True Positive	False Positive	0.96	0.04
Predicted Negative	False Negative	True Negative	0.02	0.98

records show the strategies regarding this problem, suggestions, and questions related to self-isolation, etc. The sign and symptoms category present all the symptoms: fever, cough, breath problem, etc. The treatment category gives information regarding the treatments of this disease. The transmission category presents the details of disease transmission, and finally, the other information category shows the records of helpful comments and information regarding this problem.

C. PERFORMANCE EVALUATION

F1-Score in this process evaluates based on Equation 3, and Table 6 presents the confusion matrix of the evaluated process based on actual positive and negative values.

$$Precision = \frac{X}{X + Y} \qquad (3)$$

Figure 7 presents the F1-Score of various values regarding the presented event and related shared contents. In total, three groups of shared events were created with five, ten, and twenty shared posts based on geotagging. F1-Score grows

TABLE 7. Simulation parameters of the proposed event detection.

Parameters	Description	Value
Geotagged Posts	Percentage of geotagged posts	[10,15,20] %
Generic Posts	Percentage of generic posts	[10,20,30] %
Distribution of geotagged posts	(PoI, district, street,city)	[0.1,0.1,0.2,0.6],...
Number of events	-	5,10,20
Number of posts	Posts based on topics	[10,50],...
Events radius	Event area radius	[50,100],...
Event distribution	(PoI, district, street, level of city)	[0.6,0.2,0.1,0.1],...

based on increasing the number of posts related to the event. If the set event is ten, the number of captured F1-Score is 0.59 based n using the min and max number of configured posts which shows as [10,50]. Ten is min, and 50 is max. The other side presents the same process based on the posts with the related tags to the event.

Figure 8 shows the records of shared generic posts and tagged posts. This process presents the behavior of multiple events in the same area.

Figure 9 shows the captured results from the changes of geotagged contents in four levels of PoI, district, street, city. The attentiveness of the PoI level contains the higher F1-Score and can create more related categories. The figure shows the distribution of textual information percentage on the right side that evaluates the post coordination among the mentioned four levels. The total process shows that higher accuracy means more estimated PoI levels and a higher possibility of extracting and discovering accurate events.

The applied values in the Figure 7, 8, 9 summarized in Table 7 for further detail.

Acceptable results are created during the data enrichment process, the allows for topic identification with higher accuracy results. Figure 10 shows the process of enrichment data using a few geotagged datasets.
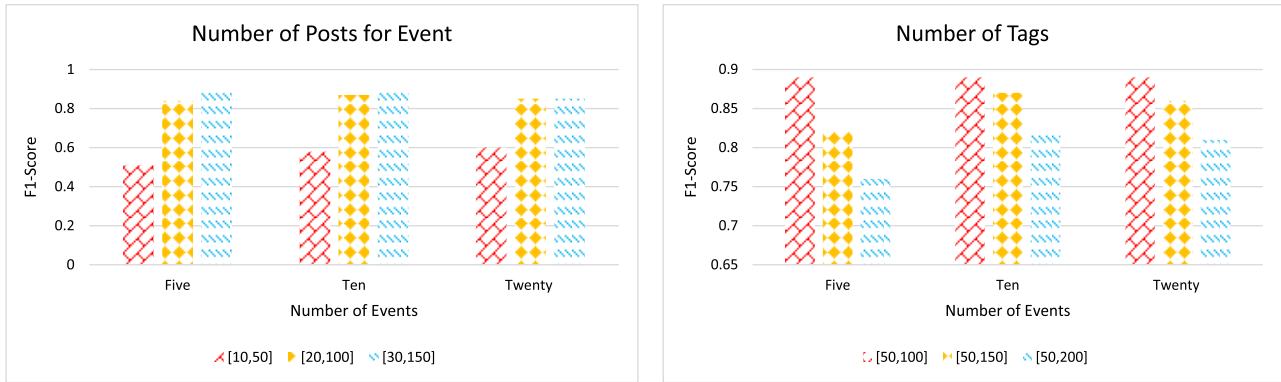
FIGURE 7. Different values F1-score for the number of shared posts related to event.
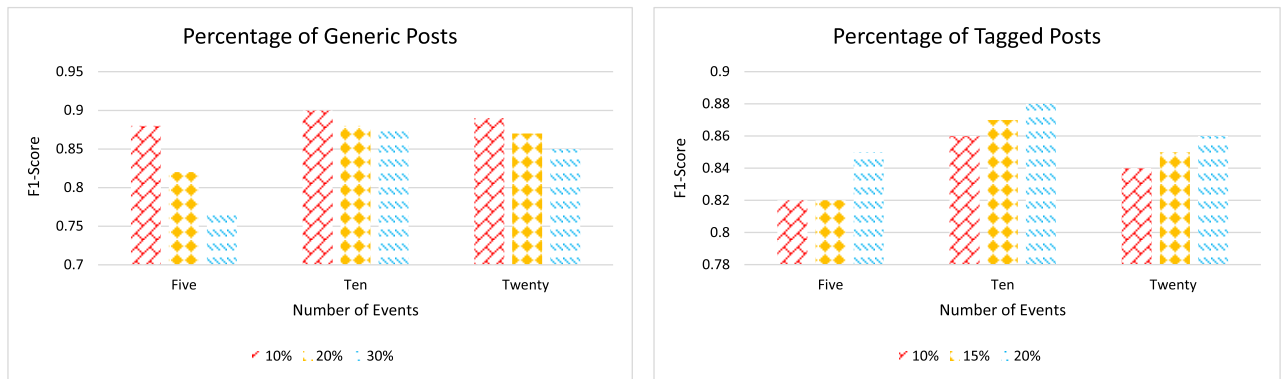

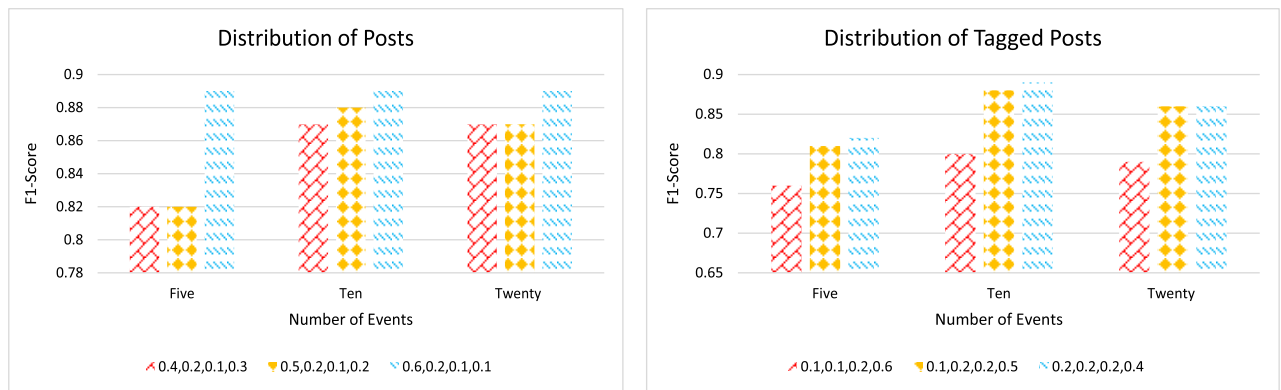**FIGURE 8. Different values of the generic posts percentage.**

FIGURE 9. Different values of the distribution of posts in four levels.

D. BLOCKCHAIN RESULTS

The phase of transaction in blockchain framework, confirms the events and make the procedure more impressive. The transactions in this system are divided into two stages regarding the geographical regions. In the first step, the local blockchain synchronizing is required and in the next step the global blockchain synchronizing which helps to the maintenance of message delivery. Figure 11 shows the successful events rate regarding the impact of threshold value and Figure 12 shows the false event detection rate based on percentage of attackers in blockchain framework.
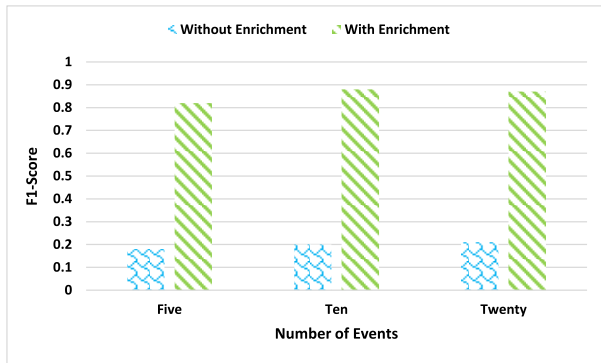
FIGURE 10. F1-Score comparison between the post with enrichment and without enrichment.
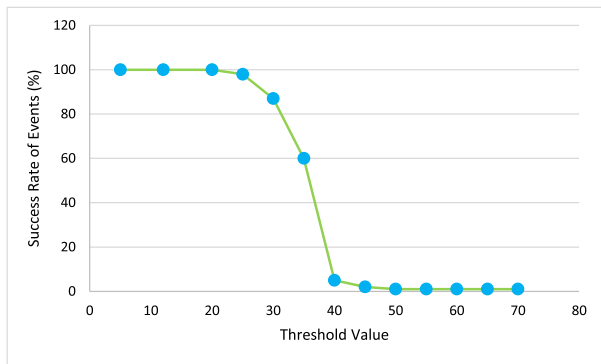

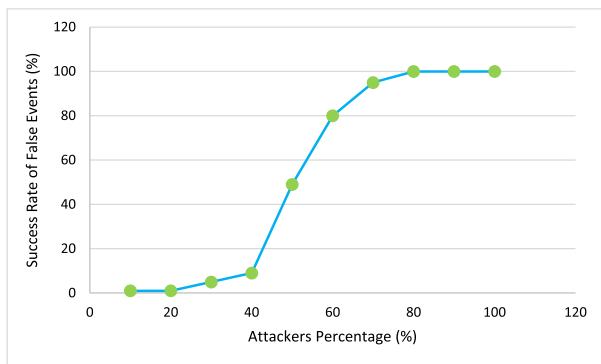**FIGURE 11.** Success event rates based on threshold impact.

FIGURE 12. Success rates of the false events based on attackers percentage.

VI. CONCLUSION

The presented system is designed based on the blockchain and machine learning pipeline to automatically map the crises and disasters with various humanitarian organizations supporting the relief efforts. The defined pipeline is categorized into event detection, classification, mapping the contents using various humanitarian categories, clustering and trust verification. The presented pipelines represent the case study of the shared information on social media and Twitter dataset. The final results are summarized as detecting suitable topics, comparing traditional techniques and recently applied techniques, and predicting and learning modules to improve system performance and avoid sharing the wrong information.

VII. DISCUSSION AND FUTURE WORK

The presented blockchain and machine learning pipeline in this system gives a significant direction for the future research work. We can extend this process to apply for different type of disasters in future in various pipelines. The deficiency in the category of broad humanitarian might weaken the process across the other disasters. Integration of various intelligent techniques detects the awareness of many situations e.g. the areas which are effected from disaster, the shared posts and information and further extra contents can support the system. Data integration from different sources is also the option for increasing the awareness of system.

REFERENCES

[1] P. Williams, "Crisis management," in *Contemporary Strategy*. Evanston, IL, USA: Routledge, 2021, pp. 152–171.

[2] L. Ardito, M. Coccia, and A. M. Petruzzelli, "Technological exaptation and crisis management: Evidence from COVID-19 outbreaks," *RD Manage.*, vol. 51, no. 4, pp. 381–392, Sep. 2021.

[3] J. Abbas, D. Wang, Z. Su, and A. Ziapour, "The role of social media in the advent of COVID-19 pandemic: Crisis management, mental health challenges and implications," *Risk Manag. Healthcare Policy*, vol. 14, p. 1917, May 2021.

[4] S. Wang, Z. Yang, and Y. Chang, "Bringing order to episodes: Mining timeline in social media," *Neurocomputing*, vol. 450, pp. 80–90, Aug. 2021.

[5] S. G. Arapostathis, "A methodology for automatic acquisition of flood-event management information from social media: The flood in Messinia, South Greece, 2016," *Inf. Syst. Frontiers*, vol. 23, pp. 1127–1144, Jan. 2021.

[6] A. Kruspe, J. Kersten, and F. Klan, "Detection of informative tweets in crisis events," *Natural Hazards Earth Syst. Sci.*, 2021. [Online]. Available: https://nhess.copernicus.org/articles/21/1825/2021/

[7] M. Fedoryszak, B. Frederick, V. Rajaram, and C. Zhong, "Real-time event detection on social data streams," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 2774–2782.

[8] F. Jamil, L. Hang, K. Kim, and D. Kim, "A novel medical blockchain model for drug supply chain integrity management in a smart hospital," *Electronics*, vol. 8, p. 505, Apr. 2019.

[9] F. Jamil, S. Ahmad, N. Iqbal, and D.-H. Kim, "Towards a remote monitoring of patient vital signs based on IoT-based blockchain integrity management platforms in smart hospitals," *Sensors*, vol. 20, no. 8, p. 2195, Apr. 2020.

[10] B. Zaabar, O. Cheikhrouhou, F. Jamil, M. Ammi, and M. Abid, "Health-Block: A secure blockchain-based healthcare data management system," *Comput. Netw.*, vol. 200, Dec. 2021, Art. no. 108500.

[11] F. Jamil, F. Qayyum, S. Alhelaly, F. Javed, and A. Muthanna, "Intelligent microservice based on blockchain for healthcare applications," *Comput., Mater. Continua*, vol. 69, no. 2, pp. 2513–2530, 2021.

[12] F. Jamil, H. K. Kahng, S. Kim, and D.-H. Kim, "Towards secure fitness framework based on IoT-enabled blockchain network integrated with machine learning algorithms," *Sensors*, vol. 21, no. 5, p. 1640, Feb. 2021.

[13] I. Jamal, Z. Ghaffar, A. Alshahrani, M. Fayaz, A. M. Alghamdi, and J. Gwak, "A topical review on machine learning, software defined networking, Internet of Things applications: Research limitations and challenges," *Electronics*, vol. 10, no. 8, p. 880, Apr. 2021.

[14] F. Jamil and D. Kim, "Enhanced Kalman filter algorithm using fuzzy inference for improving position estimation in indoor navigation," *J. Intell. Fuzzy Syst.*, vol. 40, no. 5, pp. 8991–9005, Apr. 2021.

[15] F. Jamil, O. Cheikhrouhou, H. Jamil, A. Koubaa, A. Derhab, and M. A. Ferrag, "PetroBlock: A blockchain-based payment mechanism for fueling smart vehicles," *Appl. Sci.*, vol. 11, no. 7, p. 3055, Mar. 2021.

[16] F. Jamil and D. H. Kim, "Improving accuracy of the alpha–beta filter algorithm using an ANN-based learning mechanism in indoor navigation system," *Sensors*, vol. 19, no. 18, p. 3946, 2019.

[17] M. H. Bin Waheed, F. Jamil, A. Qayyum, H. Jamil, O. Cheikhrouhou, M. Ibrahim, B. Bhushan, and H. Hmam, "A new efficient architecture for adaptive bit-rate video streaming," *Sustainability*, vol. 13, no. 8, p. 4541, Apr. 2021.

[18] I. Jamal, F. Jamil, and D. Kim, "An ensemble of prediction and learning mechanism for improving accuracy of anomaly detection in network intrusion environments," *Sustainability*, vol. 13, no. 18, p. 10057, Sep. 2021.

[19] F. Jamil and D. Kim, "Payment mechanism for electronic charging using blockchain in smart vehicle," *Korea*, vol. 30, p. 31, May 2019.

[20] I. Jamal, N. Iqbal, S. Ahmad, and D. H. Kim, "Towards mountain fire safety using fire spread predictive analytics and mountain fire containment in IoT environment," *Sustainability*, vol. 13, no. 5, p. 2461, Feb. 2021.

[21] I. Jamal, S. Ahmad, and D. H. Kim, "Quantum GIS based descriptive and predictive data analysis for effective planning of waste management," *IEEE Access*, vol. 8, pp. 46193–46205, 2020.

[22] A. Ali, M. M. Iqbal, H. Jamil, F. Qayyum, S. Jabbar, O. Cheikhrouhou, M. Baz, and F. Jamil, "An efficient dynamic-decision based task scheduler for task offloading optimization and energy management in mobile cloud computing," *Sensors*, vol. 21, no. 13, p. 4527, Jul. 2021.

[23] S. Ahmad, I. Ullah, F. Jamil, and D. Kim, "Toward the optimal operation of hybrid renewable energy resources in microgrids," *Energies*, vol. 13, no. 20, p. 5482, Oct. 2020.

[24] H. J. Li, Z. Bu, Z. Wang, and J. Cao, "Dynamical clustering in electronic commerce systems via optimization and leadership expansion," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5327–5334, Aug. 2020.

[25] H.-J. Li, L. Wang, Y. Zhang, and M. Perc, "Optimization of identifiability for efficient community detection," *New J. Phys.*, vol. 22, no. 6, Jun. 2020, Art. no. 063035.

[26] H.-J. Li, W. Xu, S. Song, W.-X. Wang, and M. Perc, "The dynamics of epidemic spreading on signed networks," *Chaos, Solitons Fractals*, vol. 151, Oct. 2021, Art. no. 111294.

[27] R. McCreadie, C. Buntain, and I. Soboroff, "TREC incident streams: Finding actionable information on social media," in *Proc. Int. Conf. Inf. Syst. Crisis Response Manage. (ISCRAM)*, Valencia, Spain, May 2019, pp. 691–705.

[28] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: Survey summary," in *Proc. Companion Web Conf. Web Conf.*, 2018, pp. 507–511.

[29] Y. Cao, H. Peng, J. Wu, Y. Dou, J. Li, and P. S. Yu, "Knowledge-preserving incremental social event detection via heterogeneous GNNs," in *Proc. Web Conf.*, Apr. 2021, pp. 3383–3395.

[30] H. Ullah, I. U. Islam, M. Ullah, M. Afaq, S. D. Khan, and J. Iqbal, "Multi-feature-based crowd video modeling for visual event detection," *Multimedia Syst.*, vol. 27, no. 4, pp. 589–597, Aug. 2021.

[31] H. Dinkel, M. Wu, and K. Yu, "Towards duration robust weakly supervised sound event detection," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 887–900, 2021.

[32] L. Huang, G. Liu, T. Chen, H. Yuan, P. Shi, and Y. Miao, "Similarity-based emergency event detection in social media," *J. Saf. Sci. Resilience*, vol. 2, no. 1, pp. 11–19, Mar. 2021.

[33] A. Pateria and K. Anyanwu, "Towards event-driven decentralized marketplaces on the blockchain," in *Proc. 15th ACM Int. Conf. Distrib. Event-Based Syst.*, Jun. 2021, pp. 43–54.

[34] E. Alomari, I. Katib, A. Albeshri, T. Yigitcanlar, and R. Mehmood, "Iktishaf+: A big data tool with automatic labeling for road traffic social sensing and event detection using distributed machine learning," *Sensors*, vol. 21, no. 9, p. 2993, Apr. 2021.

[35] K. Wüst st and A. Gervais, "Do you need a blockchain?" in *Proc. Crypto Valley Conf. Blockchain Technol. (CVCBT)*, Jun. 2018, pp. 45–54.

[36] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1495–1505, Apr. 2019.

[37] Z. Shahbazi and Y.-C. Byun, "Fake media detection based on natural language processing and blockchain approaches," *IEEE Access*, vol. 9, pp. 128442–128453, 2021.

[38] Z. Shahbazi and Y.-C. Byun, "A framework of vehicular security and demand service prediction based on data analysis integrated with blockchain approach," *Sensors*, vol. 21, no. 10, p. 3314, May 2021.

[39] A. D. Dwivedi, R. Singh, K. Kaushik, R. R. Mukkamala, and W. S. Alnumay, "Blockchain and artificial intelligence for 5G-enabled Internet of Things: Challenges, opportunities, and solutions," *Trans. Emerg. Telecommun. Technol.*, Jul. 2021.

[40] D. Xenakis, A. Tsiota, C.-T. Koulis, C. Xenakis, and N. Passas, "Contract-less mobile data access beyond 5G: Fully-decentralized, high-throughput and anonymous asset trading over the blockchain," *IEEE Access*, vol. 9, pp. 73963–74016, 2021.

[41] S. Khatoon, M. A. Alshamari, A. Asif, M. M. Hasan, S. Abdou, K. M. Elsayed, and M. Rashwan, "Development of social media analytics system for emergency event detection and crisis management," *Comput., Mater. Continua*, vol. 68, no. 3, pp. 3079–3100, 2021.

[42] X. Huang, C. Wang, Z. Li, and H. Ning, "A visual–textual fused approach to automated tagging of flood-related tweets during a flood event," *Int. J. Digit. Earth*, vol. 12, no. 11, pp. 1248–1264, Sep. 2018.

[43] A. Cossu, A. Carta, V. Lomonaco, and D. Bacciu, "Continual learning for recurrent neural networks: An empirical evaluation," 2021, *arXiv:2103.07492*.

[44] S. Madichetty and M. Sridevi, "A neural-based approach for detecting the situational information from Twitter during disaster," *IEEE Trans. Comput. Social Syst.*, vol. 8, no. 4, pp. 870–880, Aug. 2021.

[45] U. Qazi, M. Imran, and F. Ofli, "GeoCoV19: A dataset of hundreds of millions of multilingual COVID-19 tweets with location information," *SIGSPATIAL Special*, vol. 12, no. 1, pp. 6–15, Jun. 2020.

[46] J. M. Banda, R. Tekumalla, G. Wang, J. Yu, T. Liu, Y. Ding, E. Artemova, E. Tutubalina, and G. Chowell, "A large-scale COVID-19 Twitter chatter dataset for open scientific research—An international collaboration," *Epidemiologia*, vol. 2, no. 3, pp. 315–324, Aug. 2021.

[47] S. Smith. (2020). *Coronavirus (COVID19) Tweets-Early April*. [Online]. Available: https://kaggle.com

[48] A. Kruspe, J. Kersten, and F. Klan, "Detection of actionable tweets in crisis events," *Natural Hazards Earth Syst. Sci.*, vol. 21, no. 6, pp. 1825–1845, Jun. 2021.

[49] J. Kersten, A. Kruspe, M. Wiegmann, and F. Klan, "Robust filtering of crisis-related tweets," in *Proc. 16th Int. Conf. Inf. Syst. Crisis Response Manage. (ISCRAM)*, 2019, pp. 1–11.

[50] J. Kersten and F. Klan, "What happens where during disasters? A workflow for the multifaceted characterization of crisis events based on Twitter data," *J. Contingencies Crisis Manage.*, vol. 28, no. 3, pp. 262–280, Sep. 2020.

[51] K. Stowe, M. Palmer, J. Anderson, M. Kogan, L. Palen, K. M. Anderson, R. Morss, J. Demuth, and H. Lazrus, "Developing and evaluating annotation procedures for Twitter data during hazard events," in *Proc. Joint Workshop Linguistic Annotation, Multiword Expressions Construct. (LAW-MWE-CxG)*, 2018, pp. 133–143.

[52] F. Alam, S. Joty, and M. Imran, "Domain adaptation with adversarial training and graph embeddings," 2018, *arXiv:1805.05151*.

[53] A. Olteanu, C. Castillo, F. Diaz, and S. Vieweg, "CrisisLex: A lexicon for collecting and filtering microblogged communications in crises," in *Proc. 8th Int. AAAI Conf. Weblogs Social Media*, 2014, pp. 1–10.

[54] A. J. Mcminn, Y. Moshfeghi, and J. M. Jose, "Building a large-scale corpus for evaluating event detection on Twitter," in *Proc. 22nd ACM Int. Conf. Conf. Inf. Knowl. Manage.*, 2013, pp. 409–418.

ZEINAB SHAHBAZI received the B.S. degree in software engineering from Pooyesh University, Iran. In March 2017, she moved to the Republic of Korea for M.S. studies and started working with the Internet Laboratory, Chonbuk National University (CBNU). After completing her master's in 2018, she moved to Jeju-do, in March 2019, and started working as a Ph.D. Research Fellow with the Machine Learning Laboratory (MLL), Jeju National University. Her research interests include artificial intelligence and machine learning, natural language processing, deep learning, knowledge discovery, data mining, and blockchain.

YUNG-CHEOL BYUN received the B.S. degree from Jeju National University, in 1993, and the M.S. and Ph.D. degrees from Yonsei University, in 1995 and 2001, respectively. He worked as a Special Lecturer with Samsung Electronics, in 2000 and 2001. From 2001 to 2003, he was a Senior Researcher with the Electronics and Telecommunications Research Institute (ETRI). He was an Assistant Professor with Jeju National University, in 2003, where he is currently an Associate Professor with the Computer Engineering Department. His research interests include AI and machine learning, pattern recognition, blockchain and deep learning-based applications, big data and knowledge discovery, time series data analysis and prediction, image processing and medical applications, and recommendation systems.

• • •