Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 3185391, 19 pages
https://doi.org/10.1155/2023/3185391

Research Article

Effectual Seed Pick Framework Focusing on Maximizing Influence in Social Networks

Sonia Choudhary[1], **Kapil Sharma**[1], **and Monika Bajaj**[2]

[1]*Delhi Technological University, New Delhi, India*
[2]*University of Delhi, New Delhi, India*

Correspondence should be addressed to Kapil Sharma; kapil@ieee.org

Received 10 July 2022; Accepted 6 September 2022; Published 13 April 2023

Academic Editor: Akshi Kumar

Today, the emergence of social media is helpful for the healthcare system where everyone is closely connected. Large numbers of people can be reached by using seed nodes to provide medical advice, facilities, new changes in the treatment, and any health ministry guidelines. As today's world is dealing with COVID-19, the main objective is to provide healthcare services to many people irrespective of time and locality. As people suffering from corona are dealing with mental health issues, in order to deal with it, a seed pick framework using machine learning for the influence maximization technique is proposed, which will be helpful to provide pervasive healthcare. For pervasive healthcare, an effectual seed pick framework is required focusing on influence maximization using machine learning. The proposed algorithm Fuzzy-VIKOR is helpful to identify the targeted node to spread information at a high rate. Consequently, the proposed structure effectively addresses different issues related to a large number of patients, and thus, increased influence maximization using seed nodes is helpful for pervasive healthcare. The experiments show that the proposed framework has high precision, accuracy, F1-score, and recall compared to other existing algorithms employed to find influence maximization seeds.

1. Introduction

Public mass media has become a vital forum for entities to newscast their strategies, businesses to promote their merchandises, and individuals en route to spread their viewpoints. Information technology is a powerful medium that influences decision-making. Marketers can outperform their competitors if they have access to timely information. Advances in telephone networks, mobile phones, laptops, online services, and the Internet allow people to acquire information more quickly and easily [1]. As the information technology has evolved, it has provided a platform for internet marketing; whatever happens in the digital world is monitored to find out the public opinion. Reference [2] has shown how the marketing concept got influenced by the behavioural patterns. Almost everyone has personal social media accounts that provide access to both private and public information. Social media (Facebook, YouTube, Twitter, Instagram, and so on) can be used to distribute information regarding products and services. In general, social media has a communicative nature; the appropriate strategy is where users of this service can supply as much detail to users as feasible. Furthermore, the quick response makes social media significantly helpful in promoting services and products [3]. Through social media platforms, healthcare professionals can connect with patients and the public, discuss topics related to their care, and improve their health habits [4–8]. Social media can help healthcare professionals improve patient outcomes and enhance their personal network [9]. Online groups allow physicians to connect with other doctors and hear about new medical developments [10]. They can also discuss their practice management issues and develop effective marketing techniques [11]. With the rise of social networking, medical professionals have created specialized communities where members are guarded against the general public. These networks are usually secret and are often only used by members of highly specialized professions [12]. HCPs can also

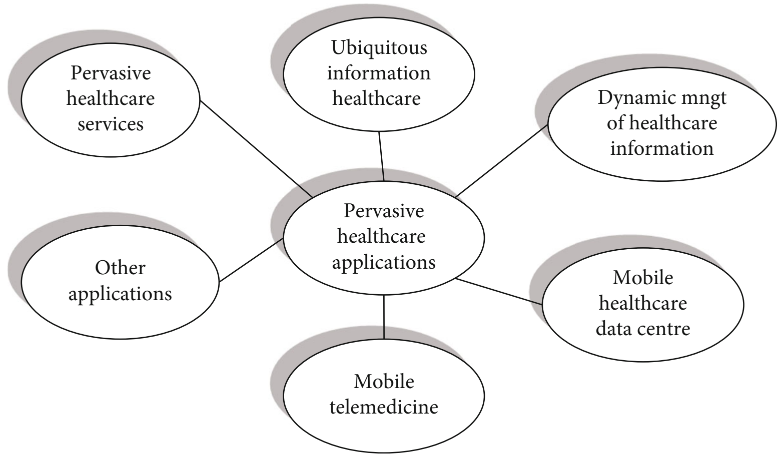

Figure 1: Pervasive healthcare applications.

communicate using "general purpose" social media sites such as Facebook, Twitter, and LinkedIn. Hospitals, health systems, and professional groups use social media to interact with patients and the general public. They are also promoting their numerous services and products through it [13]. Communication with the patients and the community is an important aspect of any healthcare facility. Establishing a venue for communicating with the public and the media is also an example of a healthcare facility's use [11, 12]. Many hospitals and medical facilities also use blogs [14]. Increasing the number of people using social media can influence public health goals and habits.

The World Health Organization pronounced the flare-up of SARS-CoV-2 as a worldwide danger. The incident highlights the importance of information dissemination in an epidemic [15, 16]. The phrase infodemic [17, 18] was used to describe the dissemination of misinformation during an epidemic. It was used to illustrate the risks of misinformation during the control of disease outbreaks [19–21], because it can hasten the spread of an epidemic by influencing and fragmenting social responses [22]. Understanding how people obtain and avoid information is a basic research topic [23]. It can be observed in the COVID-19 outbreak, which exhibited the disintermediated spread of information. The transmission of knowledge can affect how people behave and how government agencies respond to them. For instance, the transmission of a virus can affect how people respond to public health measures and how effective these are [24, 25]. During the COVID-19 episode, the transmission of data dispersed by the infection impacted the manner in which individuals acted. The rise of social media platforms has given users access to large volumes of content and has the potential to spread false information and rumours. Algorithms mediate and facilitate content promotion and information transmission by taking into account user interests and attitudes [26]. This is a considerable departure from the typical news paradigm, which greatly impacts social perception formation [27]. High polarisation, according to a study, can lead to the spread of disinformation [28, 29]. According to certain studies, fake news and false information may travel faster and further than fact-based news [30]. Because political discourse frequently labels opposing news as phony or untrustworthy, this could be a platform-specific effect [31]. The impact of social media on the perception of tough topics is also being explored in COVID-19. A novel technique has been proposed to determine the truthfulness of the information shared through social media. Identifying a seed group of k users who can maximize the propagation of influence over a social network is stated as the job of influence maximization (IM) [32, 33]. Furthermore, to find the influencer in a community to maximize the right information in the social media viral marketing, a novel method has to be developed.

Pervasive healthcare is both a solution to a range of problems and a vision for the future of healthcare. Pervasive healthcare addresses the challenge of geographic barriers for everyone, everywhere, and anytime. Pervasive healthcare has expanded both the number of individuals treated and the relevance of healthcare services. The main healthcare services include preventive and maintenance care, periodic health investigations or checkups, home healthcare monitoring (short term monitoring), nursing home monitoring (long term monitoring), individual healthcare monitoring, incidence detection and management, and emergency intervention, transportation, and treatment. These objectives could be met with universal, efficient, and dependable access to healthcare services, providers, and biomedical information available at all times. Some of the applications of pervasive healthcare as in Figure 1 are listed below:

(i) *Pervasive healthcare services*: the services include monitoring the patient anytime, irrespective of the patient's location

(ii) *Ubiquitous healthcare information*: it includes access of patient-related information at any time by the healthcare provider or the patient

(iii) *Dynamic management of healthcare information*: it restricts patient-related information access. It includes the privacy- and secrecy-related issues of the patient data or patient-related information stored by the healthcare provider

(iv) *Mobile healthcare data centre*: this application is used for a large amount of data (healthcare data). This application helps access, store, analyse, and update healthcare data. Mobile decision makers are provided the healthcare data to make decision for healthcare. Pharmaceutical companies generally use this data for the analysis and for reporting any problem as soon as possible related to any new drug launched by them in the market. This data is also helpful in predicting demands for any drug in the market for the current as well as future demands. Through these data accessions, different healthcare research can be conducted on the patients without knowing them by the healthcare researchers. Different researches can be carried out using this data related to the use of a new drug on patients, new treatments on the patients, and their survival rates. The patient data is kept secure by not allowing disclosure in any format

(v) *Mobile telemedicine*: mobile telemedicine has made it possible to reach a large number of patients irrespective of their location. It has covered a large number of patients for treatment

(vi) *Other applications*: the novel coronavirus (COVID-19) outbreak these days is a natural calamity for which humans are trying to fight, and pervasive healthcare is the significant solution, as it has affected the whole world as a pandemic by making sick a large number of people. For the outbreak of pandemic, the researchers have found out that over-population, globalization, and hyperconnectivity are the main factors. Through pervasive healthcare, a number of patients can be reached simultaneously to provide new corona treatments and guidelines by the health ministry and manage the COVID-19 crisis

As a result, in today's world, where everyone is connected through social media; social media and the data gathered through it play a crucial part in providing pervasive healthcare. The relationships between social network users are studied using social network analysis. A social network can be significantly used to find the relationship-related attributes between social network users and the extent to which they influence outcomes related to health.

SNA is being researched to see whether it can be used to analyse various aspects of pandemics. Important definitions for social network analysis include the following:

Degree centrality: in a social network, it measures the number of links or connections a node has with the other social network nodes or social media user. It is an integer value or count.

Indegree: this is the value of total incoming links or connection of a node from source nodes.

Outdegree: the number of connections or links between the targeted node and a source node is called outdegree. It also counts the number of secondary cases infected by a single sick patient.

Betweenness: it is a metric that shows how many times a node appears on the shortest path between two other nodes. It is an important measure for demonstrating a patient's participation in bridging a path between patients without direct interaction.

Closeness centrality: the total of the shortest path lengths from one node to every other in the social network is known as closeness centrality. The inverted sum of the path lengths from a node to every other node in the network is called closeness centrality. Harmonic closeness can also be used to calculate closeness centrality because of unconnected nodes in the network.

Edge betweenness: edge betweenness is a measure of the shortest paths that pass through an edge that links two parts of a social network, which are vital for communication between two users or a social network node.

Clustering coefficient: it is the count of the degree to which nodes in a graph are liable to cluster together.

Network diameter: it is the shortest path between the two most distant social network users.

Super-spreader: any node with an outdegree $\geq a$ threshold value can be operationally defined as a super-spreader.

So the significant contributions of this research paper are as follows:

(1) It provide the effectual seed for providing pervasive healthcare

(2) Through opinion leader, pervasive healthcare for the users having negative impact due to corona can be provided healthcare for positive mental health

(3) Effectual seed pick framework helps to find the opinion leaders helpful for providing medical healthcare advice to positively influence other people in the network

The main challenges to find effectual seed pick for content based dataset is very difficult to find out. In this manuscript, we will overcome this problem by using Exact Extract Algorithm (EEA); then, the proposed algorithm Fuzzy-VIKOR is helpful to identify the targeted node to spread information at a high rate. As in this section, the introduction part is discussed. The rest of the manuscript is organised as the related research work is discussed in the subsequent section. After that, in Section 3, the methodology and framework are discussed for the effectual seed pick for the influence maximization in the social network for providing pervasive healthcare.

2. Related Work

Opinion leader's/influencer's qualities make them eminent members of their community and differentiate them from the other members. Influencer influence consists of (a) the influencer's principles and qualities, (b) the influencer's proficiency or capability, and (c) the influencer's social position [34, 35]. The social position of a node or user in a social network is the topic of this research. The significance of social networks in determining the diffusion of ideas and the adoption of new practices cannot be overstated. The success of an opinion leader depends on the people who know them and how many people the opinion leader knows and accessibility of the opinion leader [35]. A node's centrality value is very important for identifying an opinion leader like the degree of a node in the social network, betweenness, and closeness centrality [36]. The nodes connected to the appropriate number of people are best suited for the influence maximization in the social network of a new idea or service. For viral marketing, nodes are selected as opinion leaders who are well connected in the social network. The selection of opinion leader depends on different methods of selection of opinion leader and the degree through which they identify nodes for different roles [37, 38]. The sociometric methods can be used to identify opinion leader by interviewing different users in the social network by asking different question like who they will seek for advice; this gives a good measure of opinion leader position in the social network [39]. The most reliable method to identify the opinion leader was "whole community" which was not prone to bias and was helpful to nominate a more credible opinion leader in the society as compared to other methods [39, 35]. The process of opinion leader selection is expensive and time consuming. So the main stress is on the low-cost method, which is simpler and less time consuming through staff selection [40], self-selection [41], or the use of both [42]. In diffusion influence maximization, importance is given to the position of the opinion leader and their role in the social network. Diverse methods are employed to classify influencers in the field of marketing and sociology [43, 44] [36, 45].

Different strategies can be employed to categorize opinion leaders. However, there are not enough studies on opinion leaders in the medical field. It has not been investigated how different methodologies are utilised to classify opinion leaders in the healthcare arena based on their social strategic position. Using many methodologies, including machine learning, social network analysis is one way to discover opinion leaders. The degree is a measurement of how well a node connects to other nodes. The degree of a node provides information about a node's strategic location in the social network. Different measurements of centrality provide different information about the nodes or users in a social network. These metrics are useful in identifying the social network's opinion leaders.

The centrality [46] of a node/user in a social network is determined by the position of the node/user in relation to other nodes/users in the network. One of the most essential factors in determining an opinion leader in a social network is the centrality rating. The number of users/nodes adjacent to a node/user is the degree of that node/user. Degree centrality is a useful metric for determining a user's or node's popularity or impact in a social network. Outdegree is the number of connections or links created by a node/user to other nodes/users (a count which tells about the influence of a node). In a social network, the number of connections or links from other nodes/users to a node indicates the information received by other nodes/users, and the term for that number of connections is indegree. The proportion of times node/user x requires node y to reach z through the shortest path is known as betweenness centrality Betweenness centrality, which represents a node's influence in the network, is the number of times nodes in a social network rely on x to connect other nodes. In a social network, a node's proximity centrality reflects how near that node/user is to other nodes/users in the network. Due to short communication channels to other nodes, users with high proximity centrality are good at distributing information about services or products to other nodes in the network.

Valente and Pumpuang [35] discuss different techniques to find an opinion leader in a network and discuss their pros and cons. In particular, 10 techniques to identify an opinion leader are used like celebrities, self-selection, staff selection, self-identification, positional approach, judge's rating, snowball method, expert selection, and sociometric and social sociometric.

Holliday et al. [47] have discussed about the selection of opinion leader for the intervention in unusual behaviour and the outcomes of those interventions in assisting smoking prevention. Centrality measures are employed to categorize the opinion leader in a social network like degree centrality, closeness centrality, betweenness, and mean distance. So the opinion leader identified through different centrality measures assists in smoking prevention. The method used is less costly and less time consuming and effective to assist smoking prevention.

This section explains about the influence of social media in healthcare and the maximation methods done in the existing works. Fatema and Lariscy [48] demonstrate media exposure that is positively associated with the usage of primary care. Organizations should explore using media to disseminate information about maternal health. In Smaldone et al. [49], the findings of the article show a significant relationship between a few aspects of patient strengthening and the distorted side of media, demonstrating a positive effect on patient experience change, but also conceivable critical dangers and issues due to distortion of online information featuring the distorted side of media. Yan et al. [50] explained that an individuals' acquaintance with emerging infectious illnesses can be altered through media inclusion, influencing the general population's perceptions and practices.

The impact of the media on the propagation of the COVID-19 flare-up, however, is a major general medical issue. Kempe et al. [51] suggested a greedy approach for selection of k influential nodes which depends on hill-climbing.

Furthermore, the greedy technique approximates more than 63% influence maximization for selected seeds.

The cost-efficient lazy forward approach (CELF), proposed by Leskovec et al. [52], is a further extension of a greedy algorithm. To enhance the CELF, Goyal et al. [53] derived an extension of CELF called CELF++, which greatly reduces the number of repetitions. Guo et al. [54] explored a cluster-based technique that divides the entire network into

Table 1: Related research on opinion leader in healthcare sector.

S no.	Author & year	Dataset	Methodology used	Pros	Cons
1	Mohamad et al. 2017 [62]	Instagram data	Centrality measures are used, mainly degree	Influence people regarding health issues	They have studied only the impact of opinion leader on local issues.
2	Chew, Mohammad, and Salleh 2019 [63]	545 articles from different journals related to healthcare	They have discussed the approach of parasocial opinion leader in health sector	Influences follower in terms of opinion, emotion, and action	New methodology needs to be discussed.
3	Saraswathi et al. 2020 [64]	Contact tracing data from daily health department bulletins released by Karnataka Govt.	Used centrality measures like degree, indegree, outdegree, mean distance	Outbreak of corona in Karnataka is studied using social network analysis.	Find out super spreaders of disease using centrality measures. Cannot reflect universally field reality
4	Hossain 2020 [65]	Twitter dataset	NLP systems used	Dataset for COVID is analysed to find misleading information.	Other dataset needs to be discussed from other social media platform.
5	Maheshwari and Albert 2020 [66]	10,000-node network with 5 infected nodes created	Social network analysis is used to detect the outbreak of COVID	Using opinion leader concept to find infecting node in the network	Impact of social distancing on spread is considered here only.
6	Nagarajan et al. 2020 [67]	Contact tracing data of SARS-CoV-2 patients in the Indian state of Karnataka	Find spreader of disease using social network analysis and different centrality measures	Spread of disease is studied using social network analysis.	Used opinion leader concept to find spreader
7	Karaivanov 2020 [68]	Constructed dynamic social network model for COVID	Embed SIR epidemiology model	It predicts the peak of infection based on lockdown.	Work is not on real data it simulated
8	Wong et al. 2020 [69]	Twitter, Facebook, LinkedIn	Social media use in medical domain	It has provided a large number of citations to research articles.	Confidentiality and security of patient data are point of concern.
9	Zaplotnik, Gavrić, and Medic 2020 [70]	Constructed a social network of 2 million nodes representing inhabitants of Slovenia	SIR model is used	Forecast of COVID progression is compared to real data of Slovenia.	For forecast of COVID, a real virus transmission model with accurate social network models is required.
10	Hung et al. 2020 [71]	Twitter data related to COVID-19	Applied machine learning models to find out the sentiments of users of social network	Outcome was regarding the positive, neutral, and negative sentiments of tweets.	Requires constant monitoring of tweets as real-time posted tweets are monitored
11	Pandey, Misra et al. 2021 [72]	Data collected by Google Forms	Analysis of data collected to find the connections of techno experts	Qualitative study explored the mitigating factors of COVID-19.	It is a time-consuming process.
12	Pandey, Astha et al. 2021 [73]	Data from 38 persons was collected	Qualitative study of data related to COVID survivors	Thematic analysis comes with four conclusions.	It cannot be applied on a vast scale.
13	Azad and Devi 2020 [74]	Travel history of infected patients from 30 January to 6 April 6, 2020	Social network analysis, eigenvector centrality and degree was used	Transmission of COVID tracing is analysed.	It is not a generic approach.
14	Jo et al. 2021 [75]	Contact tracing information of 3283 patients in Seoul	Social network analysis, different centrality measures were analysed	Root of infection in South Korea is traced.	It provide limited results on real data.
15	Tsao et al. 2021 [76]	Studies related to COVID-19 from	The machine learning approach is discussed for COVID-19 data from	Social media can be used for surveillance	Different research papers are analysed to get an overview

Table 1: Continued.

S no.	Author & year	Dataset	Methodology used	Pros	Cons
		November 2019 to November 2020	social media. Latent Dirichlet analysis and random forest is used.	and monitoring, as disease controller.	of COVID-19 for a specific time period.

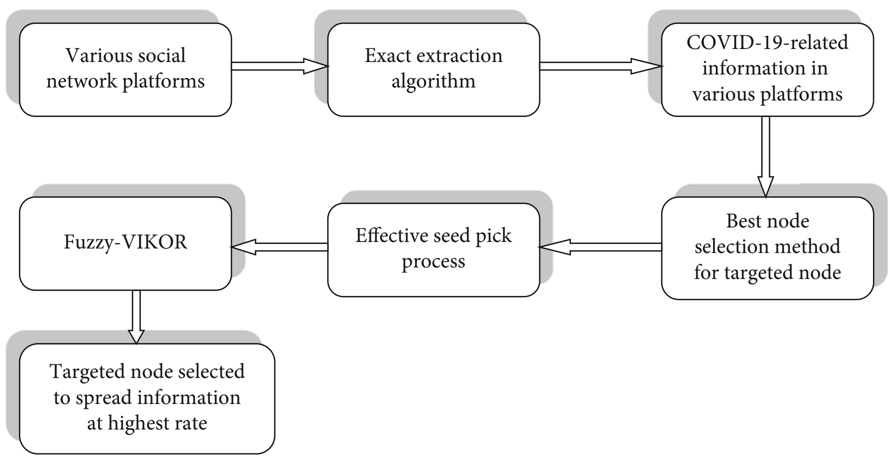

Figure 2: Effectual seed pick framework for pervasive healthcare.

(1) Find the various social network platforms.
(2) To analyse the user engagement in social networks, apply EEA to each social media platform using the skip-gram neural network to build the distribution of words for the text corpus and represent the word as vector representation to assess the COVID-19 topics.
(3) Cluster the words by running the partitioning around medoid algorithm using the closeness metric, the cosine matrix, in the particular vector representation.
(4) Now, the contents related to COVID-19 are separated.
(5) The BNS method is used for distinguishing the network nodes by using the centrality relationship with other nodes and a set of unique attributes such as age, sex, and gender.
(6) Best target nodes are selected in each attribute.
(7) ESP for selecting the seed in every attribute is based on their highest potential of spreading the information to the targeted node which has been selected in the BNS method.
(8) Next, we have to select the targeted node with the highest ranking in the social networks by Fuzzy-VIKOR method based on their relative distance to their solution for spreading the information to the maximum users.
(9) Step 8 has to be repeated whenever the new information diffuses in the social networks.

Algorithm 1: Proposed algorithm.

communities. The influence of this algorithm is maximized across trajectory databases. The algorithm's policy differs from IM techniques in that it does not use a diffusion model to propagate influence.

Stein et al. [55] developed a heuristic technique for increasing network influence spread by randomly selecting border nodes as seeds and then operating seed tweaking based on the friendship paradox to increase influence maximization in the network. Approximating random walks is a two-step community-based technique presented by Wilder et al. [56] to affect a socially explored network (ARISEN) by first weighting each community and then offering a weight refinement procedure to maximize impact.

Zareie et al. [57] identified two groups of techniques, one focusing on user attributes and social information and the other on more sophisticated structural interactions such as overlap. They employ, among other things, user trust and cost. The strategy depends on the message's interest. Yang et al. [58] proposed TOPSIS in the Susceptible-Infected-Recovered (SIR) model to vigorously identify the influential

Table 2: Experimental setup.

Tool	Python
OS	Windows 7 (64-bit)
Processor	Intel premium
Ram	8 GB RAM
Dataset	Social network data, Twitter data

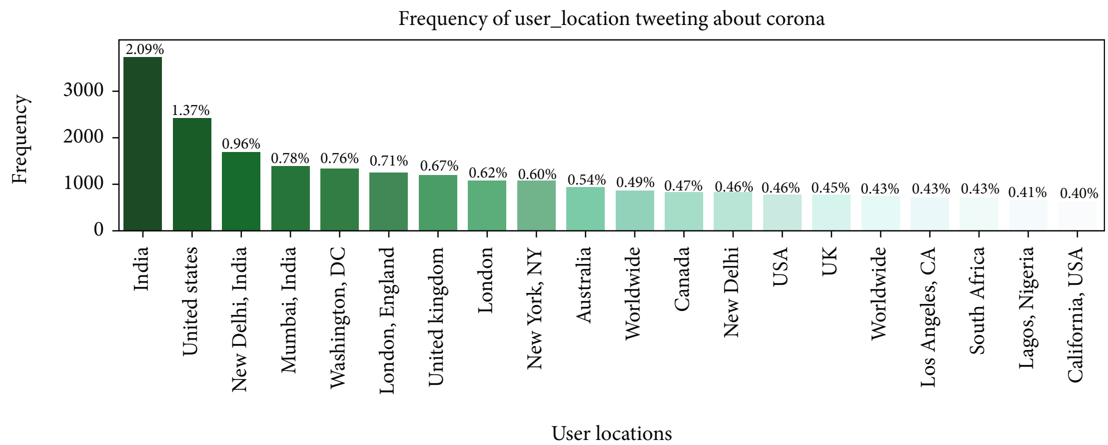

Figure 3: Frequency of tweets versus location of users.

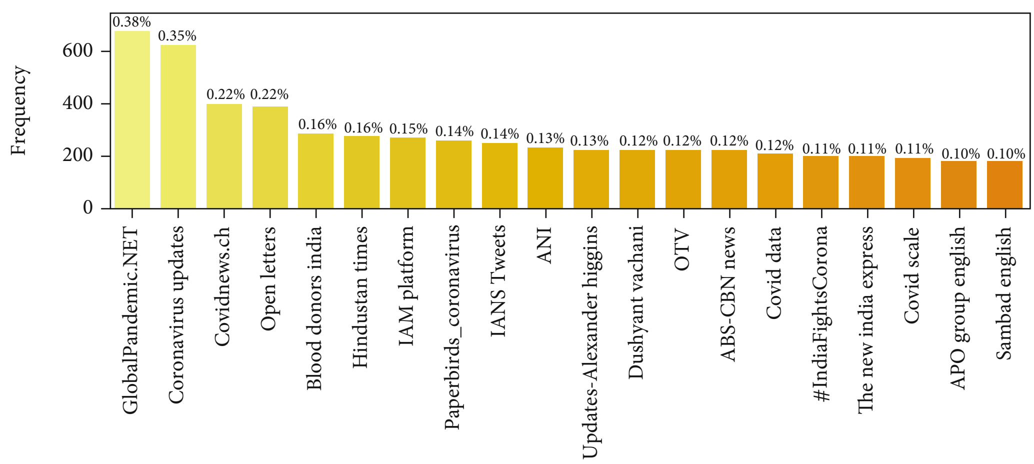

Figure 4: Frequency of user_name tweeting about corona.

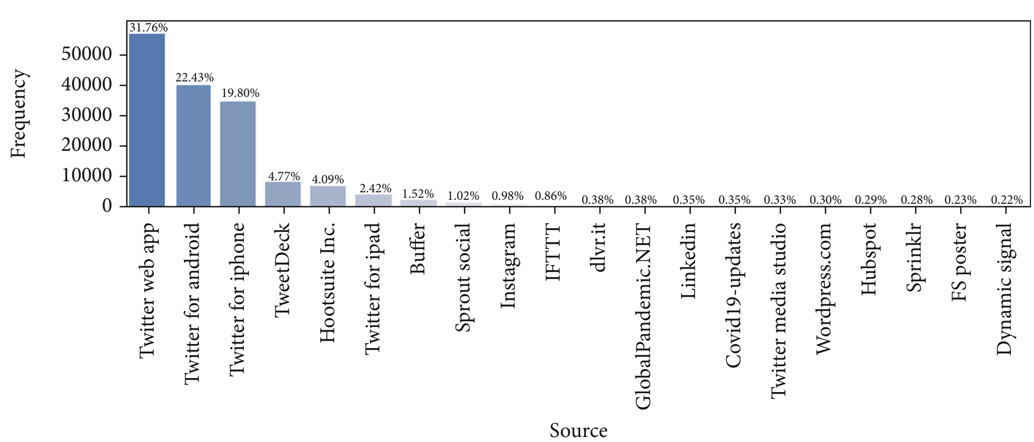

Figure 5: Frequency of source tweeting about corona.

leader in social networks, while Yang et al. [59] determined the values of the weights using entropy weighting. Robles et al. [60] used multiobjective optimization algorithms to increase income while minimizing expenditures in viral marketing efforts. Karczmarczyk et al. [61] used the PROMETHEE II method to assess the effectiveness of social network viral marketing initiatives and provide campaign planners with information.

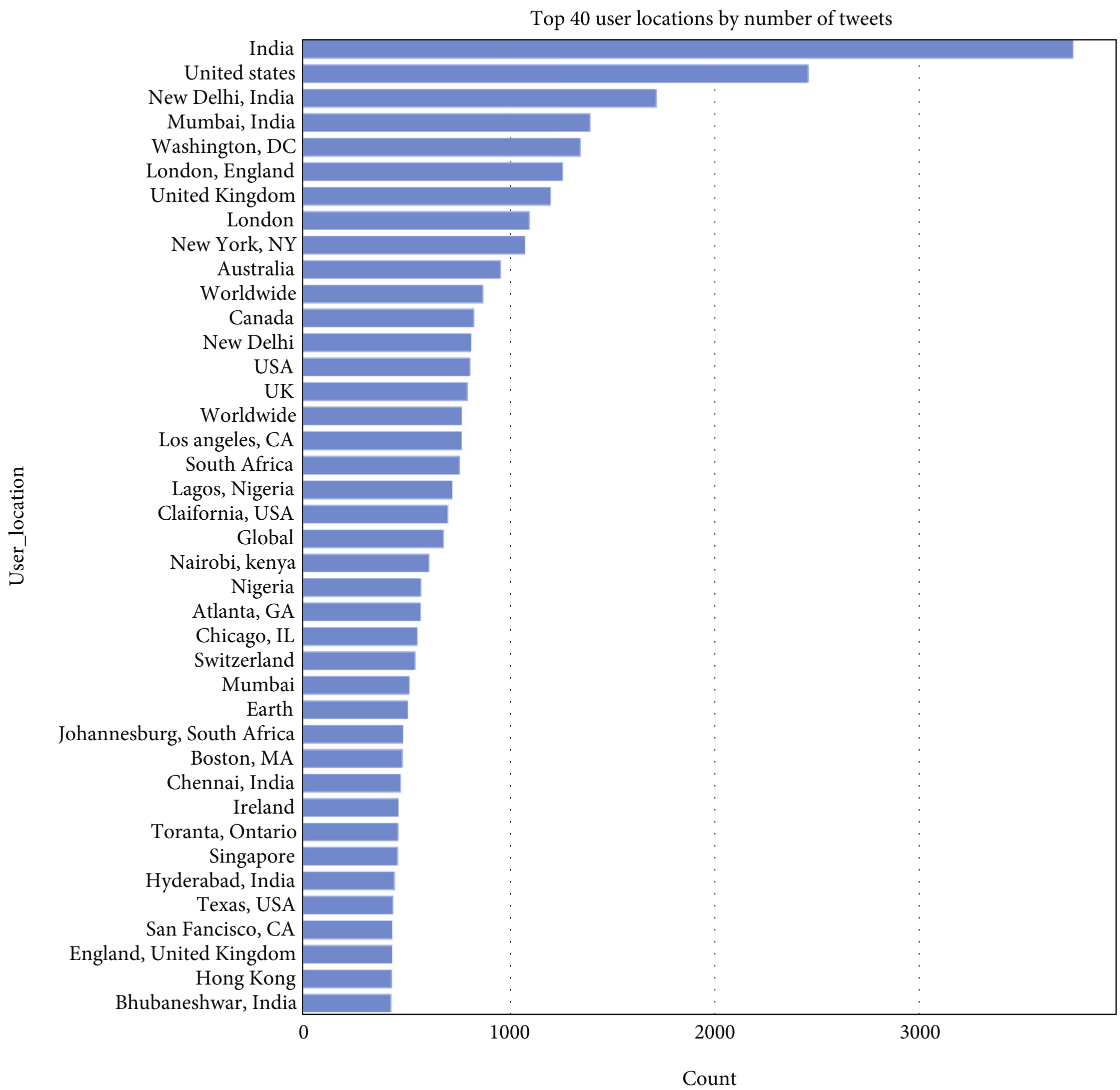

Figure 6: Graph between user_location and the count of tweets related to corona.

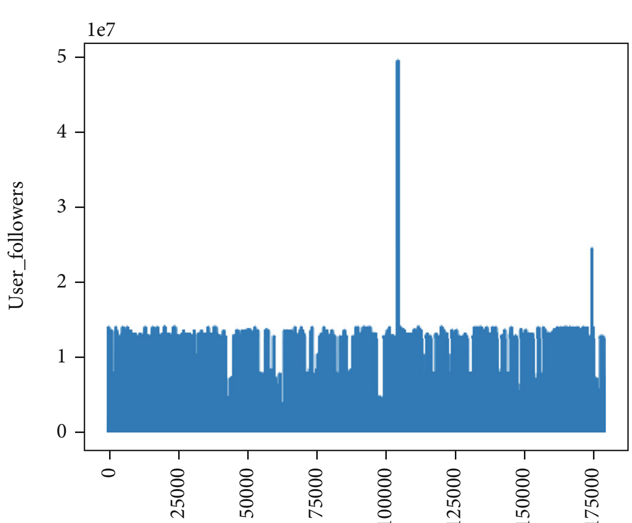

Figure 7: Graph between users and user_followers.

According to a survey of existing works, only a tiny fraction of the enormous number of research connected to information propagation and effect maximization is focused on targeting consumers with specific features. To overcome the above issues, a novel method has been developed.

2.1. Motivation. Nowadays, social media plays an essential role in all fields, particularly healthcare, due to the pandemic. Most of the information are shared via social media related to the diseases, especially for COVID-19-related information. Some of the information is shared by the medical people which may be trusted. Several methods have been developed to separate the information about COVID-19 from the different social media platforms.

However, the extraction may not give exact results. After extracting the information from the various social media platforms, we need to find the exact seed from the set of users in social media to maximize the information spreading about the COVID-19, and that should be very accurate information from the people, while the existing works

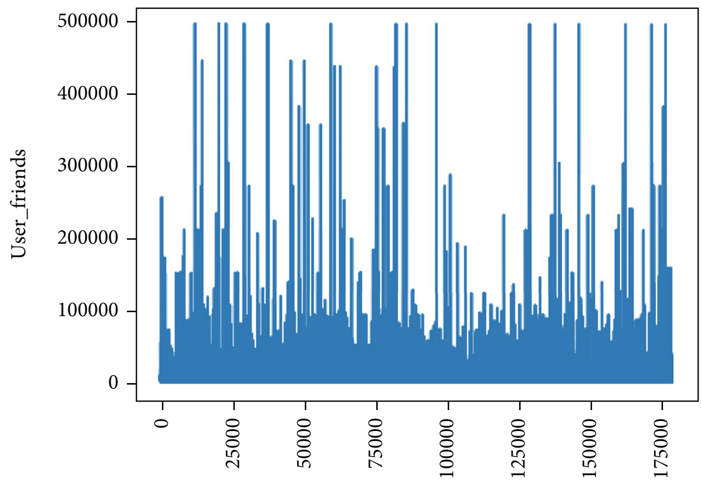

Figure 8: Graph between users and user_friends.

Table 3: Accuracy of proposed framework and other previous algorithm.

Methods	Accuracy(%)
Rand	82.17
LDAG	81.08
MC-CELF	64.9
CEF++	82.87
CIM	81.4
ComPath	78.1
PaS	81
Proposed	97.9

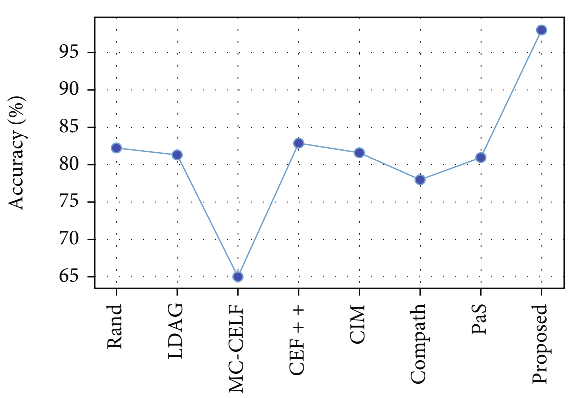

Figure 9: Graph showing accuracy of proposed framework.

focused on single attributes. To overcome the gaps in the existing works, there is a need to develop a new novel approach. Other related research for finding opinion leaders in health related area are discussed in Table 1.

3. Effectual Seed Pick Framework for Pervasive Healthcare

3.1. Proposed Method. Social media is an interactive, computer-mediated platform that enables people to create and share information. Its use in healthcare has been widely debated. Despite the increasing number of healthcare professionals who use it, the advantages of social media remain unclear. In the pandemic scenario, the utilisation of social media in healthcare was touchy. There is still considerable debate about the payback of social media in actual learning and improving the quality of treatment offered. We extract and evaluate every COVID-19-related topic in order to provide a summary of the conversation around the viral eruption on the various platforms. The skip-gram neural network is used to create distributed representations of words for the corpus of the text from each social media platform after applying the Exact Extract Algorithm (EEA) to the textual content on each platform. The term is represented as a vector to consider the themes around which the COVID-19 debate is focused, and the partitioning around medoid technique is used to cluster the words whose vector representations use the cosine distance matrix as a closeness measure. To optimise the influence across various social media platforms, we must first separate the content that is linked to COVID-19. To achieve the presumptive communities, the prior strategies concentrated on single qualities and node attributes for opinion leader selection. However, social media success is often based on choosing qualities of the target audience with variable weights, such as age, gender, or location, depending on how important they are in terms of the effectiveness of the campaign. The proposed novel Multicriteria-Based Effectual Seed Pick Framework with a unique opinion value for each criterion employed to choose the first seeds to successfully attain the network's targeted nodes by Best Node Selection (BNS) method in which the network users/nodes are distinguished not only by their centrality relationships with other nodes/users but also by a set of unique attribute parameters addresses this intriguing research gap as in Figure 2.

After reaching the targeted multiattribute nodes, the information spread in a social network proceed by the Effective Seed Pick Process (ESP) in which multiple attributes are taken into account to choose the seeds with the best chance of eventually propagating the information to the targeted node. To obtain a network node ranking with the highest

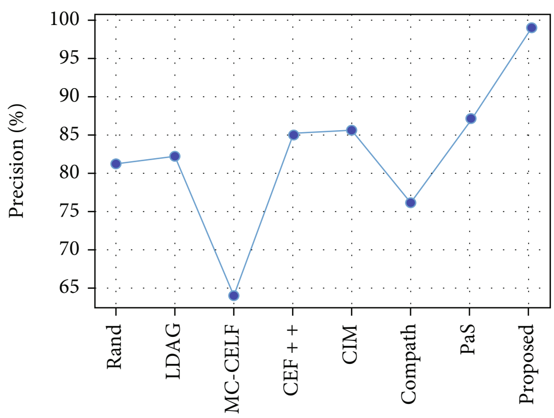

Figure 10: Precision of previous algorithm and the proposed framework.

Table 4: Precision of proposed framework and the previous algorithm.

Methods	Precision (%)
Rand	81.17
LDAG	82.08
MC-CELF	63.09
CEF++	84.87
CIM	85.4
ComPath	76.1
PaS	87
Proposed	98.9

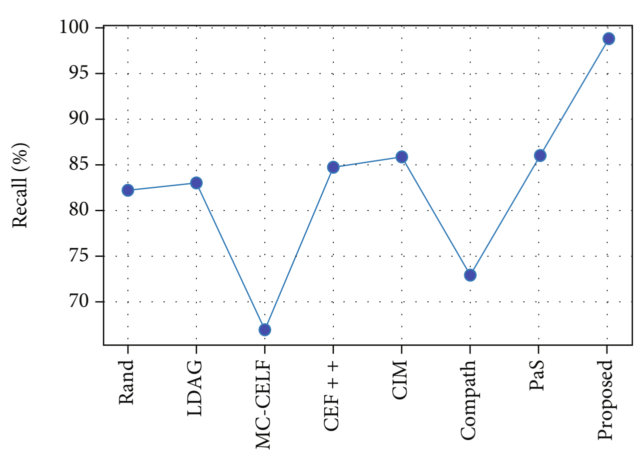

Figure 11: Graph showing recall of proposed framework compared to previous algorithm.

to reach the destination nodes in the social networks, the multicriterion evaluation of the nodes has been done in which Fuzzy-VIKOR is used based on their relative distance to their solution. As a result, the proposed approach gives exact information about COVID-19 from the various social networks and improves the node selection to increase the influence.

Table 5: Recall value of proposed as well as other previous algorithms.

Methods	Recall (%)
Rand	82.17
LDAG	83.08
MC-CELF	66.9
CEF++	84.67
CIM	85.8
ComPath	73.1
PaS	86
Proposed	98.7

3.2. Process Flow

3.2.1. Find Various Social Network Platforms. We are going to use COVID-19 Twitter comment dataset which contains tweets having hashtags #coronavirus, #coronavirusoutbreak, #coronavirusPandemic, #covid19, and #covid_19 from 17 March onwards. The dataset also consist of the additional hashtags #epitwitter and #ihavecorona.

The different variables of the dataset consist of the text tweeted by the tweeter user, location of the user account, country code, and the hash tags used by the users. As the volume of the tweets is large, there can be some gaps for some hashtags as all the tweets with the hashtags cannot be captured. To make the dataset file manageable, the data is split into half-month data in one file and the rest in the other file. Some hashtags are frequently used, and some are used less frequently. As a result, less frequently used hashtags will remain for a longer time.

The most frequently used hashtag is #coronavirus, which seems to be the most popular. Despite scraping over 500,000 tweets, there were still gaps between the number of retweets and the hashtag. This dataset excludes retweets since the retweet option has been set to FALSE (although a count of retweets is provided as a variable).

3.2.2. For Analysing the User Engagement in Social Networks. This paper introduce the skip-gram neural network to learn word embedding for the text of social media platforms. It learns word embedding for the topics related to the COVID-19. To represent the text corpus of social media platforms, we apply the skip-gram neural network. If the given content is represented by the sequence of words w_1, $w_2, \cdots, w_t$, we use stochastic gradient descent with gradient computed through backpropagation rule for maximizing the average log probability

$$\frac{1}{T}\sum_{t=1}^{T}\left[\sum_{j=-k}^{k}\log p\left(w_t+\frac{j}{w_t}\right)\right], \quad (1)$$

where k is the size of the training window. As a result, during training, comparable word vector representations are close to one another.

Each word in the skip-gram model is connected to its input and output vectors, u_w and v_w, in some way. Given

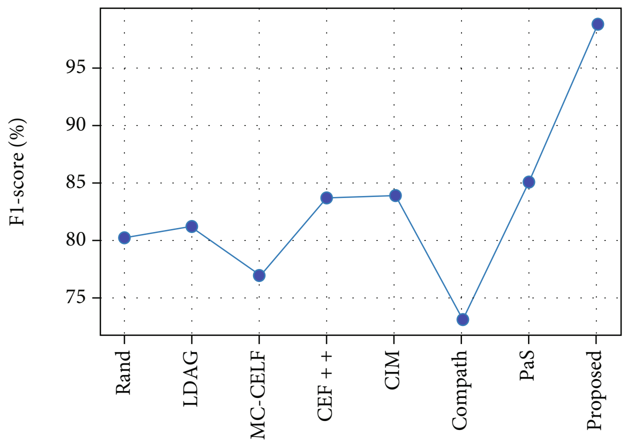

Figure 12: Graph showing F1-score for proposed as well as existing algorithms.

Table 6: F1-score for proposed as well as other existing algorithms.

Methods	F1-score (%)
Rand	80.17
LDAG	81.08
MC-CELF	76.9
CEF++	83.67
CIM	83.8
ComPath	73.1
PaS	85
Proposed	98.7

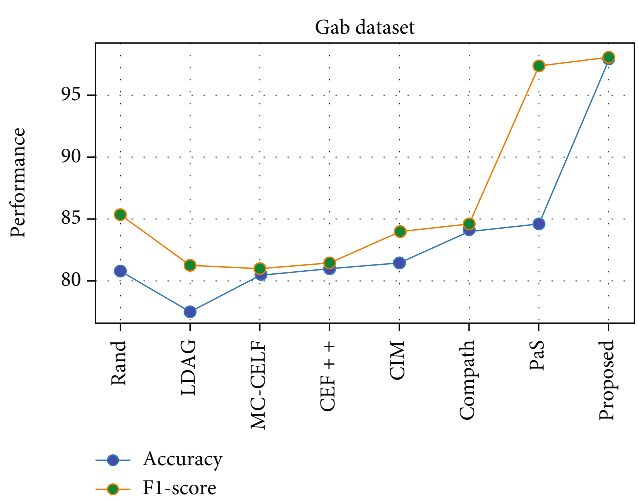

Figure 13: Performance of proposed algorithm and other algorithms for Gab dataset.

the word w_j, the likelihood that the word w_i would be accurately predicted is given by

$$p\left(\frac{w_i}{w_j}\right) = \frac{\exp\left(u_{w_i}^T v_{w_j}\right)}{\sum_{i=1}^{V} \exp\left(u_l^T v_{w_j}\right)}. \tag{2}$$

Table 7: Accuracy and F1-score for proposed as well as other algorithms.

Methods	Accuracy (%)	F1-score
Rand	80.82	85.31
LDAG	77.56	81.32
MC-CELF	80.45	81
CEF++	81	81.5
CIM	81.5	84
ComPath	84	84.6
PaS	84.6	97.34
Proposed	97.7	98

One corpus vocabulary contains V words. The dimensionality of word vectors and the size of the surrounding word window are two important factors that have an impact on training effectiveness. Here, we have utilised 6 words for the window and 200 as the vector dimension, which is a conventional choice for training large datasets, where $p(w_i/w_j)$ is the probability of correctly predicting word, V is the number of words in the corpus vocabulary, u_w is the input vector, and v_w is the output vector.

3.2.3. Clustering the Words. The cluster of the words is done by executing the partitioning around medoid algorithm, which uses the proximity metric and the cosine matrix in the particular vector representation. To evaluate the topics related to corona, we used PAM to cluster words and used the cosine distance matrix as the proximity metric of words' vector representations. After locating the k clusters, figure out the typical silhouette width for each value of k. We must compute the average pairwise Jaccard similarity between clusters using 90% of the data to establish the cluster's stability. The subject of each cluster is then established by creating word clusters. In order to analyse the conversations surrounding the coronavirus pandemic, we define the distribution over themes _c for a given content c as the distribution of its words over word clusters. To assess the significance of each topic inside a corpus, we restrict ourselves to contents c with max $c > 0.5$ and regard each of them as a separate topic.

$$I = \left[\frac{R_0}{(1+d)^t}\right]^t, \tag{3}$$

where I is an incidence, t is the addition of days, R_0 is a basic reproduction number, and d is a damping factor.

$$\begin{aligned} \partial_t S &= -\beta S \cdot \frac{I}{N}, \\ \partial_t I &= \beta S \cdot \frac{I}{N} - \gamma I, \\ \partial_t R &= \gamma I, \end{aligned} \tag{4}$$

where S is the count of those who are susceptible, I is the addition of those infected, and R is the addition of those who recovered.

Table 8: Performance of proposed as well as previous algorithm for Reddit dataset.

Methods	Precision (%)	Recall
Rand	71.82	81.22
LDAG	71.88	77.40
MC-CELF	73.78	81.01
CEF++	81	81
CIM	82	82.5
ComPath	84.4	84
PaS	84.6	83.6
Proposed	97.8	98

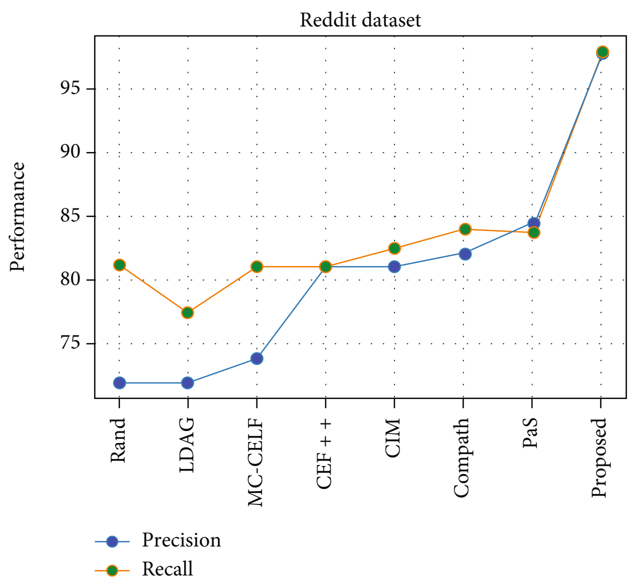

Figure 14: Performance graph of proposed as well as previous algorithms for Reddit dataset.

Table 9: Performance of proposed as well as previous algorithm for Instagram dataset.

Methods	Precision (%)	Recall (%)
Rand	70.82	80.2
LDAG	71.88	78.40
MC-CELF	75.78	82.01
CEF++	89	81
CIM	81	82.5
ComPath	82	83
PaS	83.4	88.6
Proposed	97.8	98

3.2.4. BNS Method for Distinguishing the Network Nodes. The BNS method is used to distinguish the network nodes by using the centrality relationship with other nodes/users and a set of unique attributes such as a gender. The Best Node Selection (BNS) method distinguishes network nodes and users by a variety of distinctive characteristics, such as user friends, followers, and favourites, in addition to their centrality interactions with other nodes and users. ESP for selecting the seed in every attribute is based on their highest potential of spreading the information to the targeted node which has been selected in BNS method.

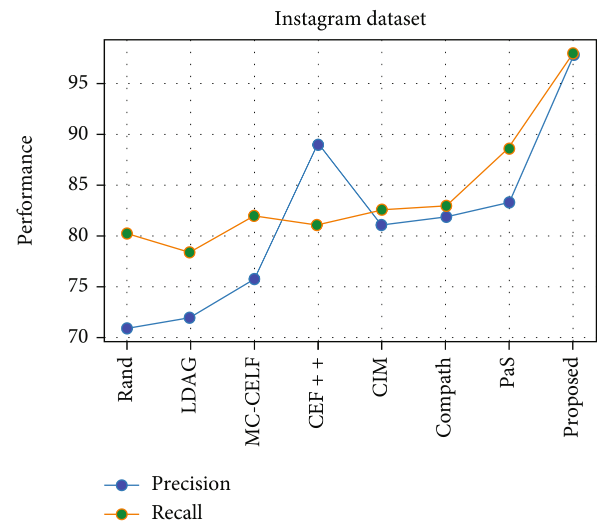

Figure 15: Performance graph of proposed as well as previous algorithms for Instagram dataset.

3.2.5. Selection of Targeted Node Using Fuzzy-VIKOR. Next, we have to select the targeted node with the highest ranking in the social networks by the Fuzzy-VIKOR method based on their relative distance to their solution for spreading the information to the maximum users.

The criterion values of all the network's vertices are then used to construct a decision matrix, $D[x_{ij}]$, where the vertices are represented by the m rows and the criteria are represented by the n columns.

$$D[x_{ij}] = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mn} \end{pmatrix}, \quad (5)$$

where $D[x_{ij}]$ is decision matrix, m row represents the vertices, and n columns represent the criteria.

In the algorithm, the second step is to normalize the decision matrix. Different formulas are employed for the cost and benefit criteria: the ability to alter the weights of various decision criteria makes the Fuzzy-VIKOR-based techniques more versatile than the conventional aggregating procedures.

The analyst adjusts the weights of each choice criterion in accordance with the preferences of the decision maker. To maximize the chance of reaching the targeted network nodes through the seeded network nodes as much as is practical, the marketer modifies the weights of several criteria in the case of the seed selection problem under consideration.

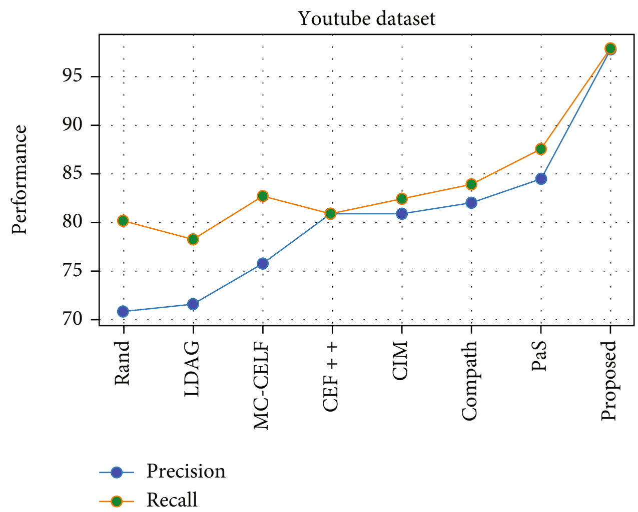

FIGURE 16: Performance graph of proposed as well as previous algorithms for YouTube dataset.

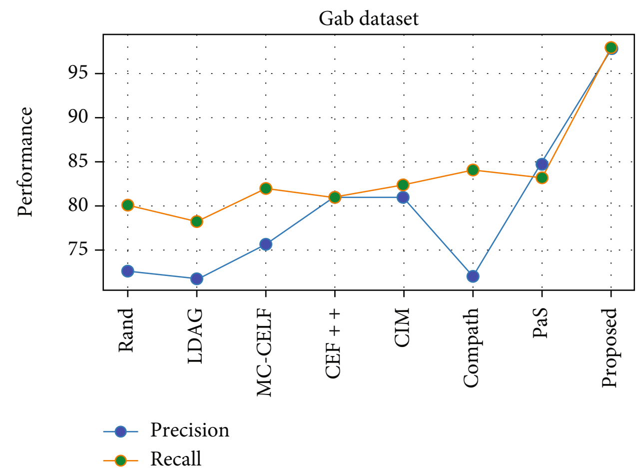

FIGURE 17: Performance graph of proposed as well as previous algorithms for Gab dataset.

TABLE 10: Performance of proposed as well as previous algorithm for YouTube dataset.

Methods	Precision (%)	Recall (%)
Rand	70.82	80.22
LDAG	71.58	78.38
MC-CELF	75.78	82.71
CEF++	81	81
CIM	81	82.5
ComPath	82	84
PaS	84.4	87.6
Proposed	97.8	98

TABLE 11: Performance of proposed as well as previous algorithm for Gab dataset.

Methods	Precision (%)	Recall (%)
Rand	72.62	80.12
LDAG	71.78	78.30
MC-CELF	75.78	82.01
CEF++	81	81
CIM	81	82.4
ComPath	72	84
PaS	84.8	83.2
Proposed	97.8	98

The weights are established depending on the analyst's expertise, knowledge, and experience.

The ideal solutions for the algorithm's fourth phase, V_j and V_j^+, are computed. The positive ideal solution for the seed selection problem under study would be a vertex that had the highest possible values for each criterion. The negative ideal solution, in contrast, would be a vertex with the lowest values for each criterion.

$$\begin{aligned} V_{ij} &= w_j \cdot r_{ij}, \\ V_j^+ &= \{v_1^+, v_2^+, v_3^+, \cdots, v_n^+\}, \\ V_j^- &= \{v_1^-, v_2^-, v_3^-, \cdots, v_n^-\}, \end{aligned} \tag{6}$$

where V_j^+ and V_j^- are positive and negative ideal solutions, respectively,

$$\begin{aligned} D_i^+ &= \sqrt{\sum_{j=1}^{n} \left(V_{ij} - V_j^+\right)^2}, \\ D_i^- &= \sqrt{\sum_{j=1}^{n} \left(V_{ij} - V_j^-\right)^2}, \end{aligned} \tag{7}$$

where D_i^+ and D_i^- are Euclidian distances between each network vertex

$$CC_i = \frac{D_i^-}{D_i^- + D_i^+}. \tag{8}$$

The vertices are then ranked using the acquired CC_i scores to create the final ranking, which can subsequently be used to choose the vertices for the first network.

(1) Extract COVID-19-Related Words.

(Step 1) Tweeter dataset j is cleaned of stop words, special character, and hashtags

(Step 2) Build word embedding for test corpus of Tweeter

(Step 3) PAM $= \mathrm{P}(w_i/w_j) = \exp\,(u_{wi}^T v_{wj})/\sum_{l=1}^{v}\exp(u_l^T v_{wj.}$).

(Step 4) Distributed representation of words by the skip-gram model

Table 12: Performance of proposed as well as previous algorithm for Instagram dataset.

Methods	Accuracy (%)	F1-score (%)
Rand	73.82	80.22
LDAG	73.88	76.40
MC-CELF	74.78	82.01
CEF++	80	82
CIM	80.5	81.4
ComPath	82	83
PaS	84.6	84.2
Proposed	97.8	98

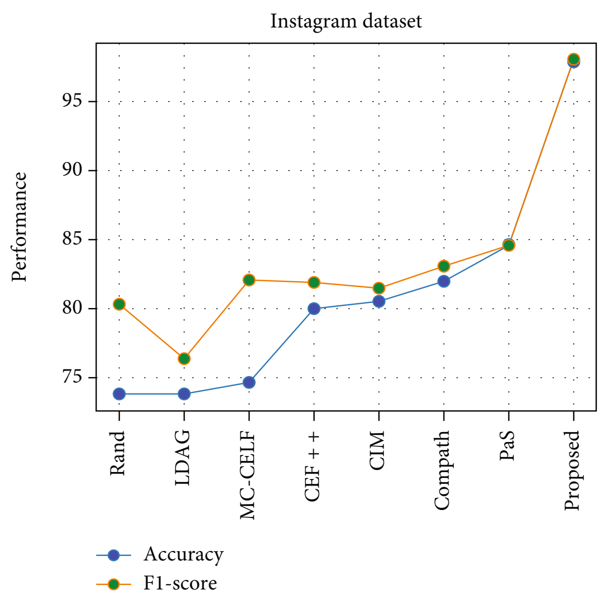

Figure 18: Performance graph of proposed as well as previous algorithms for Instagram dataset.

$$\text{Stochastic gradient decent with gradient} = \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{j=-k}^{K}\log P\left(\frac{w_{t+j}}{w_t}\right)\right]. \quad (9)$$

(Step 5) Restrict the contents by using $\mathrm{I} = [RO/(1+d)^t]^t$

(Step 6) Output: COVID-19-related words

(2) Best Node Selection Method.

(Step 1) Input: skip-gram model outputs COVID-19-related words

(Step 2) Set attributes: followers, favourite, and friends

(Step 3) Assign numerical values of set attributes

(Step 4) Compute more centrality measures between the attributes

Table 13: Performance of proposed as well as previous algorithm for Reddit dataset.

Methods	Accuracy (%)	F1-score (%)
Rand	72.82	80.22
LDAG	71.88	78.40
MC-CELF	75.78	82.01
CEF++	81	81
CIM	81	82.4
ComPath	82	84
PaS	84.4	83.6
Proposed	97.8	98

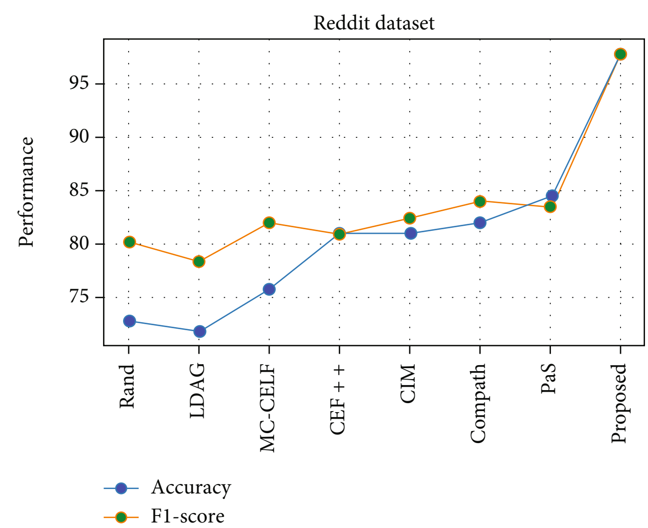

Figure 19: Performance graph of proposed as well as previous algorithms for Reddit dataset.

(Step 5) Add steps 3 and 4

(Step 6) Find maximum values of the node

(Step 7) Output: best node selected

(3) Effective Seed Pick Process.

(Step 1) Input: network G by the skip-gram model

(Step 2) Consideration: followers, friends, and favourite

(Step 3) Find the highest potential using centrality measures and numerical values of attributes

(Step 4) Output: seed pick for the maximum propagation of information in multiple attributes

(4) Fuzzy-VIKOR: To Select Best Node from the Multiple Attribute.

(Step 1) Input: seed pick nodes in effective seed pick process

(Step 2) Constructed nodes make a fuzzy decision

$$x_{ij}\check{} = \frac{1}{k}\left\{\widetilde{xij}^{1} + \widetilde{xij}^{2} \cdots \widetilde{xij}^{k}\right)\right\}. \quad (10)$$

$x_{ij}\check{}$ is the input from the seed nodes

(Step 3) Determine best and worst $x_{ij}\check{}$

$$f_1^* = \max\left[f_{j,j=1..J}\right],$$
$$f_1^\wedge = \min\left[f_{j,j=1..J}\right]. \quad (11)$$

(Step 4) $S_j = \text{sum}[w_i(f_i^* - f_{ij})/f_1^* - f_1^\wedge]$,

$$R_j = \max\left[\frac{w_i\left(f_1^* - f_{ij}\right)}{f_1^k - f_1^\wedge}\right], \quad (12)$$

where w_i is the weights and f_i is x_{ij}

(Step 5) Compute the values

$$Q_j = v\left(\frac{s_j - s^*}{s^\wedge - s^*}\right) + (1 - v)\left(\frac{R_j - R^*}{R^\wedge - R^*}\right), \quad (13)$$

where $s^* = \min(s_j)$, $s^\wedge = \max(s_j)$, $R^* = \min(R_j)$, $R^\wedge = \max(R_j)$, $(1 - v)$ is the weight of the individual best node, and $v = (n + 1)/2n$

(Step 6) Rank the alternatives sorting by S, R, and Q from minimum value

(Step 7) C1: check $Q(A(2)) - Q(A(1)) \geq DQ$

$$\text{DQ} = \frac{1}{J - 1}. \quad (14)$$

C2: $A(1)$ must also be the best ranked by S or R/S and R

(Step 8) Output: best node selected from the multiple attribute to propagate the information in a selected social network pervasive healthcare

4. Result and Discussion

This section provides a thorough analysis of the implementation outcomes, highlights the effectiveness of our suggested system, and concludes with a comparison to make sure that our suggested system outperforms the alternatives.

4.1. Experimental Setup. The system requirements for this work were implemented in Python, and the results of the simulation are given below. The experimental design for using an efficient seed select framework for influencing maximization is described in Table 2.

4.2. Evaluation Metrics and Simulation Output. This section presents the network simulation and the resulted outputs of implementations. The results are as follows in the following figures. Figures 3–8 show the frequency of tweets with respect to the location from where tweeter users are tweeting related to corona.

The accuracy of the proposed methodology and the obtained image is detected by the following equation.

The accuracy of the clinical text data is calculated using

$$\text{Accuracy} = \left[\frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}\right] * 100, \quad (15)$$

where TP is the true positive value, TN is the true negative value, FP is the false positive value, and FN is the false negative value.

F1-score is defined as

$$\text{F1-score} = \frac{2 \times (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}}, \quad (16)$$

where

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (17)$$

Precision is defined as how closely two or more measurements match one another. The equation is displayed as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (18)$$

where TP is true positive and FP is false positive.

The capacity of a model to correctly predict an outcome is known as recall. The definition of the recall formula is

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (19)$$

where TP is true positive and FN is false negative.

The accuracy of the Twitter dataset is taken for comparison with the proposed system and other existing techniques like Rand, LDAG, MC-CELF, CEF++, CIM, ComPath, and PaS as in Table 3. The graph shows the accuracy of proposed work; proposed technique accuracy seems to be 97.9% and existing techniques are below the proposed technique's accuracy as in Figure 9. The existing technique Rand is 82.17%, LDAG is 81.08%, MC-CELF is 64.9%, CEF++ is 82.87%, CIM is 81.4%, ComPath is 78.1%, and PaS is 81%. From this result, it is determined that the proposed technique is more efficient.

The precision of the Twitter dataset is taken for comparison with the proposed system and other existing techniques. The graph shows the precision of the proposed work; the proposed technique precision seems to be 98.9% and existing techniques are below the proposed technique's precision as in Figure 10. The existing technique Rand is 81.17%, LDAG is 82.08%, MC-CELF is 63.09%, CEF++ is 84.87%, CIM is 85.4%, ComPath is 76.1%, and PaS is 87% as in Table 4. From this result, it is determined that the proposed technique is more efficient.

The recalls of the Twitter dataset are taken for comparison with the proposed system and other existing techniques as in Figure 11. The graph shows the recalls of proposed work; the proposed technique recall seems to be 98.7% and existing techniques are below the proposed technique's precision. The existing technique Rand is 82.17%, LDAG is 83.08%, MC-CELF is 66.9%, CEF++ is 84.67%, CIM is 85.8%, ComPath is 73.1%, and PaS is 86% as in Table 5. From this results, it is determined that the proposed technique is more efficient as in Figure 12.

The F1-score of the Twitter dataset is taken for comparison with the proposed system and other existing techniques as in Table 6. The graph in Figure 12 shows the F1-score of the proposed work; the proposed technique's F1-score seems to be 98.7% and existing techniques are below the proposed technique's precision. The existing technique Rand is 80.17%, LDAG is 81.08%, MC-CELF is 76.9%, CEF++ is 84.67%, CIM is 83.8%, ComPath is 73.1%, and Pas is 85%. From this result, it is determined that the proposed technique is more efficient as in Figure 13.

The accuracy and F1-score of the Gab dataset is taken for comparison with the proposed system and other existing techniques as in Table 7. The graph shows the accuracy and F1-score of the proposed work; the proposed technique accuracy seems to be 97.7%, the F1-score seems to be 98%, and existing techniques are below the proposed technique's accuracy and F1-score. From these results, it is determined that the proposed technique is more efficient as in Figure 13.

The precision and recall of the Reddit dataset is taken for comparison with the proposed system and other existing techniques like as in Table 8. The graph in Figure 14 shows the precision and recall of the proposed work; the proposed technique precision seems to be 97.8%, recall seems to be 98%, and existing techniques are below the proposed technique's precision and recall. From these results, it is determined that the proposed technique is more efficient.

The precision and recall of the Instagram dataset is taken for comparison with the proposed system and other existing techniques as in Table 9. The graph shows the precision and recall of the proposed work; the proposed technique precision seems to be 97.8%, recall seems to be 98%, and existing techniques are below the proposed technique's precision and recall. From these results, it is determined that the proposed technique is more efficient as in Figure 15.

The precision and recall of the YouTube dataset is taken for comparison with the proposed system and other existing techniques as in Figure 16. The graph shows the precision and recall of the proposed work; the proposed technique precision seems to be 97.8%, recall seems to be 98%, and existing techniques are below the proposed technique's precision and recall as in Table 10. From these results, it is determined that the proposed technique is more efficient.

The precision and recall of the Gab dataset is taken for comparison with the proposed system and other existing techniques as in Table 11. The graph shows the precision and recall of the proposed work; the proposed technique precision seems to be 97.8%, recall seems to be 98%, and existing techniques are below the proposed technique's precision and recall as in Figure 17. From these results, it is determined that the proposed technique is more efficient.

The accuracy and F1-score of the Instagram dataset is taken for comparison with the proposed system and other existing techniques as in Table 12. The graph as in Figure 18 shows the accuracy and F1-score of the proposed work; the proposed technique accuracy seems to be 97.8%, F1-score seems to be 98%, and existing techniques are below the proposed technique's accuracy and F1-score. From these results, it is determined that the proposed technique is more efficient.

The accuracy and F1-score of the Reddit dataset is taken for comparison with the proposed system and other existing techniques as in Table 13. The graph as in Figure 19 shows the accuracy and F1-score of the proposed work; the proposed technique accuracy seems to be 97.8%, F1-score seems to be 98%, and existing techniques are below the proposed technique's accuracy and F1-score. From these results, it is determined that the proposed technique is more efficient.

5. Conclusion

As we find the opinion leader using our proposed algorithm, it is having maximum influence maximization in the network. The proposed work can be applied in any area or field like politics, product promotion, or service promotion, so in this paper, we have applied our proposed algorithm in the field of pervasive healthcare. The proposed framework is efficient to provide pervasive healthcare. When for different datasets like YouTube, Facebook, Reddit, Instagram, Gab, and Twitter, the proposed framework is compared with the existing algorithms to find out the seeds for maximum influence maximization, and the accuracy, recall, F1-score, and precision are high as compared to other existing algorithms.

Data Availability

No data were used to support this study

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Ditto and B. Pille, "Marketing on the internet," *Healthcare Executive*, vol. 13, no. 5, pp. 54-55, 1998.

[2] M. Fojt, "Strategies for service quality," *Journal of Services Marketing*, vol. 9, no. 3, pp. 1–72, 1995.

[3] M. Program, S. Magister, and I. Komunikasi, "Social strategy pada media sosial untuk promosi pariwisata daerah istimewa

Yogyakarta," *Interaksi: Jurnal Ilmu Komunikasi*, vol. 4, no. 2, pp. 195–201, 2015.

[4] R. Ferguson, "Word of mouth and viral marketing: taking the temperature of the hottest trends in marketing," *Journal of Consumer Marketing*, vol. 25, no. 3, pp. 179–182, 2008.

[5] I. Technology, *ASHP statement on use of social media by pharmacy professionals*, Best practice, 2020.

[6] J. M. Bernhardt, J. Alber, and R. S. Gold, "A social media primer for professionals," *Health Promotion Practice*, vol. 15, no. 2, pp. 168–172, 2014.

[7] N. S. Fogelson, Z. A. Rubin, and K. A. Ault, "Beyond likes and tweets," *Clinical Obstetrics and Gynecology*, vol. 56, no. 3, pp. 495–508, 2013.

[8] S. A. Moorhead, D. E. Hazlett, L. Harrison, J. K. Carroll, A. Irwin, and C. Hoving, "A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication," *Journal of Medical Internet Research*, vol. 15, no. 4, pp. e85–117, 2013.

[9] D. R. George, L. S. Rovniak, and J. L. Kraschnewski, "Dangers and opportunities for social media in medicine," *Clinical Obstetrics and Gynecology*, vol. 56, no. 3, pp. 453–462, 2013.

[10] S. P. Singh, A. K. Rai, A. Wal, G. Tiwari, R. Tiwari, and A. Parveen, "Effect of social media in health care: uses, risks, and barriers," *World journal of pharmacy and pharmaceutical sciences*, vol. 5, no. 7, pp. 282–303, 2016.

[11] K. C. Chretien and T. Kind, "Social media and clinical care," *Circulation*, vol. 127, no. 13, pp. 1413–1421, 2013.

[12] K. Grindrod, A. Forgione, R. T. Tsuyuki, S. Gavura, and D. Giustini, "Pharmacy 2.0: a scoping review of social media use in pharmacy," *Research in Social and Administrative Pharmacy*, vol. 10, no. 1, pp. 256–270, 2014.

[13] D. S. Dizon, D. Graham, M. A. Thompson et al., "Practical guidance: the use of social media in oncology practice," *Journal of oncology practice*, vol. 8, no. 5, pp. e114–e124, 2012.

[14] F. J. Grajales, S. Sheps, K. Ho, H. Novak-Lauscher, and G. Eysenbach, "Social media: a review and tutorial of applications in medicine and health care," *Journal of medical Internet research*, vol. 16, no. 2, p. e13, 2014.

[15] "Coronavirus disease (COVID-19)," [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=CjwKCAjw_o-HBhAsEiwANqYhp0uAq0egxgu2ONO-i3n9FvL-OoR7T-yG8L9uIDqrtEZz-Ys-vK55hxoCFdgQAvD_BwE. [Accessed: 07-Jul-2021].

[16] "Global Risks Report 2017- reports-World Economic Forum," [Online]. Available: https://reports.weforum.org/global-risks-2017/part-2-social-and-political-challenges/2-1-western-democracy-in-crisis/. [Accessed: 07-Jul-2021].

[17] "Coronavirus disease (COVID-19) situation reports," [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/?gclid=CjwKCAjw_o-HBhAsEiwANqYhp-WmeVhL4J10tNCQPMwRgPpL0gFjeGhxw4nmG3DiqL_IX4fQ6-DIgBoCi9wQAvD_BwE. [Accessed: 07-Jul-2021].

[18] "How to fight an infodemic | Elsevier Enhanced Reader," 2021, https://reader.elsevier.com/reader/sd/pii/S014067362030461X?token=F4887619FED82B55ABE4F6F71165E4606C4464D9C0E0BCCD60C1E9BE7A8515E5934D6DA0F9866A9EF49900567C7FB60B&originRegion=eu-west-1&originCreation=20210706225708.

[19] "Director-General's remarks at the media briefing on 2019 novel coronavirus on 8 February 2020," [Online]. Available: https://www.who.int/director-general/speeches/detail/director-general-s-remarks-at-the-media-briefing-on-2019-novel-coronavirus—8-february-2020. [Accessed: 07-Jul-2021].

[20] M. Mendoza, B. Poblete, and C. Castillo, "Twitter under crisis: can we trust what we RT?," in *SOMA '10: Proceedings of the First Workshop on Social Media Analytics*, pp. 71–79, Washington D.C., District of Columbia, 2010.

[21] K. Starbird, J. Maddock, M. Orand, P. Achterman, and R. M. Mason, "Rumors, false flags, and digital vigilantes: misinformation on twitter after the 2013 Boston marathon bombing," in *IConference 2014 proceedings*, Boston Marathon Bombing, 2014 Mar 1.

[22] L. Kim, S. M. Fast, and N. Markuzon, "Incorporating media data into a model of infectious disease transmission," *PLoS One*, vol. 14, no. 2, pp. 1–13, 2019.

[23] T. Sharot and C. R. Sunstein, "How people decide what they want to know," *Nature Human Behaviour*, vol. 4, no. 1, pp. 14–19, 2020.

[24] J. Shaman, A. Karspeck, W. Yang, J. Tamerius, and M. Lipsitch, "Real-time influenza forecasts during the 2012-2013 season," *Nature communications*, vol. 4, no. 1, pp. 1–10, 2013.

[25] C. Viboud and A. Vespignani, "The future of influenza forecasts," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 8, pp. 2802–2804, 2019.

[26] J. Kulshrestha, M. Eslami, J. Messias et al., "Quantifying search bias: investigating sources of bias for political searches in social media," in *CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 417–432, Portland, Oregon, USA, 2017.

[27] A. L. Schmidt, F. Zollo, M. del Vicario et al., "Anatomy of news consumption on Facebook," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 12, pp. 3035–3039, 2017.

[28] M. Del Vicario, W. Quattrociocchi, A. Scala, and F. Zollo, "Polarization and fake news: early warning of potential misinformation targets," *ACM Transactions on the Web*, vol. 13, no. 2, 2019.

[29] C. Wardle and H. Derakhshan, *Information disorder: toward an interdisciplinary framework for research and policy making*, Council of Europe Strasbourg, 2017.

[30] S. Vosoughi, D. Roy, and S. Aral, *Community development and gender in the 21st century*, vol. 1151, Gender Studies, 2018.

[31] D. Ruths, "Social science: the misinformation machine," *Science (80-.)*, vol. 363, no. 6425, p. 348, 2019.

[32] M. M. Daliri Khomami, A. Rezvanian, N. Bagherpour, and M. R. Meybodi, "Minimum positive influence dominating set and its application in influence maximization: a learning automata approach," *Applied Intelligence*, vol. 48, no. 3, pp. 570–593, 2018.

[33] J. Zhu, J. Zhu, S. Ghosh, W. Wu, and J. Yuan, "Social influence maximization in hypergraph in social networks," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 4, pp. 801–811, 2019.

[34] L. Katz, "A new status index derived from sociometric analysis," *Psychmetrika*, vol. 18, no. 1, pp. 39–43, 1953.

[35] T. W. Valente and P. Pumpuang, “Identifying opinion leaders to promote behavior change,” *Health education & behavior*, vol. 34, no. 6, pp. 881–896, 2007.

[36] E. M. Rogers and D. C. Cartano, “Methods of measuring opinion leadership,” *Public Opinion Quarterly*, vol. 26, no. 3, pp. 435–441, 1962.

[37] J. Goldenberg, B. Libai, and E. Muller, “Using complex systems analysis to advance marketing theory development: modeling heterogeneity effects on new product growth through stochastic cellular automata,” *Academy of Marketing Science Review*, vol. 9, no. 3, pp. 1–18, 2001.

[38] L. Locock, S. Dopson, D. Chambers, and J. Gabbay, “Understanding the role of opinion leaders in improving clinical effectiveness,” *Social Science & Medicine*, vol. 53, no. 6, pp. 745–757, 2001.

[39] R. L. Davis, “Accelerating the diffusion of innovations using opinion leaders,” *The Annals of the American Academy of Political and Social Science*, vol. 566, no. 1, pp. 55–67, 1999.

[40] A. R. Mellanby, J. B. Rees, and J. H. Tripp, “Peer-led and adult-led school health education: a critical review of available comparative research,” *Health Education Research*, vol. 15, no. 5, pp. 533–545, 2000.

[41] J. M. Stephenson, V. Strange, S. Forrest et al., “Pupil-led sex education in England (RIPPLE study): cluster-randomised intervention trial,” *Lancet*, vol. 364, no. 9431, pp. 338–346, 2004.

[42] E. Harrin, “Peer education in practice,” *Health Education*, vol. 97, no. 4, pp. 132–135, 1997.

[43] R. Iyengar, C. van den Bulte, and T. W. Valente, “Opinion leadership and social contagion in new product diffusion,” *Marketing Science*, vol. 30, no. 2, pp. 195–212, 2011.

[44] L. L. Jacoby and C. Kelley, “Unconscious influences of memory: dissociations and automaticity,” in *The neuropsychology of consciousness*, pp. 201–233, Academic Press, 1992.

[45] C. Te Li, S. De Lin, and M. K. Shan, “Finding influential mediators in social networks,” in *WWW '11: Proceedings of the 20th international conference companion on World wide web*, pp. 75-76, Hyderabad, India, 2011.

[46] L. C. Freeman, “A set of measures of centrality based on betweenness,” *Sociometry*, vol. 40, no. 1, p. 35, 1977.

[47] J. Holliday, S. Audrey, R. Campbell, and L. Moore, “Identifying well-connected opinion leaders for informal health promotion: the example of the ASSIST smoking prevention program,” *Health Communication*, vol. 31, no. 8, pp. 946–953, 2016.

[48] K. Fatema and J. T. Lariscy, “Mass media exposure and maternal healthcare utilization in South Asia,” *SSM-Population Health*, vol. 11, p. 100614, 2020.

[49] F. Smaldone, A. Ippolito, and M. Ruberto, “The shadows know me: exploring the dark side of social media in the healthcare field,” *European Management Journal*, vol. 38, no. 1, pp. 19–32, 2020.

[50] Q. Yan, Y. Tang, D. Yan et al., “Impact of media reports on the early spread of COVID-19 epidemic,” *Journal of Theoretical Biology*, vol. 502, p. 110385, 2020.

[51] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” *Journal of Chemical Theory and Computation*, vol. 11, pp. 105–147, 2015.

[52] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, “Cost-effective outbreak detection in networks,” in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 420–429, San Jose, California, USA, 2007 Aug 12.

[53] A. Goyal, W. Lu, and L. V. S. Lakshmanan, “CELF++: optimizing the greedy algorithm for influence maximization in social networks,” in *WWW '11: Proceedings of the 20th international conference companion on World wide web*, pp. 47-48, Hyderabad, India, 2011.

[54] L. Guo, D. Zhang, G. Cong, W. Wu, and K. L. Tan, “Influence maximization in trajectory databases,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 3, pp. 627–641, 2017.

[55] S. Stein, S. Eshghi, S. Maghsudi, L. Tassiulas, R. K. E. Bellamy, and N. R. Jennings, “Heuristic algorithms for influence maximization in partially observable social networks,” in *CEUR Workshop Proceedings*, vol. 1893, pp. 20–32, Melbourne, Australia, 2017.

[56] B. Wilder, N. Immorlica, E. Rice, and M. Tambe, “Maximizing influence in an unknown social network,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, pp. 4743–4750, New Orleans Louisiana USA, 2018.

[57] A. Zareie, A. Sheikhahmadi, and M. Jalili, “Identification of influential users in social networks based on users' interest,” *Information Sciences*, vol. 493, pp. 217–231, 2019.

[58] P. Yang, X. Liu, and G. Xu, “A dynamic weighted TOPSIS method for identifying influential nodes in complex networks,” *Modern Physics Letters B*, vol. 32, no. 19, 2018.

[59] Y. Yang, L. Yu, Z. Zhou, Y. Chen, and T. Kou, “Node importance ranking in complex networks based on multicriteria decision making,” *Mathematical Problems in Engineering*, vol. 2019, Article ID 9728742, 12 pages, 2019.

[60] J. F. Robles, M. Chica, and O. Cordon, “Evolutionary multiobjective optimization to target social network influentials in viral marketing,” *Expert systems with applications*, vol. 147, p. 113183, 2020.

[61] A. Karczmarczyk, J. Jankowsk, and J. Watrobski, “Multi-criteria approach to viral marketing campaign planning in social networks, based on real networks, network samples and synthetic networks,” in *2019 Federated Conference on Computer Science and Information Systems (FedCSIS)*, vol. 18, pp. 663–673, Leipzig, Germany, 2019.

[62] E. Mohamad, A. L. Ahmad, S. Mohamed Salleh, and U. K. Wan Sulaiman, “Breaking the traditional communication flow: exploration of social media opinion leaders in health,” *SHS Web of Conferences*, vol. 33, article 00044, 2017.

[63] S. T. Chew, E. Mohamad, and S. M. Salleh, “The quality of health parasocial opinion leaders on social media. A literature review,” *E-BANGI Journal*, vol. 16, no. 7, pp. 1–21, 2019.

[64] S. Saraswathi, A. Mukhopadhyay, H. Shah, and T. S. Ranganath, “Social network analysis of COVID-19 transmission in Karnataka, India,” *Epidemiology and Infection*, vol. 148, pp. e230–e310, 2020.

[65] T. Hossain, *COVIDLies: detecting COVID-19 misinformation on social media*, University of California, Irvine, 2020.

[66] P. Maheshwari and R. Albert, “Network model and analysis of the spread of COVID-19 with social distancing,” *Applied network science*, vol. 5, no. 1, 2020.

[67] K. Nagarajan, M. Muniyandi, B. Palani, and S. Sellappan, “Social network analysis methods for exploring SARS-CoV-2 contact tracing data,” *BMC Medical Research Methodology*, vol. 20, no. 1, pp. 1–10, 2020.

[68] A. Karaivanov, “A social network model of COVID-19,” *Plos one*, vol. 15, no. 10, article e0240878, 2020.

[69] A. Wong, S. Ho, O. Olusanya, M. V. Antonini, and D. Lyness, “The use of social media and online communications in times of pandemic COVID-19,” *Journal of the Intensive Care Society*, vol. 22, no. 3, pp. 255–260, 2021.

[70] Ž. Zaplotnik, A. Gavrić, and L. Medic, “Simulation of the COVID-19 epidemic on the social network of Slovenia: estimating the intrinsic forecast uncertainty,” *PLoS One*, vol. 15, no. 8, article e0238090, 2020.

[71] M. Hung, E. Lauren, E. S. Hon et al., “Social network analysis of COVID-19 sentiments: application of artificial intelligence,” *Journal of Medical Internet Research*, vol. 22, no. 8, article e22590, 2020.

[72] V. Pandey, N. Misra, R. Greeshma et al., “Techno trend awareness and its attitude towards social connectedness and mitigating factors of COVID-19,” *Frontiers in psychology*, vol. 12, article 637395, 2021.

[73] V. Pandey, A. Astha, N. Mishra et al., “Do social connections and digital technologies act as social cure during COVID-19?,” *Frontiers in psychology*, vol. 12, article 634621, 2021.

[74] S. Azad and S. Devi, “Tracking the spread of COVID-19 in India via social networks in the early phase of the pandemic,” *Journal of Travel Medicine*, vol. 27, no. 8, pp. 1–9, 2020.

[75] W. Jo, D. Chang, M. You, and G. H. Ghim, “A social network analysis of the spread of COVID-19 in South Korea and policy implications,” *Scientific Reports*, vol. 11, no. 1, pp. 1–10, 2021.

[76] S. F. Tsao, H. Chen, T. Tisseverasinghe, Y. Yang, L. Li, and Z. A. Butt, “What social media told us in the time of COVID-19: a scoping review,” *The Lancet Digital Health*, vol. 3, no. 3, pp. e175–e194, 2021.