Computing and Informatics, Vol. 42, 2023, 1428–1444, doi: 10.31577/cai_2023_6_1428

NOVEL TECHNIQUE OF HEALTHCARE RECORD INDEXING AND RECOMMENDATION BASED ON TRENDING QUERIES IN SOCIAL MEDIA

Maruthavani Elankannan,
Shantharajah Sinnasamy Periyasamy*

School of Computer Science Engineering and Information Systems
Vellore Institute of Technology, Vellore, 632 014, Tamil Nadu, India
e-mail: shantha_raja78@outlook.com

Abstract. Recommendation of services and applications based on user-data analytics is the most common approach to understanding user requirements. In this article, a novel technique for user recommendation is proposed and validated. The technique uses a Twitter Application Programming Interface (API) handle-based dataset for evaluating and computing the recommendations. The technique uses an open platform Graphical User Interface (GUI) for keyword categorization and building a reliable support system for query analysis. API driven queries from Twitter are cross-validated with labeling techniques and trending hashtags. Typically, the defined tweets are validated to build a Healthcare Record Indexing (HRI) data structure. The HRI is used to support the decision-making and recommendation of services of various healthcare applications and tweets. The technique has trained 750 datasets of categorized clusters with 150 000 tweets (dynamic) datasets from Twitter API. The technique has performed 92.68 % in accuracy and 91.72 % in sensitivity of given datasets.

Keywords: Social media, Twitter analysis, healthcare tweets, sentiment validation, hashtags

* Corresponding author

1 INTRODUCTION

Social media is the most advantageous and widely available platform with global users. The social media platforms connect the users via multiple attributes and features and has been developed on a common agenda of providing a collective networking environment. The user's count has tremendously increased over time and the scope of social media has reached beyond trivial networking and communication. By maintaining consistent outreach, providing patients with handy mobile options, encouraging favorable reviews, expanding the offerings, and remembering to schedule follow-up assignments, the user count has tremendously improved in the course of time. It assists in growing brand awareness, which implies more inbound traffic, increased search engine results, higher conversion rates, and better customer satisfaction. And these are all the things a firm is anticipating from their marketing efforts. It keeps the responses brief and always refers back to them instead of giving lengthy responses that the other person is likely not interested in. The social media application provides a wide range of features and processes among which the users have the privilege to share and discuss multiple topics and subjects such as social eHealth, tweets on medical emergencies and much more. The process of recommendation and computing data shares and tweets can be optimized and a pattern can be evaluated for reliable decision making. Maintaining and managing data across the organization and ensuring the greatest effectiveness of the database management system are the responsibilities of specialists known as database administrators. A database administrator plans, installs, manages and watches over data management systems while guaranteeing their security, quality, and consistency. If access is made available, multiple users may access the data stored in the database management system. Health service providers and other stakeholders in the healthcare sector will be able to provide patients with more accurate and insightful diagnoses, personalized treatment, monitoring of the patients, preventive medicine, and support of medicine by deciphering this maze of given association rules, patterns, and trends.

In the current media state, the stakeholders of larger group discussions and processing are centric to the users. These users are communicating with a multilayered self-built networking framework with tweets and shares of information. Data from public healthcare surveillance are provided and interpreted to help with illness prevention and control. Clear objectives should be set for surveillance of a disease or other health issue in order to accomplish this goal. To discuss problems with health care policy and practice, encourage healthy habits, interact with the public, and instruct and communicate with patients, caregivers, students, and coworkers. The categorization of information is based on labeling and hashtags on subject lines. The recommendation framework must be developed based on these labeling tags to extract the interdependencies from multiple users. The User to User (U2U) communication is based on the hand-shake rule in accepting and acknowledging the data and tweets shared via the approved channel of social media network. These users are interconnected and form a group of clusters with each user node represented as a root or parent node. Hospital records, patient medical records, test findings, and Internet of things-enabled devices are some of the big data sources used in the healthcare sector. Big data pertinent to public healthcare is also produced in great quantities by the biomedical research field. It will assist in the support of clinical treatment decisions, the delivery of real-time alerts to healthcare professionals, the provision of in-depth insight into a patient's health and treatment, and the improvement of patient disease risk identification speed and accuracy.

The process of communication and reliability addition to the social media is blocked via irregularity monitoring. In those scenarios the user information is streamlined and abbreviated via multiple optimizing algorithms. In the existing state, the social media users can avail dynamic services and recommendations from multiple tweets and layered information extraction process. It provides a user's approach recommending to leap across beneficial information. In this research the gap of providing a reliable recommendation is a novel technique of record indexing. The Healthcare Index provided using Queries (HIQ) is proposed and thus provides a sustainable recommendation framework for users via eHealth queries. The proposed approach has redefined the process of recommendation of eHealth record and indexing scenario via existing labeling technique.

The manuscript is organized with a brief introduction on social media application and platform with the research setbacks and challenges in Section 1. The current literature reviews and survey-based observations are made in Section 2, followed by the proposed methodology in Section 3. Section 4 and 5 are briefed on evaluation scenarios of computing the data-tweets via Twiter APIs. The proposed technique's recommendation approach is defined and justified in Section 6 followed by the experimental setup and observations from the defined technique in Section 7. The manuscript is concluded with the observation remarks and justification from the proposed technique.

2 LITERATURE SURVEY

Social media is associated with multiple research ventures and venues with reference to sentimental analysis, user-user networking protocols, social media network standards and much more. In the current survey, the scope is restricted to build and provide recommendation systems via tweets and shares extracted from Twitter APIs [1]. Doctors and nurses are less likely to make errors by updating charts in real time. They are more adept at adhering to procedures, reading patients' medical records, and giving them the right care. Clinicians' time management is improved by this efficiency. The APIs from Twitter has retained a way-exit to provide real-time tweets and treading hashtags from exiting topic of search. These tweets are used in justifying the sentimental analysis as discussed in [2]. The use of supervised machine learning algorithms is one of the most widely used approaches for sentiment analysis. To connect to Twitter, we begin by requesting the Twitter key. Wondering, now that we have imported all the required modules, we can begin the coding portion. The GUI is then created using Tkinter and Create a function to extract data from user-provided input. From time to time, concerns raised in using these tweets from academic research and the challenges in providing a reliable solution for real-time justification [3]. The scope of API based tweets and academic challenges are discussed in optimizing and customizing informatics analysis. The term "Requests per 100 seconds per user" refers to a comparable quota in the API Console. It can be changed up to a maximum value of 1000 and is initially set at 100 requests per user per 100 seconds. Nevertheless, each user is only permitted to submit a maximum of 10 requests to the API each second. Seeing the notification "rate limit exceeded" is not the end of the world. It simply means that until the hour is up, Twitter will not post any updates. Up until the new hour begins and the rate restriction is reset, some parts of Twitter will appear to be frozen. The recommendation models and techniques for Twitter based API system is customized for user profile optimization and grouping [4]. An application programming interface (API) is a set of routines, protocols, and tools for building software applications. An API expresses a software component in terms of its operations, inputs, outputs, and underlying types. These techniques have defined and developed a labeling approach for user grouping and clustering based on primary parameters. The recommendation model can further justify the subject of news and user new consumption ratio to provide valid recommendations. The scope of recommendation models has been extended to HIV tweets user driven recommendation [5, 6]. These recommendations are considered the primary approaches for classifying tweets based on eHealth services. The healthcare system provides four main service categories: rehabilitation, disease prevention, diagnosis and treatment, and promotion of good health. The Twitter data logs are extracted and built using a dynamic Multi Objective Optimal Medical (MooM) Datasets [7, 8, 9, 10]. The MooM dataset provides a customized approach of filtering API driven application data streams. These claims are justified in [11] with series of supports and training provided in API driven data logs. Your IP address, browser type, operating system, referring website, pages viewed, location, mobile carrier, device information (including device and application IDs), search keywords, or cookie information could all be included in this log data on the projects. The process assures the data is not manipulated in larger extent and hence originality is preserved during the process of computation. Process synchronization, often known as synchronization, is the method an operating system uses to manage processes that share the same memory space. By limiting the number of processes that can modify the shared memory at once using hardware or variables, it helps ensure the consistency of the data.

With the outbreak of global CoVID-19 pandemic and an unprecedented lockdown, the users have found a way to connect via social media platforms and the trend of outreaching medical assistance and support via these platforms can be recorded in early 2013 by [12] and justified via collection of speech signals from multiple online platforms recording [13]. The works [14, 15, 16] have included a structural representation and roles of Twitter in CoVID-19 era. The scope of validation is limited to various constraints and parameters of Twitter and its operational principles. Thus, the recommendation framework can be included in the platform to provide reliable and consistent decision support for healthcare facilities.

3 PROPOSED METHODOLOGY

The proposed technique is built to provide a reliable and consistent recommendations of healthcare tweets via indexing and labeling. Information such as your username, any email or phone numbers connected to your account, and information on the process of creating account. Since millions of users under a social media platform are active and have privilege to tweet irrespective of subject line. The proposed technique provides a centric indexing approach for grouping and categorizing the tweets based on healthcare services. The proposed technique is shown in Figure 1, the technique extracts dynamic datasets from active Twitter APIs and stores under a computational buffer defined under Hadoop framework. The processing technique that helps to extract the dynamic datasets are as follows: incremental batch extraction, incremental stream extraction, and full extraction. With a full extraction, all of the data from the source system is taken out and loaded into the destination system. The initial keywords are extracted and processed via a keyword search technique. Users can learn more about the needs of your audience by conducting keyword research. Then you can write about that or produce content based on it. The keywords are user driven and has customized via hashtags to provide effective and dynamic results. A hashtag generator is a tool that creates pertinent hashtags automatically for a specified keyword or phrase. In order for the relevant users to find your material on social media, hashtags are crucial. The keyword is grouped in primary archives of Healthcare Indexing Queries (HIQ) from dynamic API data stream and has relatively extended the operations of labeling and classifications for subclustering the tweets driven by a defined User Interface (UI). Using keyword research tools, innovative research techniques, YouTube and Google suggestions, multiple metrics analysis, keyword length difficulty analysis, and seasonality analysis are some of the suitable process for solving the keyword search techniques. The HIQ tweets include record time, status and intervals of operations from Twitter server. The record log aims to project and develop an interfacing customization of tweets. Further the customizations are classified and labelled via hashtags and indexing parameter of HIQ. The dictionary needs to be expanded with new words, and postings lists for already existing words need to be revised. Dynamic indexing is the easiest method for accomplishing this, as it involves periodically rebuilding the index from scratch. Bibliographic and database indexing, genealogical indexing, geographic indexing, book indexing, legal indexing, periodical and newspaper indexing are a few examples of indexing types. The process is termed as Tweet Mapping (TM). On successfully mapping the archives, the hashtags and phrases will be shown on an interactive hashtag map by Tweetsmap, where you may zoom in for more information. Where links, images, quotes, retweets and questions are the main types of tweet mapping, the Twitter API data stream is paused as it restricts the auto-updating of tweets to perform further recommendation operations. The recommendation process includes validation of HIQ with respect to the API driven priorities and attributes of tweets.

The recommendation is provided based on regional parameters such as trending queries and HIQ stacks. Making or executing something without prior planning while utilizing whatever is available is known as improvisation. The standard records' paperwork enables additional providers to comprehend the patient's background so they can keep giving each person the best care feasible. The recommendation building process assures the improvisation of tweet recommending quality of healthcare records. Based on predictions and recommendations for the patients, the recommender system uses predictive analytics. Where patient engagement is the main focus, patient outcomes are analyzed, and connections to or collaborations with other institutions are made in order to improve the quality of records in healthcare system. The process of organizing medical records and data into a searchable, userfriendly system is known as healthcare record indexing. In order to analyze user data and prescribe medication, data processing is employed in healthcare.

Relational databases can be used to monitor a patient's care in the form of treatments, their results, and important signs of their current condition. The proposed technique is further justified using a mathematical representation in successive sections. For the analysis and design of control systems, the mathematical models are helpful. Finding the output after knowing the input, and the mathematical model is what control system analysis entails. Whereas, in the design of control systems, linear ordinary differential equations with constant coefficients in the time do main and transfer functions obtained from time domain descriptions via Laplace transforms in the frequency or transform domain are the most frequently used mathematical models of the behavior of interest.

4 SCENARIO EVALUATION

The coordination of information is based on the interference and attribute mapping from the API driven Twitter ecosystem. The process is standard and carries authenticated requirements and data coordinations. The process of determining a user's identification is called a uthentication. It is the system that links a set of identifying credentials to an incoming request. The most popular types of authentication include Symmetric-Key Authentication, Password Authentication Protocol (PAP), Authentication Tokens, and Biometric Authentication. Typically, the information sent and requested from API interface is required via primary server exchange. The most popular and extensively used methods in APIs and websites are GET requests. Simply, data retrieval from a server at the provided resource is accomplished via the GET technique. Simple browser viewing of data obtained from an API end-point is the most basic method of accessing information. Using a set of definitions and protocols, APIs are techniques that let two software components communicate with one another. Any request made to the Twitter API should always have the oauth version argument set to 1.0. The validation of information exchange is reported in Figure 2 for ease in understandability. The primary objective of keyword-based Twitter extraction is to realign and configure the requesting servers via authentication servers. Processes that control access to a network, application, or system are managed by an authentication server. Users can access numerous different systems after just one authentication. One of the most crucial responsibilities when working with text is keyword extraction because it allows users to determine whether the text is worthwhile reading more rapidly. An automated method for finding and extracting the words that best characterize a document's subject is called keyword extraction. It aids in distilling a text's information and identifying its major themes. Consider the queries of keyword with the following scenario.

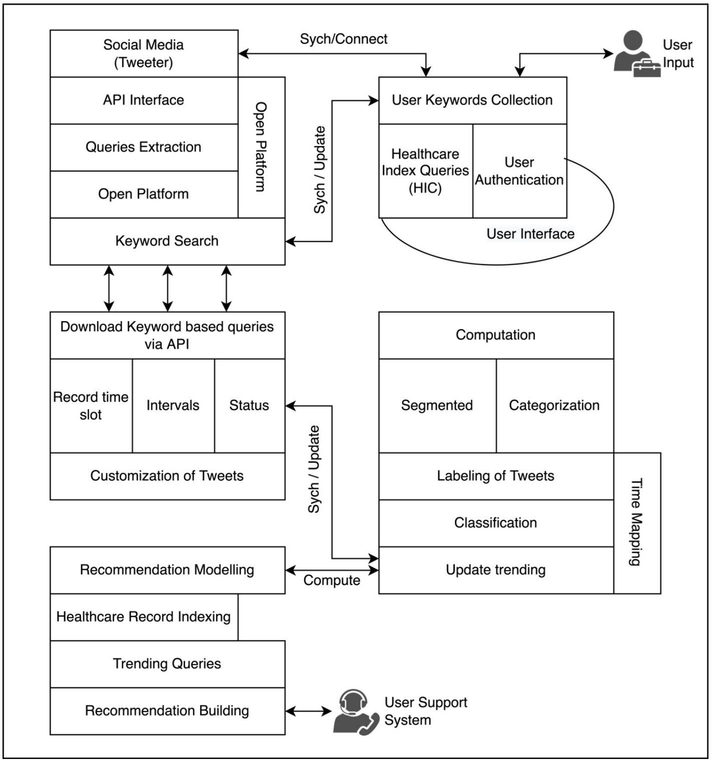

Figure 1. Proposed system architecture and representation diagram

Definition 1. Consider the health keywords under an open search user-driven API. The userset $\Delta U \Rightarrow \|Q\|_0^n \to \lim(H)$. Where $\|Q\|$ are the user cases under open-order of keyword collection with H healthcare keywords provided in a direct access-set.

The process of collecting and aligning dataset of keywords is based on direct set H as $H \to H_1, H_2, H_3 \dots H_n$ such that $\forall H_i \Rightarrow (\Delta Q)_0^n$ at each $\Delta H \Rightarrow \Delta Q$ for ease in mapping and relationship validation. The process in the considered case-study is reflected with limited keyword and queries coordinated. Case studies can be classified as intrinsic, instrumental, or collective. Although quantitative methods are occasionally utilized, qualitative methods are typically used in case study research designs. Choose a case, create a conceptual framework, gather your info, explain and evaluate the case. Case studies have several limitations, including a lack of scientific rigor and a lack of a solid foundation for extrapolating findings to a larger population. Hence, the process of $\Delta H_i \Rightarrow H_i \subseteq \Delta T$ where ΔT is the trending indexing array of keywords on a given time interval t. Hence the validation of ΔH with ΔT is projected and represented.

$$\mathrm{Val}(\Delta H) \Rightarrow \left[\int_0^\infty \delta(t) \oplus \left\{\lim_{n\to\Delta T}\left(\sum_{i=1}^{n}\frac{\delta\left(\Delta H_i\right)}{\delta t}\right)\right\}\right]. \tag{1}$$

On responsive validation with ΔT trending variance, the Equation (1) is rerepresented as Equation (2).

$$\mathrm{Val}(\Delta H) \Rightarrow \left[\lim_{n\to\Delta T}\left(\sum_{i=0}^{n}\prod_{j=i+1}^{\Delta T}\frac{\delta\left(\Delta H_i\right)\oplus\delta\left(\Delta T_j\right)}{\delta t}\right)\times\Delta U\right], \tag{2}$$

$$\mathrm{Val}(\Delta H) \Rightarrow \left[\frac{\delta(\Delta U)}{\delta t} \cong \left\{\lim_{n\to\Delta T}\left(\sum_{i=0}^{n}\prod_{j=i+1}^{\Delta T}\frac{\delta\left(\Delta H_i\right)\oplus\delta\left(\Delta T_j\right)}{\delta t}\right)\right\}\right]. \tag{3}$$

According to Equation (3), the relevance of user ΔU on open Twitter is based on the parallel influence of ΔH provided under the computational time. The inference drawn in Definition 1 is to provide a reliable set of keywords trending in Twitter in terms of healthcare surveillance. Typically, the process is aligned and coordinated with administrator privilege driven query management. With the creation of pointers to the locations of data within a database, indexing speeds up querying of columns where data structure that accelerates data retrieval. By reducing the number of disk accesses needed when a query is completed, indexing helps a database perform better.

Definition 2. Proposing Healthcare Indexing Queue (HIC) for dynamic mapping of queries under open ΔH coordination with $\forall H_i \Rightarrow \Delta T \oplus \Delta U$ and $\Delta H_C \Rightarrow \Delta H$ on given time interval t.

Consider the dynamic tweets based on trending ΔT with a calibrated rating order O_f such that $\Delta T \Rightarrow O_f$ at instance of API collaboration. Typically, the formulation of dependency between the tweets and trending keywords is computed. The defined process is elaborated in secondary mapping of data attributes from user-defined keywords ΔH such that $\forall \Delta H_i \Rightarrow \Delta T \subseteq O_f$. To organize test steps under a single term, we use user-defined keywords. So, we may write the keyword only once rather than repeating it across numerous lines. When numerous test cases employing the same test steps are present, this is quite helpful. Similarly, descriptive data about a user entry is stored in attributes. Each attribute has a label, one or more values, and a standard syntax for the kind of data it can hold in its value. It is used to keep confidential information specific to the page or application. The Data Attributes are primarily divided into two parts: Attribute Name: Must begin with the prefix "data-", be at least one character long, and not contain any capital letters. Value for the attribute: Any string. This relationship improves the quality of tweet harvesting from the global server. The process of tweets harvesting C is represented in Equation (4).

$$C = \lim_{\Delta n \to \infty} \left(\frac{\delta(H) \cup \delta\left(\Delta T \subseteq O_f\right)}{\delta t} \right) \times \Delta H. \tag{4}$$

On calibration function

$$C = \lim_{\Delta n \to \infty} \left(\sum_{i=1}^{n} \prod_{j=i+i}^{\Delta T} \left\{ \frac{\delta(H) \cup \delta\left(\Delta T \subseteq O_f\right)}{\delta t} \right\} \right). \tag{5}$$

In Equation 5, the influence of ΔH externally as a dependency component is eliminated since the iterative approach of reducing the dependency on directly defined set ΔH with extension to retrieve data elements O_f on trending ΔT.

4.1 Computational eHealth Tweeting and Labeling

According to Definitions 1 and 2, process of validation is defined and calibrated via the tweets downloading and extracting approach. Typically, the tweets are downloaded at local computational servers for labeling and defining the categories. By selecting your account from the menu after clicking the more symbol in the navigation bar, you may access your account settings. To get an archive of your data, click Download. Under Download an archive of your data, enter your password, and then click Confirm. Where, Twitter Analytics provides you with your followers' fundamental information. You can more effectively approach your audience by being aware of the demographics. The aim of this section is to provide a reliable labeling framework for categorizing eHealth tweets and services. The guidelines for managing and carrying out health and social care research that take into account legislative requirements and other standards are outlined in the framework for health and social care research. The six "aims" of the IOM framework for highquality healthcare are equity, efficiency, safety, effectiveness, and timeliness. Several health professionals have accepted this approach. The issues that should be covered in an extensive set of quality measurements are generally acknowledged to be covered by this model. The labeling L is defined as Equation (6).

$$L = \int_0^n \sum_{i=1}^{\Delta T} \left(\frac{(\Delta H_i \varrho_\Delta (\Delta U)}{\Delta T} \right) \Rightarrow L_1, \tag{6}$$

$$L = \left[\int_0^n \sum_{i=1}^{\Delta T} \left(\frac{(\Delta T_i \subseteq O_f) \cup (\Delta H_J)}{\Delta T} \right) \times \Delta U \right] \Rightarrow L_2. \tag{7}$$

According to Equation (7), the representation of labeling is computed via L_1 and L_2 as shown. The defined parameter of labeling vector is functionally dependent on the model of tweets extracted from global server. The categorization of labels is redefined and processed with a synchronizing approach as $\Delta L = (L_1 \cup L_2) \cap L$ at an individual progressive iteration. The process of labeling assures the data alignment. The act of detecting unlabeled raw data (such as photos, text files, videos, etc.) and adding one or more insightful labels to give context so that a machine learning model may learn from it is known as data labeling. On completion of alignment, the categorization is computed. Consider the categorization as $F_L = \forall L \subseteq \Delta C$ shown in Equation (8).

$$F_L = \Delta U \oplus \left[\lim_{\Delta n \to \infty} \left(\sum_{i=1}^{n} \sum_{j=i+1}^{\Delta T} \left(\frac{\delta\left(L_1\right) \cup \delta\left(L_2\right)_j}{\delta t} \right) \cap \left(\frac{\delta(L)}{\delta t} \right) \right) \right]. \tag{8}$$

The process assures the computational difference of L_1 and L_2 with respect to user derived keywords. Consider the user categories of keywords K, the parametric differences of trending are T_1, T_2, T_3 such that $\forall K \Rightarrow T_1, T_2, T_3$ and hence the labeling is formatted and categorized for ease in labeling. The fundamental understanding is drawn with API based accessing of tweets from social media platforms. The progressive computation is resultant of dependency variables based in labeling and categorization.

Hence the defined labels are paradigms of future evaluation and construction of healthcare record indexing (HRI). The purpose of the proposed framework is to extract and evaluate the tweets based on trending labels and categories. The recommendation function is based on computation F_L parameters with user-driven requirements, as shown in Equation (8). The recommendation is based on primary and secondary processing dataset of Twitter API.

4.2 Recommendation Model

Based on computational validation of categories F_L with accordance to levels L_1 and L_2 the recommendations string is processed. The recommendation model produces the HRI with dataset processing. The training queries are parallel in computation for rectifying and generating the threshold as shown in Equation (9).

$$\Delta Thr = \left[\frac{\delta(\Delta K)}{\delta t} \cup \lim_{n \to \infty} [API(\Delta K)] \times \Delta Thr \right]. \tag{9}$$

An API-driven architecture enables connected components and services to remain flexible, processing requests and ensuring corporate systems operate without interruption. Cluster analysis is generally used to create groups or clusters while making sure that the observations are as similar as possible within each group. Strategic planning is made possible by the use of big data in healthcare since it provides better understanding of people's motivations.

Where ΔThr are the threshold queries of keywords from user driven and API are based keyword extraction. The process of extracting API driven keywords is collected in buffer server S_f with each server interconnected to form a coordinated matrix M of relationship between the labeled values of tweets under eHealth category, as shown in Equation (10).

$$S_f = \int_0^n \Delta Thr \Rightarrow \left[\frac{\delta\left(M_K\right)}{\delta t} \cup \frac{(L_1 \cap L_2)}{F_L} \right], \tag{10}$$

where the threshold values of each server L_1 and L_2 are fragmented and processed with S_f under a naive based arrangement for probability extraction in providing the most relevant recommendations from the keywords. The process of server S_f optimization is based on ΔThr factor and matrix vector M. This is to provide a base-line server for computing optimized recommendations of trending keywords.

Consider R as the recommendation variable associated with optimized recommendation server S_f. The influencing parameter of recommendation R is bound to cluster multiple values of keywords, as shown in Figure 2. The recommendation process is further optimized with current trending parameters to extract and rebuild the categorization in the series of L_1 and L_2. In this context, the extracted parameters of tweets under trend ΔT_t is subjected, as shown in Equation (11).

$$\Delta T_t \Rightarrow \lim_{n \to \Delta T} \left[(\Delta Thr)_0^\infty \cup (\Delta S_f)_0^K \right], \tag{11}$$

$$\Delta T_t \Rightarrow \sum \left(\lim_{n \to \Delta T} \left[(\Delta Thr)_0^\infty \cup (\Delta S_f)_0^K \right] - F_L \right). \tag{12}$$

The possibility of Twitter trending extraction can dependent on multiple references with mapping on a single keyword. Typically, the orientation is subjected to eliminate the redundancies. The process of ΔT_t as by Equation (12) is re-subjected for recommendation alignment and calibration, as shown in Equation (13).

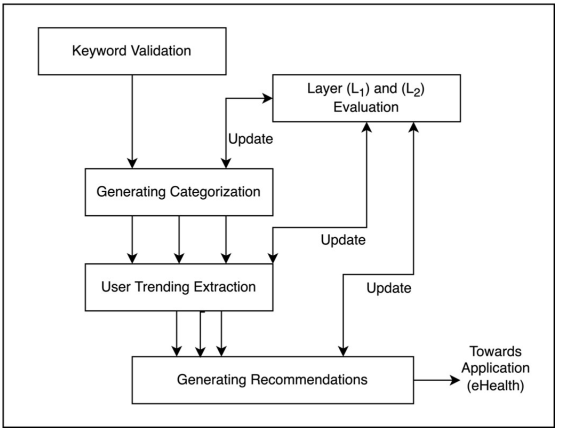

Figure 2. Recommendation evaluation process

$$R = \Delta Thr \left(\sum_{i=1}^{\infty} \prod_{j=i+1}^{\Delta K} \left(\frac{\delta \left(M_K \right)_i}{\delta t} \times \frac{\delta \left(\Delta T_t \right)_j}{\delta t} \right) \right), \tag{13}$$

$$R = \Delta Thr \left(\sum_{i=1}^{\infty} \Pi_{j=i+1}^{\Delta K} \left(\frac{\delta \left(M_K \right)_i \cup \delta \left(\Delta T_t \right)}{\delta t} \right) \right) \tag{14}$$

$$\Downarrow$$

$$R = \Delta Thr . \left(\Delta S_f \right)_0^K . \tag{15}$$

According to Equation (14), the back tracked influence is drawn on the references provided. Typically, the formulation of ΔS_f based optimized server creates an ecosystem for reliable recommendations in terms of alignment with co-relationship matrix M and optimized trending tweets ΔT_t, as shown in Equation (15).

5 RESULTS AND DISCUSSIONS

The proposed technique has successfully extracted and processed on dynamic API Twitter based tweets. The proposed technique has limited the tweet selection to healthcare driven tweets and user customization of tweet extraction. At the beginning, Twitter used MySQL as its main data store; from a single instance, the persistence layer expanded to include many clusters. From its start, Twitter has had one of the largest MySQL deployments. There are MySQL clusters there that can handle millions of queries per second on thousands of servers. The process of tweet extraction is shown in Figure 3. The extracted and customized tweets represent the tweet strength and ratio of extraction.

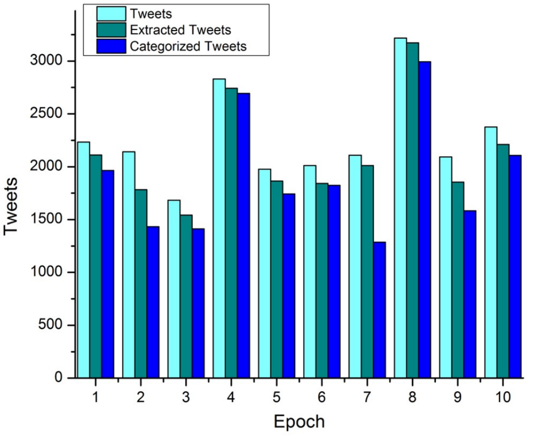

Figure 3. Pre-process representation of tweet evaluation

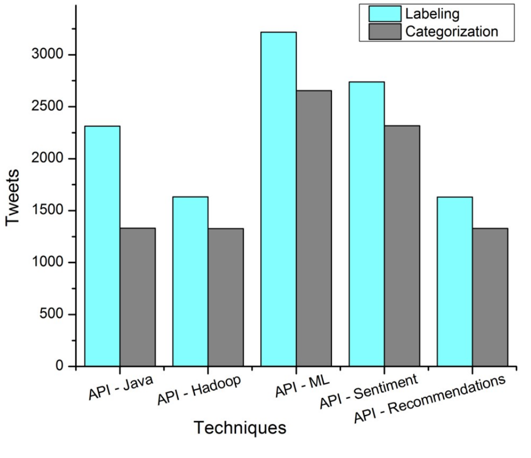

Figure 4. Tweets labeling and categorization of dynamic APIs

The validation of tweets and customization of labels is shown in Figure 4, the labeling process includes the process of labeling and categorization based on API driven systems on JAVA, Hadoop, ML and recommendations. With hashtags, Twitter's main goal is to link its users and enable them to communicate with one another and their followers. It can be a tool for marketing businesses as well as a source of news and entertainment. The API-Sentimental analysis is streamlined and processed for customizing approach of tweets selection. Figure 5 demonstrates

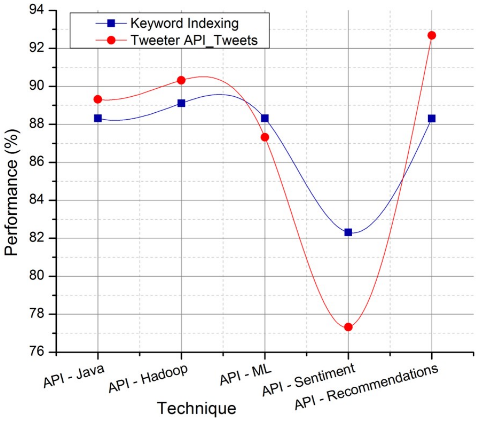

Figure 5. Comparative analysis of API driven tweets

the comparative analysis on API based recommendation and keyword based recommendation. The computational values of keyword search are enhanced with recommendation framework model. Further, a detailed validation of recommendation is processed with API, sentimental and HRI based recommendation (Figure 6). A detailed accuracy validation is shown in Table 1 for understandability.

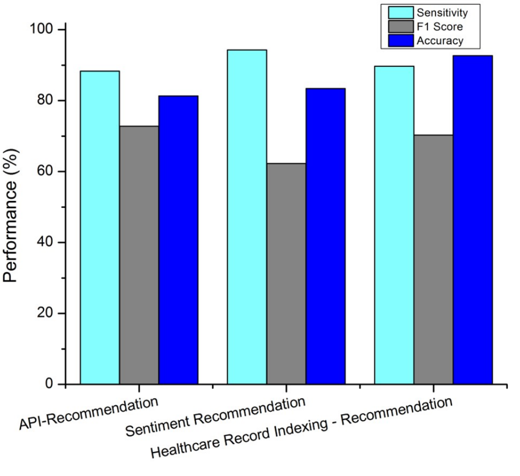

Figure 6. Comparative analysis on recommendation techniques

	Sensitivity	F1 Score	Accuracy
API-Recommendation	88.32	72.81	81.32
Sentiment-Recommendation	94.32	62.32	83.42
HRI-Recommendation	89.67	70.32	92.68

Table 1. Experimental performance validation of proposed technique

6 CONCLUSION

This article discusses providing a reliable recommendation on eHealth services based on Twitter dataset analysis. The technique consists of dynamic datasets via Application Program Interface (API) driven tweet keyword extraction. The technique is validated on 750 classified datasets of medical and eHealth labels. The technique has deployed a novel optimization and labeling approach towards categorization of tweets. The secondary validation includes the authentication via trending tweets to assure stability in the recommendation provided. The proposed Healthcare Record Indexing (HRI) improves the validation by optimizing the server and computational labels in the dataset. The technique has demonstrated 92.68 % accuracy of the dataset (tweets) classification-based recommendation and 91.72 % sensitivity of labeled dataset. In the near future, the technique can be cross-validated using dynamic on-demand classification to extract and streamline healthcare services.

7 DECLARATIONS

Funding: No funds, grants were received by any of the authors.

Conflict of interest: There is no conflict of interest among the authors.

Data Availability: All data generated or analysed during this study are included in the manuscript.

Code Availability: Not applicable.

REFERENCES

[1] Makice, K.: Twitter API: Up and Running: Learn How to Build Applications with the Twitter API. O'Reilly Media, Inc., 2009.

[2] Farsi, D.: Social Media and Health Care, Part I: Literature Review of Social Media Use by Health Care Providers. Journal of Medical Internet Research, Vol. 23, 2021, No. 4, Art. No. e23205, doi: 10.2196/23205.

[3] Ahmed, S. T.—Sandhya, M.—Sankar, S.: TelMED: Dynamic User Clustering Resource Allocation Technique for MooM Datasets Under Optimizing Telemedicine Network. Wireless Personal Communications, Vol. 112, 2020, No. 2, pp. 1061–1077, doi: 10.1007/s11277-020-07091-x.

[4] ANTHEUNIS, M. L.—TATES, K.—NIEBOER, T. E.: Patients' and Health Professionals' Use of Social Media in Health Care: Motives, Barriers and Expectations. Patient Education and Counseling, Vol. 92, 2013, No. 3, pp. 426–431, doi: 10.1016/j.pec.2013.06.020.

[5] AHMED, S. T.—SANDHYA, M.—SANKAR, S.: A Dynamic MooM Dataset Processing Under TelMED Protocol Design for QoS Improvisation of Telemedicine Environment. Journal of Medical Systems, Vol. 43, 2019, No. 8, Art. No. 257, doi: 10.1007/s10916-019-1392-4.

[6] GILL, S.—KINSLOW, K.—MCKENNEY, M.—ELKBULI, A.: Twitter and the Credibility of Disseminated Medical Information During the COVID-19 Pandemic. Vol. 87, 2021, doi: 10.1177/0003134820973391.

[7] TRUPTHI, M.—PABBOJU, S.—NARASIMHA, G.: Sentiment Analysis on Twitter Using Streaming API. 2017 IEEE 7th International Advance Computing Conference (IACC), 2017, pp. 915–919, doi: 10.1109/IACC.2017.0186.

[8] BIRJALI, M.—KASRI, M.—BENI-HSSANE, A.: A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends. Knowledge-Based Systems, Vol. 226, 2021, Art. No. 107134, doi: 10.1016/j.knosys.2021.107134.

[9] KAZAI, G.—YUSOF, I.—CLARKE, D.: Personalised News and Blog Recommendations Based on User Location, Facebook and Twitter User Profiling. Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '16), 2016, pp. 1129–1132, doi: 10.1145/2911451.2911464.

[10] CAMPAN, A.—ATNAFU, T.—TRUTA, T. M.—NOLAN, J.: Is Data Collection Through Twitter Streaming API Useful for Academic Research? 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 3638–3643, doi: 10.1109/BigData.2018.8621898.

[11] ABEL, F.—GAO, Q.—HOUBEN, G. J.—TAO, K.: Analyzing Temporal Dynamics in Twitter Profiles for Personalized Recommendations in the Social Web. Proceedings of the 3rd International Web Science Conference (WebSci '11), 2011, doi: 10.1145/2527031.2527040.

[12] PERSHAD, Y.—HANGGE, P. T.—ALBADAWI, H.—OKLU, R.: Social Medicine: Twitter in Healthcare. Journal of Clinical Medicine, Vol. 7, 2018, No. 6, Art. No. 121, doi: 10.3390/jcm7060121.

[13] LOHMANN, S.—WHITE, B. X.—ZUO, Z.—CHAN, M. P. S.—MORALES, A.—LI, B.—ZHAI, C.—ALBARRACÍN, D.: HIV Messaging on Twitter: An Analysis of Current Practice and Data-Driven Recommendations. AIDS, Vol. 32, 2018, No. 18, pp. 2799–2805, doi: 10.1097/QAD.0000000000002018.

[14] SHARMA, A. E.—MANN, Z.—CHERIAN, R.—DEL ROSARIO, J. B.—YANG, J.—SARKAR, U.: Recommendations from the Twitter Hashtag #DoctorsAreDickheads: Qualitative Analysis. Journal of Medical Internet Research, Vol. 22, 2020, No. 10, Art. No. e17595, doi: 10.2196/17595.

[15] AHMED, S. T.—SINGH, D. K.—BASHA, S. M.—ABOUEL NASR, E.—KAMRANI, A. K.—ABOUDAIF, M. K.: Neural Network Based Mental Depression Identification and Sentiments Classification Technique from Speech Signals: A COVID-19 Focused Pandemic Study. Frontiers in Public Health, Vol. 9, 2021,

Art. No. 781827, doi: 10.3389/fpubh.2021.781827.

[16] Saeed, A. M.—Hussein, S. R.—Ali, C. M.—Rashid, T. A.: Medical Dataset Classification for Kurdish Short Text over Social Media. Data in Brief, Vol. 42, 2022, Art. No. 108089, doi: 10.1016/j.dib.2022.108089.

Maruthavani Elankannan received her M.Eng. degree in software engineering and her B.Eng. degree in computer science and engineering, both from the Anna University, Chennai. Currently, she is the Research Scholar in the School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore. She has contributed 10 research papers. Her current research interest includes analytics, recommender systems. She is also a member of various bodies like the Indian Society for Technical Education (ISTE), the Computer Society of India (AENG).

Shantharajah Sinnasamy Periyasamy received his Ph.D. from the Periyar University, Salem. Currently, he is working as Professor in the Department of Software Systems Engineering, the School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore. He has contributed 60 research papers and worked on several projects. His main areas of research include the Internet of Things, networks and nature inspired optimization techniques and data analytics. He is also a member of various professional bodies like the Indian Society for Technical Education (ISTE), the Computer Society of India (CSI) and the Cryptology Research Society of India (CRSI).