Informatics Bulletin, Faculty of Computers and Artificial Intelligence, Helwan University
Published Online *Vol 7 (2025) Issue 2, July 2025*
(https://fcihib.journals.ekb.eg)

Survey: Depression Detection Using Sentiment Analysis Techniques

Afaf Hussein Abdelrahman[1], Doaa Elzanfaly[1], Mai El Defrawi[1]
1Infromation Systems Dept., Faculty of Computers and Artificial Intelligence, Helwan University, Egypt
afafmseleem@gmail.com, doaa.saad@fci.helwan.edu.eg, mai.eldefrawi@gmail.com

***Abstract*—Depression is a mood disorder that leads to feelings of permanent sadness. It affects people's social life and work. Diagnosing depression is not easy because of the unwillingness to go to a doctor, especially in Middle Eastern societies. Still, social media has facilitated this, and it has become easier for people who suffer from depression to express their feelings. The specialists follow up on their condition, study it, find out the causes of depression, and intervene to try to treat it. The outbreak of the COVID-19 virus has contributed to the rise of depression levels. This is evident in most comments on social media. Therefore, it is necessary to search in this direction. This research provides a literature review on efforts done to detect depression from social media platforms while focusing on both the English and Arabic Languages. A summarization of the findings of this research is provided.**

***Index Terms*—Depression; Machine Learning Techniques; Arabic Sentiment Analysis**

I. INTRODUCTION

HEALTH IS A VERY IMPORTANT PART OF ANY country and the outbreak of any disease causes anxiety and tension among people [1]. Emotion is an important part of our life, and it helps us understand ourselves and others [2]. Depression is a serious mental illness that affects the ability to work, sleep, eat, and have a bad mood [3].

According to the World Health Organization statistics 350 million people suffer from depression [4]. The failure to diagnose depression and not treat it is the main problem in the Arab world, and this is because they consider mental illness a stigma [5].

Social media such as Facebook and Twitter have become very important in the lives of all people. It has become a tool for communicating with family and friends, and it has also become a tool for expressing daily feelings and moods in which One can know a person's psychological state. Social media helps depressed people express their psychological state and mood, and thus the possibility of follow-up and knowledge of their condition by specialists [6].

People resort to social media to express feelings of depression and not go to a psychiatrist for several reasons, including the fear of the idea of depression. It needs treatment, and also the thought formed about the psychological disease in it, as it cannot be disclosed in Arab societies, and this made social media play a vital role. Knowing and following up on depressed people

In the field of natural language processing, sentiment analysis, and opinion mining are among the most active research areas, which analyzes people's opinions, sentiments, and growth of social media such as reviews leading to the growing importance of sentiment analysis. [7]. Fig. 2

Sentiment analysis is very important because it helps company owners improve their work by reviewing and analyzing customer opinions on social media and other sites and then making a quick and accurate decision.

Machine learning techniques are considered one of the most popular techniques used in sentiment analysis which is divided into supervised, unsupervised, and semi-supervised learning techniques. [8]. Fig. 1

Deep learning is a branch of machine learning that helps create theories that allow a machine to learn itself and stimulate nerve cells in the body. Most deep-learning research analyzes large datasets with linear and nonlinear variables.

Sentiment analysis is divided into four types, namely (1) Fine-grained sentiment that analyzes sentences by parts to identify the topic or the target of sentiment, i.e. understanding customers' comments on an item. This type is cost-intensive (2) Emotion Detection Sentiment Analysis, which is used to identify emotions in the text by using Lexicons and machine learning., This type helps the company to understand the reason for customer feelings. (3)Aspect-based sentiment analysis, this type used for one aspect of a product, like, brightness in televisions (4)

Intent analysis, this type is used to predict if a client intends to buy the product or not by understanding the intention of the client.

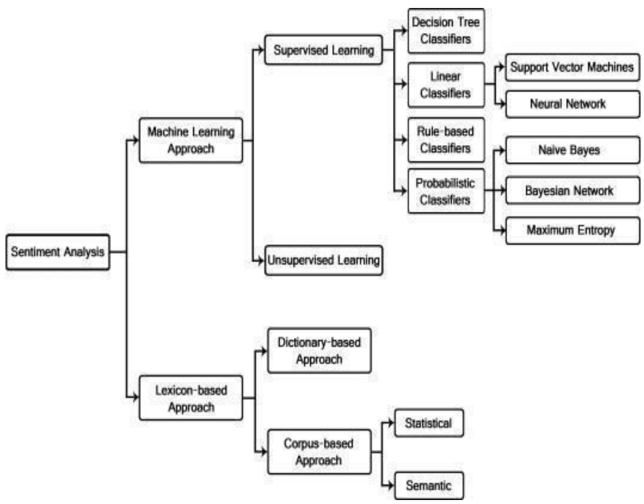

Fig. 1. Machine learning-based techniques [9]

Classification Algorithms: Logistic Regression, used to predict if something happens or not and it shows as Yes or No, Naive Bayes Classifier, determines if the data point is part of a certain category, K-nearest Neighbors used to discover the k closest relatives by using the training dataset, Decision Tree, which belongs to supervised learning used to solve classification problems, Support Vector Machines used to classify data by degree of polarity.

Sentiment Analysis is not an easy matter, and it faces many challenges, including. (1) Subjectivity and Tone, where the subjectivity and objectivity of a text depend mainly on the explicit and implicit. (2) Context and Polarity, this is because of a change in polarity (3) Irony and Sarcasm, sometimes people resort to positively expressing their negative feelings, and this is called irony, and thus it is difficult for the machine to understand the context. Sentiment analysis in Arabic also faces many challenges, including

- Returning the word to its origin.
- Using colloquial words such as those found in the Egyptian dialect.
- Spelling in the text.
- Frequent use of pronouns.
- The use of English quotes in a text written in the Arabic language.
- There are not many available resources for the Arabic language.
- Composition of sentences.

As depicted in Fig 3. Sentiment Analysis Process goes through the following steps: (1) Input Text: collect text data to train the classification model such as social media data from Twitter or Facebook. (2) preprocessing: The data is preprocessed by performing some operations on it, such as removing duplicate words and punctuation marks, which is a very important process to make classification easier, and the steps for processing the entered data are Replace Emotions, Remove stop words, Negation replacement, (3) Feature Extraction:

- Tokenization: Give a character sequence token to deal with them as a group
- Stemming: This process combines words with the same meaning to avoid repetition
- Sentiment Extraction: determine the polarity of the sentence

(4) Feature Selection: Various techniques can be used to select the best group of features that will be used, and then the selected features are used by the machine learning techniques like support vector machines and decision trees (5) Sentiment Classification: identify the opinion to decide if it is positive or negative.

Fig. 2. Taxonomy of Sentiment Analysis Tasks [10]

II. RELATED WORK

Recent research, such as [11] , compares hybrid and ensemble methods for depression detection, finding ensemble models perform better but highlighting areas for improvement. These studies enhance classification with advanced features but are limited by dataset dependency, computational complexity, and neglect of timing, frequency, and repetitive behavior. They also overlook cultural nuances and real-world application validation, focusing only on two classes. While exploring DL architectures, further work is needed on using transformer-based models, managing unbalanced datasets, and expanding the scope for more comprehensive results.

The study in [12] identifies depression by analyzing Twitter users' demographic data and responses to the Patient Health Questionnaire-9. It evaluates five machine learning techniques, with Random Forest achieving the highest accuracy. However, the model relies solely on publicly disclosed data and excludes private or undisclosed user information. Additionally, it is limited to three classes and neglects timing and frequency in its analysis. While the paper presents valuable methods, it faces challenges like scalability, dependence on high-quality datasets, and limited focus on real-time adaptability and algorithmic bias mitigation.

The research in [13] aims to improve lexicon-based methods for detecting depressive text on social media by developing a more accurate classification function. The study introduces an online depression diagnostic system that uses this enhanced lexicon for real-time assessment of depression levels, achieving an F1 score of 74% and precision of 77%. However, the research faces limitations, such as Twitter's API restriction, which limits analysis to 2,400 tweets, and the use of only publicly available data. The system also categorizes depression into two classes

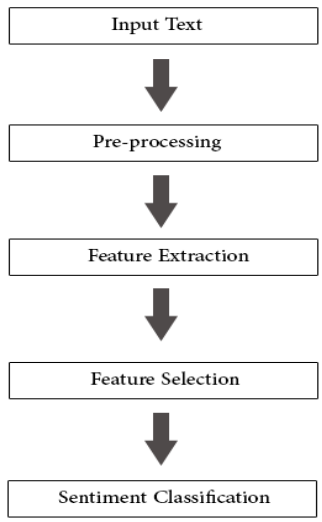

Fig. 3. Sentiment Analysis Process

and does not consider timing or frequency, which could provide deeper insights into depressive patterns. Additionally, the reliance on lexicons may miss regional dialects and cultural variations, limiting accuracy and generalizability. The model also lacks validation on real-world data and does not explore other potentially valuable features like behavioral patterns or physiological data. Lastly, the paper does not address potential biases within the lexicon that could affect the fairness and reliability of the depression detection system across different demographic groups.

The study in [6] introduces the AD prediction model, a real-time tweet-based system for detecting anxious depression with an accuracy of 85.09%. The model uses a 5-tuple feature set, integrating language signals and user posting habits, alongside time and frequency analyses. However, its limitations include reliance on a small dataset of 100 users, a restricted lexicon of 60 words, and limited classification to three classes. These factors reduce its generalizability and fail to account for linguistic and cultural nuances. Ethical concerns about data privacy and consent are also not adequately addressed. Future research could improve lexicon development, explore diverse user bases, and investigate advanced methods like deep learning to enhance prediction accuracy.

The paper [14] introduces methods for identifying depression in Arabic users based on Twitter data. However, the study has several limitations. The dataset may not fully represent the broader Arabic-speaking population, limiting generalizability. The model may also struggle with cultural and linguistic variations in Arabic, as expressions of depression differ across dialects. Additionally, the study relies solely on text data and does not account for important behavioral factors like tweet duration, interactions, or frequency. The lack of temporal analysis and ethical concerns regarding privacy and informed consent further restrict the model's effectiveness and applicability .

Research in [15] investigates the use of natural language processing (NLP) and machine learning (ML) to detect depression in Arabic text. The work employs a lexicon-based approach, creating the ArabDep lexicon to identify depressive symptoms, and a machine-learning-based approach with annotated data by psychologists. These methods aim to capture linguistic markers and predict depression-related expressions with an accuracy exceeding 80%. However, there are several limitations. The reliance on the ArabDep lexicon may overlook regional dialects and cultural variations in Arabic, impacting accuracy. The dataset, focused on a specific forum, may not be representative of broader social media platforms, limiting generalizability. Additionally, the model treats all types of depression as a single class and does not consider temporal patterns or repetitive behavior, which are essential for comprehensive analysis.

The paper [16] develops a model for identifying depression among Arabic-speaking Twitter users by analyzing their tweets. It leverages various machine learning classifiers, including Support Vector Machine (SVM), Random Forest (RF), and Naïve Bayes, and combines them with natural language processing (NLP) techniques for sentiment analysis. The model categorizes tweets into three classes: "depressed," "non-depressed," and "neutral," with the Random Forest classifier achieving the highest accuracy of 82.39%. However, the study has some limitations. It heavily relies on the quality of the dataset, which is limited in scope and size. Additionally, the model does not account for factors such as timing and frequency of tweets, which could provide more insights into depression. Moreover, the study doesn't fully address potential cultural nuances in how depression is expressed in Arabic, which may impact the model's accuracy across different Arabic-speaking communities.

The study [17] explores the use of deep learning and NLP to detect depression in social media content, proposing the "Fasttext Convolution Neural Network with Long Short-Term Memory (FCL)" model to improve text representation. However, it has limitations, such as relying on large datasets, focusing solely on text data, and lacking consideration for timing and frequency. The model does not address behavioral cues or informal language challenges, and future plans include developing an automated system with a chatbot for medication adherence and illness detection.

In [18], the authors propose a method for classifying depression severity in tweets, categorizing them into "None," "Mild," "Moderate," and "Severe" based on linguistic cues from clinical tools like DSM-5 and PHQ-9. While this approach shows promise, it has some limitations. Its reliance on dataset quality affects generalizability, and the model's focus on textual data overlooks other potential behavioral cues. The absence of manual expert supervision in annotating the data could introduce biases, and the informal language of social media presents challenges in detecting subtle depression indicators accurately.

The study [19] presents a novel approach by developing a Weibo-based dataset and employing a Multimodal Hierarchical Attention (MHA) model. The model integrates text, images, and metadata from social media posts, applying attention mechanisms both within and across modalities to improve depression detection accuracy. However, the research has several limitations. It is restricted to two classification classes and does not consider temporal aspects such as timing and frequency of posts, which are critical for understanding depressive behavior. Additionally, reliance on multimodal data poses challenges, as not all social media content includes rich text or visual information. The model's generalizability is further constrained if the dataset lacks linguistic and cultural diversity. Moreover, ethical concerns about privacy and user consent in utilizing social media data for mental health detection are not comprehensively addressed, raising concerns about its real-world application.

Table I presents a comparative summary of recent research papers. The comparison focuses on three main aspects: the number of classes used to classify depression severity, the consideration of timing and frequency of posts in depression detection, and the language adopted in each study. It is observed that most studies concentrate on the English language, while only a few explore the Arabic language. Similarly, limited attention has been given to the timing and frequency of social media activity in the context of depression analysis

III. MATERIALS AND METHODS

A. Sentiment Analysis Approaches

1) Data Collection

There are many ways to collect data, including:

- PHQ-9**:** The nine-item Patient Health Questionnaire is a depressive symptom scale and diagnostic tool introduced in 2001 to screen adult patients in primary care settings. The instrument assesses for the presence and severity of depressive symptoms and a possible depressive disorder. The data derived from the PHQ-9 are the survey response date and the depression assessment score of each participant.

- CES-D**:** The CES-D scale is a short, self-reporting scale that contains 20 questions. It is designed to measure depressive symptomatology.

- Personal information questionnaire: Twitter IDs and demographic data, including gender, age, weight, education, congenital condition, career, income, number of family members, self-couple status, and parent's marital status, are gathered from this questionnaire.

- Twitter API: is used to collect Twitter users' information from the Twitter page consisting of name, date of account creation, account verification status (verified or not), language, description, and tweet count. For each user, tweets are fetched, with the date and time of post, number of retweets, hashtags, and mentioned users.

- using a dictionary of depressive phrases. For example, which means, "I am diagnosed with depression."

Nafsany**:** It is an online platform that is aimed at encouraging people from Arab countries to publicly share their stories and to obtain feedback.

Reddit**: U**sers who are depressed as well as those who are not depressed posting on this social media platform.

eRisk Dataset: This dataset is collected from the eRisk (Early Risk Prediction) forum [The eRisk collection contains posts from depressed and non-depressed 4498 users.

2) Data Preprocessing

Cleaning and filtering the data to prepare it for feature extraction is known as pre-processing. The process includes:

• Removing numeric and empty texts, URLs, mentions, hashtags, non-ASCII characters, stop-words and punctuations

• Tokenization of tweets is done using the TreebankWordTokenizer of Python Natural Language Toolkit (NLTK) to filter the words, symbols, and other elements called tokens(Loper,& Bird, 2002) The tokens are converted to lowercase.

• Slang and emojis should be replaced with their descriptive text. As the Internet is an informal way of communication, the use of slang and emojis is a common practice. These might clarify the situation and heighten the emotion connected with it.

• Stemming to reduce the words to their root words using Porter's stemmer6. Stemming enhances the likelihood of matching to the lexicon. Tweets can at most contain 280 characters, so users tend to write in short forms. Different users can use different terms for the same word, not every synonym word can be included in the lexicon, which will increase the processing time, stemming is crucial for the accuracy of the prediction model**.**

TABLE I
A summary of the research detects depression using Sentiment Analysis

Ref	Target	Dataset	classes	Technique
[11]	Text classifiers are taught to identify depression in this paper. The main goal is to enhance the performance of depression detection via the analysis and comparison of two sets of techniques: ensemble and hybrid.	There are posts on Reddit from both depressed and nondepressed persons in the social media collection. The dataset contained 1841 users (1200 positives and 641 negatives)	positive, negative	hybrid methods (symbolic and subsymbolic) Ensemble Methods
[12]	By examining Twitter users' thoughts and demographic data collected over two months following their completion of the Patient Health Questionnaire-9, which served as an outcome measure, this study was able to identify depression.	Three sources provide the data used in this study's model construction: the PHQ-9, a personal information questionnaire, and the Twitter API.	positive, negative	This study examines five different machine learning techniques: Support Vector Machine, Decision Tree, Naïve Bayes, Random Forest, and Deep Learning.
[13]	Our research, which is described in this paper, aims to improve upon current lexicon-based methods by creating a more precise classification function for identifying depressive text on social media platforms. Additionally, we are creating an online depression diagnostic system that incorporates our improved lexicon method, enabling users to view their current state of depression in real-time	The first step involves collecting depressing and non-depressing tweets from the Twitter platform. The sample size was 2890 tweets consisting of 1445 depressing tweets and 1445 non-depressed tweets.	positive, negative	lexicon-based methods
[6]	A model was presented to predict depression in real-time tweets by using <word, timing, frequency, sentiment, contrast	The tweets for 100 students in MS India	positive, negative, and neutral polarity	Supervised Learning (Multinomial Naïve Bayes)
[14]	design a model to predict if Arabic users tweet their Anxiety feelings on Twitter and if these feelings are depression or not	The dataset is collected from tweets from Gulf region users	positive, negative	supervised learning techniques (Naïve Bayes)
[15]	predict depression in Arabic text.	The dataset was collected from posted stories from Nafsany's forum.	positive, negative	Machine learning techniques(SVM)
[16]	Create a model to evaluate tweets from Arabic users and identify sadness in Arabic Twitter users.	The Arabic Twitter users who responded to the CES-D poll provided the dataset.	positive, negative, neutral	Support Vector Machine (SVM), K-nearest Neighbors (KNN), AdaBoost, Random Forest (RF), Logistic Regression (LR), and Naïve Bayes (NB).
[17]	This study examines a wide range of earlier research that employed learning strategies to detect depression	The present study uses two different datasets from reputable sources to identify depression in text. Reddit postings are included in Dataset 1, and Twitter tweets are included in Dataset 2.	positive, negative	Neural Network for Long Short-Term Memory (FCL) and Fasttext Convolution a convolution neural network (CNN) information, and Long Short-Term Memory (LSTM)
[18]	In this work, they utilize the clinical definition of depression to construct a typology of social media texts that may be used to assess depression severity	They present a new dataset of 40191 tweets labeled by expert annotators. Each tweet is labeled as 'non-depressed' or 'depressed'. Furthermore, tweets labeled as "depressed" are categorized into three severity levels: mild, moderate, and severe.	(1) Non-depressed, (2) Mildly Depressed, (3) Moderately Depressed, and (4) Severely Depressed.	–
[19]	This study presents a Multimodal Hierarchical Attention (MHA) model for social media depression identification and builds a Weibo-based depression detection data set	They construct a depression detection data set based on Weibo	positive, negative	Deep learning

3) Approaches (Prediction Models)

a) Machine Learning Approach

Feature Engineering and Extraction: extracted features from tweets, such as bag-of-unigrams, negation handling, word Polarity, sentiment, contrast, Term Frequency–Inverse Document Frequency (TF-IDF), and N-gram.

Supervised Learning: the model is now trained using these features. Machine learning classifiers are used namely, Multinomial Naïve Bayes, Gradient Boosting, and Random Forest. An ensemble vote classifier with a majority voting mechanism, Support Vector Machine, Decision Tree, Adaboost (Ensemble) K-Nearest Neighbors (KNN) Logistic Regression (LR), And Deep Learning techniques Each classifier takes labeled(depressed- and non-depressed) tweets and develops the model to produce predictions for fresh tweets.

b) Lexicon-Based Approach: in this approach, a depression lexicon is generated, for the prediction of depression cases. A Lexicon is a collection of depression-related terms that are more likely to be found in online posts that are written by individuals who are struggling with depression The dictionary is used with each tweet or Post to determine the percentage of words indicating depression from the total sentence, and then determine whether the user is depressed or not.

People usually use two approaches to predict depression. I think the suitable approach is to gather between a LEXICON-BASED APPROACH and a MACHINE LEARNING APPROACH to detect depression.

IV. CONCLUSION

A survey is conducted in this paper for the work done in the area of depression detection using sentiment analysis. The survey shows that few papers have focused on Arabic languages recording depression. Only a few papers have focused on the frequency and timing of the tweets. The frequency and timing of the tweets are very important to detect depression accurately. Most of the work has focused on the English Language, but the Arabic language needs more focus. In the future, research is suggested to make a model that focuses on the timing and frequency of tweets. The model can be used to predict signs of depression in future users by the analysis of their comments and how frequently they post them.

REFERENCES

[1] Al-Moslmi, Tareq, et al. "Approaches to cross-domain sentiment analysis: A systematic literature review." Ieee access 5 (2017): 16173-16192.

[2] Gill, Alastair J., et al. "Identifying emotional characteristics from short blog texts." 30th Annual Conference of the Cognitive Science Society. Washington, DC: Cognitive Science Society, 2008.

[3] Nierenberg, A. A., and J. D. Amsterdam. "Treatment-resistant depression: definition and treatment approach." The Journal of Clinical Psychiatry 51 (1990): 39.

[4] Marcus, Marina, et al. "Depression: A global public health concern." (2012).

[5] Zolezzi, Monica, Maha Alamri, Shahd Shaar, and Daniel Rainkie. (2018) "Stigma Associated with Mental Illness and Its Treatment in Arab Culture: A Systematic Review." International Journal of Social Psychiatry 64 (6): 597–609.

[6] Kumar, Akshi, Aditi Sharma, and Anshika Arora. "Anxious Depression Prediction in Real-time Social Data." International Conference on Advances in Engineering Science Management & Technology (ICAESMT)-2019, Uttaranchal University, Dehradun, India. 2019.

[7] Liu, Bing. "Sentiment analysis and opinion mining." Synthesis lectures on human language technologies 5.1 (2012): 1-167.

[8] Madhoushi, Zohreh, Abdul Razak Hamdan, and Suhaila Zainudin. "Sentiment analysis techniques in recent works." 2015 Science and Information Conference (SAI). IEEE, 2015.

[9] Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. "Sentiment analysis algorithms and applications: A survey." Ain Shams engineering journal 5.4 (2014): 1093-1113.

[10] Yadollahi, Ali, Ameneh Gholipour Shahraki, and Osmar R. Zaiane. "Current state of text sentiment analysis from opinion to emotion mining." ACM Computing Surveys (CSUR) 50.2 (2017): 1-33.

[11] Ensemble Hybrid Learning Methods for Automated Depression Detection. (2023, February 1). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/9733425

[12] Angskun, J., Tipprasert, S., & Angskun, T. (2022, May 20). Big data analytics on social networks for real-time depression detection. Journal of Big Data. https://doi.org/10.1186/s40537-022-00622-2

[13] Yeow, B. Z. (2022, January 25). A Depression Diagnostic System using Lexicon-based Text Sentiment Analysis. https://journals.iium.edu.my/kict/index.php/IJPCC/article/view/250

[14] Almouzni, Salma, and Asem Alageel. "Detecting Arabic Depressed Users from Twitter Data." Procedia Computer Science 163 (2019): 257-265.

[15] Alghamdi, Norah Saleh, et al. "Predicting Depression Symptoms in an Arabic Psychological Forum." IEEE Access 8 (2020): 57317-57334.

[16] Musleh, D., Alkhales, T. A., Almakki, R. A., Alnajim, S. E., Almarshad, S. K., Alhasaniah, R. S., Aljameel, S. S., & Almuqhim, A. A. (2022, January 1). Twitter Arabic Sentiment Analysis to Detect Depression Using Machine Learning. Computers, Materials & Continua. https://doi.org/10.32604/cmc.2022.022508

[17] Tejaswini, V., Babu, K. S., & Sahoo, B. (2022, November 5). Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and a Hybrid Deep Learning Model. ACM Transactions on Asian and Low-Resource Language Information Processing. https://doi.org/10.1145/3569580

[18] Kabir, M., Ahmed, T., Hasan, M. B., Laskar, M. T. R., Joarder, T. K., Mahmud, H., & Hasan, K. M. A. (2023, February 1). DEPTWEET: A typology for social media texts to detect depression severities. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2022.107503

[19] Li, Z., An, Z., Cheng, W., Zhou, J., Fang, Z., & Hu, B. (2023, January 18). MHA: a multimodal hierarchical attention model for depression detection in social media. Health Information Science and Systems. https://doi.org/10.1007/s13755-022-00197-5