Computer Science Faculty Research & Creative Works

Computer Science

17 Jul 2023

Detecting Mental Distresses Using Social Behavior Analysis In The Context Of COVID-19: A Survey

Sahraoui Dhelim

Liming Chen

Sajal K. Das
Missouri University of Science and Technology, sdas@mst.edu

Huansheng Ning

et. al. For a complete list of authors, see *https://scholarsmine.mst.edu/comsci_facwork/1353*

Follow this and additional works at: https://scholarsmine.mst.edu/comsci_facwork

Part of the Computer Sciences Commons, and the Psychology Commons

Recommended Citation

Detecting Mental Distresses Using Social Behavior Analysis in the Context of COVID-19: A Survey

SAHRAOUI DHELIM, School of Computer Science, University College Dublin, Ireland
LIMING CHEN, School of Computing, Ulster University, UK
SAJAL K. DAS, Department of Computer Science, Missouri University of Science and Technology, USA
HUANSHENG NING, School of Computer and Communication Engineering, University of Science and Technology, Beijing, China
CHRIS NUGENT, School of Computing, Ulster University, UK
GERARD LEAVEY, Bamford Centre for Mental Health and Wellbeing, Ulster University, UK
DIRK PESCH, School of Computer Science and IT, University College Cork, Ireland
ELEANOR BANTRY-WHITE, School of Applied Social Studies, University College Cork, Ireland
DEVIN BURNS, Department of Psychological Science, Missouri University of Science and Technology, USA

Online social media provides a channel for monitoring people's social behaviors from which to infer and detect their mental distresses. During the COVID-19 pandemic, online social networks were increasingly used to express opinions, views, and moods due to the restrictions on physical activities and in-person meetings, leading to a significant amount of diverse user-generated social media content. This offers a unique opportunity to examine how COVID-19 changed global behaviors regarding its ramifications on mental well-being. In this article, we surveyed the literature on social media analysis for the detection of mental distress, with a special emphasis on the studies published since the COVID-19 outbreak. We analyze relevant research and its characteristics and propose new approaches to organizing the large amount of studies arising from this emerging research area, thus drawing new views, insights, and knowledge for interested communities. Specifically, we first classify the studies in terms of feature extraction types, language usage patterns, aesthetic preferences, and online behaviors. We then explored various methods (including machine learning and deep learning techniques) for detecting mental health problems. Building upon the in-depth review, we present our findings and discuss future research directions and niche areas in detecting mental health problems using social media data. We also elaborate on the challenges of this fast-growing research area, such as technical issues in deploying such systems at scale as well as privacy and ethical concerns.

---

1 INTRODUCTION

Mental illness is a major contributor to the global burden of diseases [49]. The economic cost of mental disorders treatment was estimated to be U.S. \$2.5 trillion in 2010, and is expected to double by 2030 [18]. One of the main goals of the **World Health Organization's (WHO)** Comprehensive Mental Health Action Plan 2013–20 was to develop strong information systems for mental well-being, e.g., increasing capacity for population health diagnosis [128]. The direct and indirect consequences of COVID-19 on the mental health of the general population were anticipated to be considerable by many commentators. Anxiety, depression, and alcohol abuse appeared to have increased in the general population. Grief, loneliness, and isolation play significant roles in all of these conditions [106]. These impacts may be amplified for already vulnerable groups who rely on other people and organizations for care and support. Moreover, front-line health workers are thought to be vulnerable to burn-out and trauma. There are likely to be longer-term psychological effects of the pandemic created by economic disruption, unemployment, and family breakdown [50]. Risk factors such as ill health, bereavement, domestic abuse, and violence, and maladaptive coping such as increased alcohol consumption, substance misuse, and gambling are likely to contribute to increased disorders [30]. The anticipated increase in problems will occur at a time when access to services is severely restricted. The challenges to health and social care service delivery cannot be underestimated, and innovative ways are needed to maintain a connection to vulnerable people.

Online Social Networks (OSNs) are well established as a data source for public opinion mining [140], business analytics [45], events detection [39], impact on quality of life [113], and population health monitoring [28]. Hence, they are increasingly used for mental health applications at both population-level and individual-level health. Social media analysis has been studied extensively [112] and is especially promising for mental healthcare, as OSNs such as Twitter and Facebook provide access to naturalistic, first-person accounts of user behavior, emotions, thoughts, and feelings that may be indicative of mental well-being. The popularity of social media where people willingly and publicly express their ideas, thoughts, moods, emotions, and feelings, and often share their daily struggles with mental distress, offers a rich source of information for studying mental illnesses, such as depression and loneliness. In this article, we survey the existing research literature on social media analysis for detection of mental distresses, with a special focus on the context of COVID-19 pandemic.

During COVID-19 and in its aftermath, the number of research articles that focused on the social media analysis for detection of mental disorders has increased sharply, including survey and review articles. As shown in Table 1, the authors of Reference [135] reviewed the literature on social media analysis for depression and suicide detection, but they limited the review only to

Table 1. Recent Reviews and Surveys on Social Media Analysis for Detecting Mental Disorders

Reference	Scope	Limitations
Skaik & Inkpen [135]	−Covers only text-based social media. −Focuses on detecting depression and suicide ideation.	Text-only
Kim, et al. [69]	−Covers 15 works prior to August 2019. −Reviews only depression detection.	Not COVID-related
Liu, et al. [80]	−Covers causes for suicidal ideation. −Analyzes from text-based social media data.	Text-only Not COVID-related
Ji, et al. [65]	−Reviews suicidal ideation detection methods using clinical data. −Covers few works of text-based social networks data.	Not COVID-related
Chancellor & Choudhury [24]	−Reviews works published between 2013 and 2018	Not COVID-related
Castillo-Sanchez, et al. [22]	−Reviews 16 suicide detection techniques using social media data	Text-only Not COVID-related

text-based social media platforms such as Twitter, Reddit, and Weibo. In Reference [69], 15 papers were summarized to discuss detection of depression from social media data prior to August 20, 2019. In Reference [80], the authors analyzed the potential causes of suicidal ideation using text-based social media data. Similarly, the authors of Reference [65] reviewed suicidal ideation detection methods using clinical data. However, they covered only a few works of suicidal detection from social media data. Since the advent of COVID-19 pandemic, numerous studies have appeared to address concerns for mental well-being, many of which examined social media analysis on the users' mental well-being related to the COVID-19 pandemic. These works have not been fully covered in the aforementioned reviews. In addition, most of the recent reviews [65, 80, 135] focus only on text-based social media.

Focusing on COVID-19 related works, the survey in this article makes the following\break contributions.

(1) In terms of the scope of mental well-being, our survey covers more mental health conditions, such as loneliness, anxiety, stress, **post-traumatic stress disorder (PTSD)**, and other mental disorders, in addition to depression and suicide ideation.
(2) In terms of the studies, we cover both text- and visual-based social media analysis, particularly the most recent research based on the visual-oriented social media analysis, such as Instagram. To the best of our knowledge, it is the first survey to cover the most recent research findings regarding the COVID-19 impact on mental well-being.
(3) We provide new approaches to organizing the literature in terms of feature classifications and detection methods, thus offering new insights and visions to the related research\break communities.
(4) We draw observations, derive findings, and point out research gaps, directions, and challenges that are most valuable to guide researchers and practitioners to future research.

The rest of the article is organized as follows. Section 2 outlines the scope of the survey and the methodologies considered. Section 3 briefly summarizes the widely used OSNs and their basic functionalities. In Section 4, we describe various features, their discrimination capabilities, and feature extraction methods of the social media content. Section 5 reviews social media-based mental distress detection techniques. Section 6 discusses recent literature on social media mining for detecting mental distress and classifies them according to their feature extraction and detection techniques. Section 7 outlines our findings and possible future directions. Section 8 presents
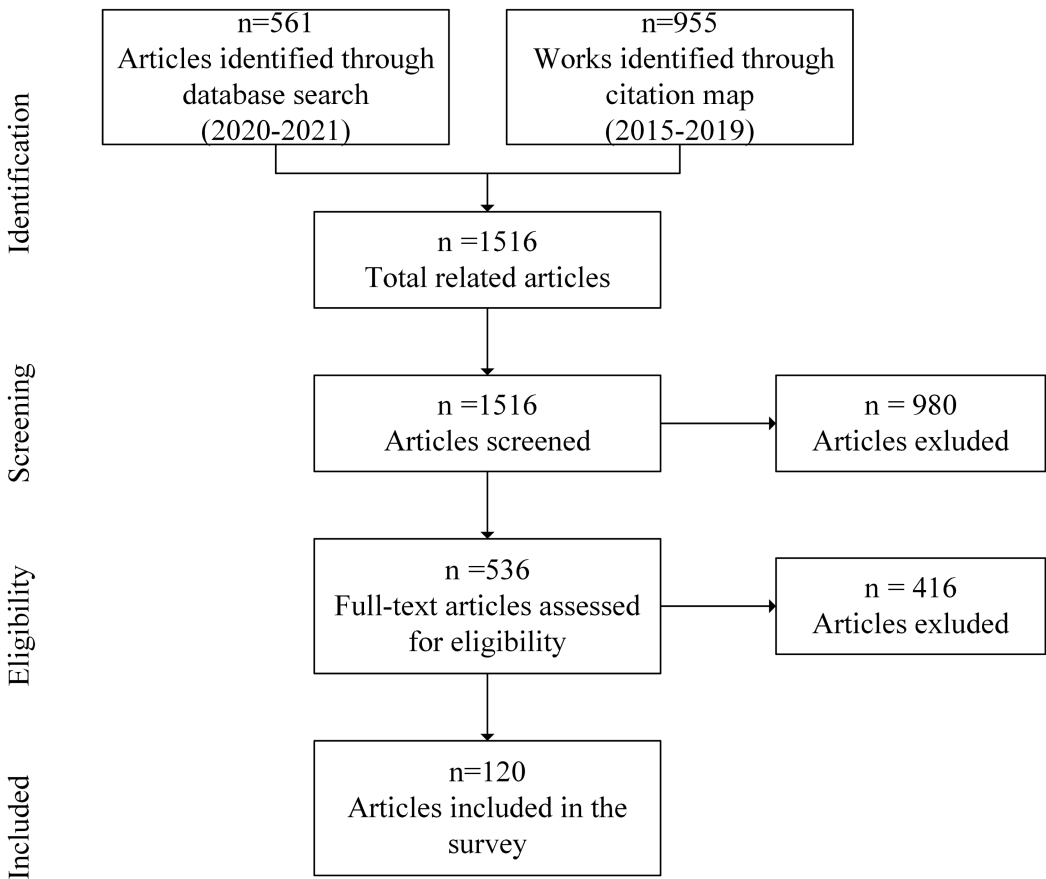
Fig. 1. PRISMA flowchart of the review phases.

some open issues and challenges of mental distress detection from social media. Finally, Section 9 concludes the survey.

2 SCOPE OF SURVEY AND METHODOLOGY

The focus of this survey article is on mental disorders, also known as psychological disorders. (These two terms will be used interchangeably throughout the rest of the article.) Mental disorders incorporate a wide range of mental illnesses, such as depression, anxiety, PTSD, and schizophrenia, to name a few. Specifically, we survey the works that use social media data for detection of mental disorder. We focus on works published during 2020–2022, with a special focus on the works conducted in the context of COVID-19 pandemic. We also cover recent key works, e.g., highly influential works in the field of mental disorder detection using social media analysis. We used the guidelines of **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)** framework [94] to select publications related to loneliness, suicide, and mental disorders detection using user-generated social media data. As shown in Figure 1, initially 561 related papers between January 2020 and April 2021 were identified after searching Google Scholar, Elsevier, IEEE Xplore digital library, ACM Digital Library, Springer, and PubMed for articles related to the following search terms: "mental disorder," "psychological disorder," "mental distress," "psychological distress," "social media," "social networks," "depression," "COVID-19," "stress," "anxiety," "loneliness," "PTSD," "suicide," and "schizophrenia."

Table 2. Most Used OSN for Mental Disorder Detection

OSN	Active Users in Millions (2023)	Main Functionality
Twitter	450	Microblogging
Reddit	430	Forum and blogging
Sina Weibo	586	Microblogging
Facebook	2,963	Social networking
Instagram	2,350	Photo and video-sharing

The searches were limited to articles written in English. An additional 955 articles were identified as related works in the field published between January 2014 and December 2022, by following the citation map of the searched articles. After removing duplicated papers, a total of 1,516 articles were gathered in the identification phase. In the screening phase, based on the title and abstract screening, 980 articles were excluded for not meeting the inclusion criteria. Most of these articles are either not based on social media data analysis (instead based on wearables and smartphone-based passive sensing), or they focus on non-mental healthcare, such as diabetes, obesity, or outbreak predictions. Thus, 416 articles were excluded in the eligibility phase after full-text reading. Finally, 120 articles were qualified for final inclusion in this survey article, 59 articles were selected to establish the research context and background.

3 ONLINE SOCIAL NETWORKS (OSN)

The popularity of OSN has increased sharply during the last few years. Following the travel restrictions and lockdown measures imposed during the COVID-19 pandemic, people have increasingly relied on social media. The OSNs differ in many aspects, some are microblogging oriented like Twitter and Sina Weibo, while others are multimedia like Instagram and Facebook. The data extracted from these OSN differ in their types and features. Table 2 presents the most used OSNs for mental disorder detection in the literature, which we introduce in the following together with their data extraction methods.

3.1 Twitter

Despite its relatively low number of active users (450 million active users by 2023) compared to other social networks, Twitter is by far the most popular OSN in the research community. Datasets extracted from Twitter are widely used in the literature of social media analysis. Most works surveyed in the current study have used Twitter as the data source. The tweets are publicly accessible and hence can be extracted and analyzed. In addition to the text body, the tweets also include user metadata, such as the user's geographical and network information, and the tweet's date and time. Thus, Twitter data can be used not only for user-level studies but also for population-level studies. Tweets can be obtained using Twitter API by searching keywords, hashtags, or other queries; the search can be limited to specific locations or time intervals. In the context of mental distress detection, there are roughly three ways to find and extract the data of mentally distressed users. Some works recruit the participants for experimental study, directly extract their Twitter data [31], analyze the extracted data, and compare them with their self-assessed scores of various mental healthcare questionnaires. While other works search for keywords or randomly look for mentally distressed users by searching for specific phrases and apply regular expression [55], for example, searching for the phrase "I was diagnosed with depression" or " I feel so down." The downside of this method is that it may require human intervention to confirm the user's mental distress [170]. A third way is to search for a specific hashtag to identify tweets related to a particular mental illness, such as #depression or #MyDepressionLooksLike [74].

3.2   Reddit

Reddit is a forum-like social network, which consists of thousands of active, user-created communities (also known as subreddits) that focus on areas of interest such as art, football, movies, politics, mental disorders, and many other topics. It had over 430 million active monthly users by 2023. Reddit users can participate in a community by sending text- or media-based posts, and by commenting on other posts or comments. The users can rate posts and comments using up-vote or down-vote functions. The score of a post or comment can be calculated as the difference between up- and down-votes. Many subreddits address mental disorders, including depression, suicide and anxiety. In these subreddits, the users share their daily life, express their opinions and feelings, and some may offer mental support to others, or COVID-19 related topics [138]. By March 2023, /r/depression subreddit had 945,000 members, r/SuicideWatch had 409,000 members and r/Anxiety had 600,000 members. Mental health-related subreddits offer a very rich source of data that can be used to train mental distress detection models, since these subreddits contain a high percentage of ground-truth data concentrated in one place, unlike Twitter data that require extensive searching to identify such mentally distressed users. Posts in the r/SuicideWatch subreddit on Reddit show the details and nature of comments and supports sought by users.[1]

3.3   Sina Weibo

Sina Weibo is the largest microblogging website in China (simply known as Weibo), whose main functionality is very similar to Twitter. Many studies in the literature of mental distress detection were conducted using Weibo datasets. With more than 586 million active users that post about their daily life struggles, Weibo offers a rich source of user mental health information. Various works have collected datasets from Weibo; these datasets are used for different mental illness detection tasks. In Reference [160] the published data of 1 million users were extracted with more than 390 million posts. A keyword-based method was employed to mark the users at suicide risk. Following that, three mental health experts manually labeled the users at risk of suicide. They identified 114 users with suicide ideation, and leveraged linguistic analysis to explore behavioral and demographic characteristics.

3.4   Facebook

Facebook is by far the largest social network in terms of the number of active users, refer to Table 2. Facebook offers the possibility to create and share text posts, as well as photos and videos. Users can also join various groups and follow Facebook pages. Unlike microblogging OSNs such as Twitter and Weibo, the Facebook textual posts are not limited by character limits (e.g., tweets have a 280 character limit). In the context of mental distress detection, user-generated data can be used for various purposes, specifically, to extract textual, visual features, and also behavioral features. See Section 4 for more details on the different types of extracted features.

3.5   Instagram

Instagram is the largest visual-based social network, where the users share photos and videos. Mental distress detection using media-based OSN may not be as straightforward as text-based OSN. Generally, photos and video require deep analysis to associate their features with mental distress markers. Instagram photos incorporate a variety of features that can be analyzed for mental state assessment. The content of photos can be represented by various characteristics: Are there people in the photos? Is the photo setting outdoors or indoors? Was the photo taken at night or

---

[1]/www.reddit.com/r/SuicideWatch/.
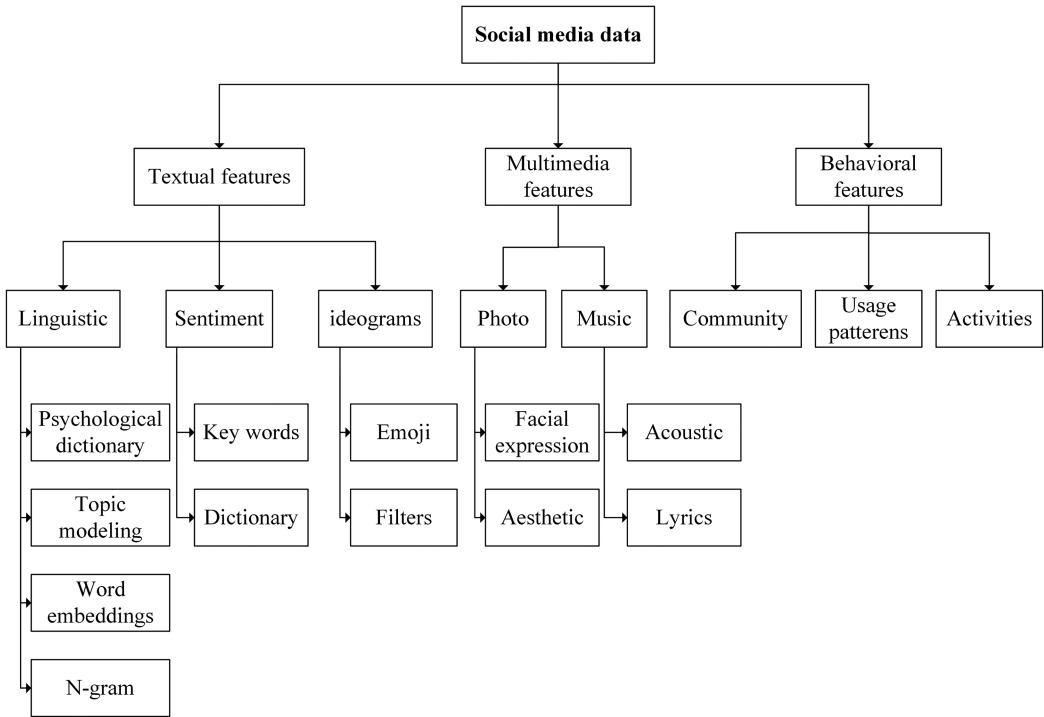
Fig. 2. Psychological features classification.

day? A photo's statistical properties can also be analyzed at the pixel level, such as average color and brightness. Instagram post metadata contains additional information about the photo: Did the post receive any comments? How many "Likes" did the photo receive? Finally, behavioral features (e.g., the usage and posting frequency) may also give some hints about the user's mental state.

4 MENTAL FEATURES EXTRACTION

Social network data contain various features associated with mental distress and can reveal the psychological state of the users. These features are either extracted from single modalities (e.g., text, image, or audio) or fused from a variety of multi-modal data sources [100]. Multi-modal feature processing generally achieves better detection accuracy of mental distress compared to the single modal schemes, as individuals tend to manifest their inner psychological states using different expression media [12]. Hence, multi-modal features can alleviate the generalization error caused by individual differences, such as personality type, ethnicity, age, and gender. Different features are extracted from the social network data depending on the type of shared content. We propose a taxonomy in Figure 2 to classify existing works on mental distress detection using social network data.

4.1 Textual Features

The user-generated text data incorporate various latent features that can be leveraged to reveal the user's psychological distress. As shown in Figure 2, textual features can be divided into three classes, namely, linguistic, sentiment, and ideogram features. The fusion of multi-modal features can be applied to the same textual features class by combing various features from the same

class (e.g., linguistic-linguistic [145]), or by combining features from various feature classes (e.g., linguistic-sentiment [119]).

4.1.1 Linguistic Features. The linguistic features are extracted from the language use patterns such as word choice, topics of interest, and sentence structure. One of the most used linguistic feature techniques is to map the user-expressed text corpus into a psychological dictionary and measure the word frequencies in each word class. One of the most widely used psychological dictionaries is **Linguistic Inquiry and Word Count (LIWC)**, which is a widely used text analysis technique in psychology, and can be easily adapted for **natural language processing (NLP)** tasks. LIWC was introduced in the early 1990s to associate linguistic dimensions of written expression with psychological states. LIWC could compute the percentage of words within 80 linguistically or psychologically meaningful categories. These categories cover various important psychological states of an individual, including personal preferences, cognition state and emotions. Over the past 30 years, LIWC has been used in several works studying the relationships between the word categories in daily language and psychological states. For instance, the correlation between the first person singular pronoun usage and depression [59], emotions and LIWC emotion-related categories [144], and LIWC positive emotion category word usage and anxiety [137]. In the context of mental distress detection using social media analysis, LIWC is by far the most used psychological dictionary-based feature extraction method [60, 78, 81, 156]. Other works used psychological dictionaries to extract linguistic features, such as Affective Norms for English Words. Another effective linguistic feature extraction method is topic modeling, the process of extracting users' topics of interest by analyzing the text corpus generated from their posts. Topic modeling is an effective technique in computational linguistics to decrease the input of textual data feature space to a determined number of topics. Using unsupervised text mining techniques can be used to extract hidden topics from a text corpus, such as topics related to psychological distress.

Unlike psychological dictionaries such as LIWC, it is not created by a fixed set of pre-categorized words. However, topic modeling techniques automatically compute and generate the set of non-labelled words that represent the user's topics of interest. In the context of topic modeling for psychological distress detection, the generated topics are considered the linguistic features of the user. The most used topic modeling method is **Latent Dirichlet Allocation (LDA)**. In LDA, a topic is represented as a multinomial distribution associated with unique words in the text corpus. Furthermore, a document is represented as a multinomial distribution over all topics. LDA is used to generate topics and provides linguistic features automatically from the user's text corpus. Many studies leverage the original unsupervised LDA algorithm for topic extraction [48, 92, 122, 153, 171]. Some works report better results using supervised or semi-supervised LDA [34, 37, 38] that guide the topic modeling process by specifying mental-related lexicons, which are likely to appear in the generated content of distressed individuals [8, 141].

Another widely used linguistic features for mental distress detection are the word frequency features that represent the original text corpus in a representational language model. The most basic models in this category are the **Bag of Words (BoW)** model and **term frequency-inverse document frequency (TF-IDF)**. In BoW, the user-generated text is represented as the bag (multiset) of its words, disregarding grammar and even word order but keeping word multiplicity. TF-IDF is a term-weighting scheme that measures the importance of a word in all text generated by the user. For example, TF-IDF might be applied to know the importance of the word "depressed" in all the tweets of a given user. One disadvantage of BoW and TF-IDF is that the order of the words is discarded, and hence important information is lost about the temporal dimension of the user's psychological state. The N-gram model [149] solves this problem, as N-gram is an adjacent sequence of entities. The entities can be syllables, phonemes, characters, words or base pairs. N-gram

models are extensively used in computational linguistics and statistical NLP for various tasks. N-gram models are used as linguistic features that represent the probability of co-occurrence of each input sentence as an unigram and bigram.

Word embedding is a deep learning technique that is used for text representation for the recognition of comparatively important words. This modelling approach is based on the mapping of every word into a corresponding low-dimensional vector, where each word is represented as a positive or negative decimal number. One of the widely used word embedding techniques is **Global Vector for word representation (GloVe)** [124], which searches for similar words in the whole context of the user-generated text. Another widely used word embedding is word2vec [33]. Generally, linguistic features fusing from more than one class can increase the mental distress detection accuracy compared to single class linguistic features, e.g., LIWC+N-gram [161], LIWC+LDA+N-gram [145].

4.1.2 Sentiment Features. Sentiment analysis is concerned with extracting emotions, opinion affects from user-generated texts on social media. Knowledge from this area can also be very useful to find the psychological state of the user. Especially, the sentiments that users describe toward their personal situation could be an important marker. Besides the linguistic features described above, the user-generated text on social media may contain semantic features related to mental distress that are difficult to extract using linguistic features, such as emotions and mood. In this context, many works have leveraged sentiment analysis to extract emotional features. The most naïve method of capturing user's sentimental features is the keyword search, in which the system looks for specific pre-defined keywords in the users' posts, e.g., "depressed," "feeling lonely," or "happy" [23, 168]. The drawback of keyword approach is the lack of context of using these words, as users may use these keywords in a sarcastic manner, for instance. A more selective method is to compute the frequency of using words that belong to LIWC emotion categories [51, 97], or LDA topics that represent basic emotions. Another way to measure the importance of emotion-related words is to measure the frequency of occurrence of these words using TF-IDF, which is often used in sentiment analysis tasks [3, 166]. In addition to this, the NRC Emotion Lexicon [11] and VADER Sentiment Lexicon [61] are used for user sentiment and emotion recognition.

4.1.3 Ideogram Features. Most social networking platforms offer users the opportunity to enrich the posted text with additional ideograms such as emojis and stickers. The usage patterns of these ideograms are rich features that can reflect the users' psychological distress. For example, continuous usage of stressed face emojis and stickers might be associated with stress. Emoji usage choices have been associated with depression [15, 89, 155], mental distress [40], and emotions [150]. One of the main advantages of the ideogram feature is the easy and straightforward extraction, unlike linguistic and sentiment features, which require various pre-processing steps.

4.2 Multimedia Features

With the popularity of visual-oriented social media platforms (e.g., Instagram), and audio streaming services such as (e.g., Spotify), many users tend to use these platforms while being passive on other text-based social networks. Previous research has proven that visual and audio features are more expressive compared to textual features regarding the psychological states of the users [109].

4.2.1 Visual Features. Visual features can be extracted from photos and videos posted on social networking websites. Due to the processing difficulties of extracting visual features from videos, most of the previous works focused on feature extraction from photos. There are two types of social media photos, namely, profile photos and posted photos. Generally speaking, profile photos convey less information about the mental state of the user, as most of the users choose socially attractive profile photos, which is known as self-presentation biases; for instance, even depressed

Fig. 3. Visual feature and its relation to depressed user.

users express positive emotions in their profile photos [52]. As shown in Figure 2, three types of features can be extracted from user-generated photos, facial expression features, object class features, and aesthetic features. Facial expression features include those of the studied users, as well as people present in the photo. Facial expression features are not only used to identify the user's sentiments but also information about their exposure to society. For example, the tendency to post group photos is a strong indicator of the social anxiety level of the user. Another indicator is the selfie posting frequency on the psychological state [172]. However, aesthetic features represent the photo's characteristics, such as the color and filter choices. For example, a photo might be analyzed at the pixel level averages to extract **Hue, Saturation, and Value (HSV)**—the three-colour properties widely used in image feature extraction. The Hue represents the photo's coloring on the light spectrum (varying from red to purple/blue). Lower hue values reflect in redder, while higher hue values result in the photo tending more to blue. Saturation represents the vividness of the photo. Low saturation results in the photo being greyer and faded. Value is the image brightness. Lower brightness scores reflect a darker photo. Previous research associated depression with the posted images being grayscale and low aesthetic cohesion across a variety of image features. Users suffering from social anxiety also tend to post grayscale and low aesthetic cohesion photos, but less so than depressed users [52] (see Figure 3).

4.2.2 Audio Features. Various features can be extracted from music listening activities using online streaming services. Thanks to the available music public dataset and online streaming services API, we can analyze and extract rich features associated with mental distress. By analyzing the user's listening history and playlists, we can extract acoustic features of these songs, such as rhythm structure, timbrel texture, pitch, and spectral features. Lyrics features are extracted by applying textual feature extraction techniques, as discussed above, on the lyrics of the listened songs. For instance, computing the LIWC categories of the lyrics can be associated with the mood of the users' who listened to this song.

4.3 Behavioral Features

In addition to the user-generated content feature, the online behaviors convey valuable hints about the users' daily life activities and their interests.

4.3.1 Community Affiliation. Social networks offer the users the ability to join social network groups and communities (e.g., Facebook groups and Reddit communities). Analyzing the type of groups and communities a user joins might help detect their topics of interest. Mental distress-related groups and communities can specifically help identify users with mental health

issues. My previous studies have shown that topics and psycholinguistic features were found to be accurate predictors of mental disorders like depression [122], bipolar [70], and suicide ideation [130].

4.3.2 Usage Patterns. The patterns of social media usage and the variation of usage across different periods is a strong indicator of the user's psychological state. Usage features include the distribution of social media usage within the day, the daily usage time and the usage time increase/decrease. It has been found in Smith et al. [136] that the sudden increase in Facebook posting activity is positively correlated with the depression symptoms. For instance, there is a marginal positive correlation between Instagram hourly usage time and depressive symptoms [84].

4.3.3 Online Activities. Besides the usage time and community affiliations, the interaction activities with posted content (e.g., likes, comments, and follows) might be correlated with certain psychological markers. For example, the authors of Reference [99] found that the depressive symptom is negatively correlated with the number of Facebook friends. Reference [84] investigated the number of strangers the user follows, and its relation with depression markers, and the number of strangers followed slightly correlated with depression. Another way to reveal the psychological problem is by analysing keyword searching behaviors [63]. Analysis of Baidu search data reveals that searches related to psychological problems increased during the COVID-19 pandemic, keywords like depressed, sad, panic, fear, insomnia, obsessive-compulsive disorder, and psychological counselling [26].

5 MENTAL DISTRESS DETECTION TECHNIQUES

Due to the large size and diversity of user-generated data on social networks, many detection models have been used to learn from such rich data. Mental distress detection algorithms are divided into three main classes: machine learning, deep learning, and statistical analysis methods. In this section, we survey various learning models used in the literature on social media mental distress detection. Figure 4 presents the classification of the detection methods.

5.1 Machine Learning-based Detection

Machine learning detection schemes are the most used techniques for mental distress detection.

5.1.1 Classification and Clustering. In machine learning, classification is the problem of finding to which group/class/category a new observation belongs, based on existing knowledge of a training data set that includes instances whose class affiliation is already known. For example, given three sets of user groups (low, medium, high) based on their depression level, a classification algorithm is used to analyze the user-generated data on social media and assign the user to one of the three groups (classes). One of the most effective machine learning classifiers is the **support-vector machine (SVM)** that has been proven to perform well with short informal text and produces promising results when applied to mental health classification tasks [31]. SVM is widely used in mental disorders detection using social networks data, including depression detection [31, 108, 145], social anxiety [25, 41, 165], and suicide risk assessment [27, 85, 146]. Other widely used classifiers for mental health assessments include **K-Nearest Neighbor (KNN)** [62] and **Naïve Bayes (NB)** classifier [19].

5.1.2 Regression Methods. Regression is a commonly used statistical procedure to iteratively refine a measure of the error in the prediction made by the model to find relationship between the variables. It can be used to predict the risk of mental distress based on the extracted social media features of the user. Specifically, the logistic regression model can calculate the probability of a categorical variable (depressed or non-depressed) from a set of predicted features. Linear regression
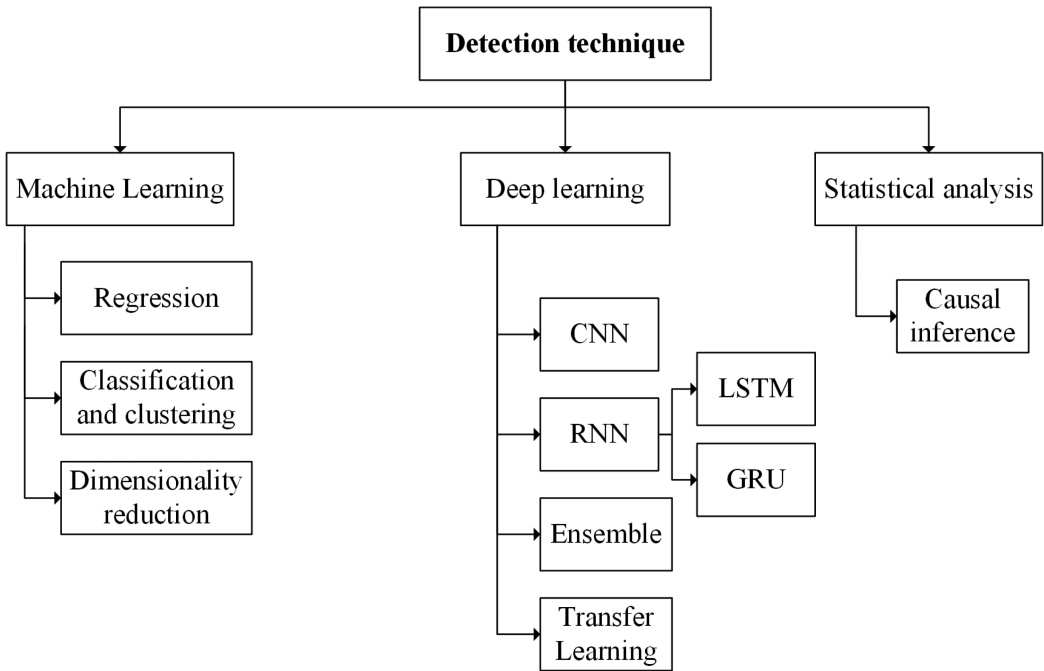
Fig. 4. Classification of mental distress detection methods.

is also used in various mental disorder detection methods. In linear regression, the relationships between the mental disorder and user data are modelled using linear predictor functions whose unknown model parameters are estimated from the user-generated social media data. For example, in Reference [119], a linear regression depression predictive model was constructed, linguistic features from the Instagram comments and post captions including emoji sentiment analysis results, and multiple sentiment scores and meta-features (e.g., likes count and average comment length) were obtained.

5.1.3 Dimensionality Reduction. Dealing with raw data in high-dimensional spaces can be undesirable for various reasons. Raw data are often sparse because of the problem of dimensionality, and analyzing such raw data is usually computationally expensive. Dimensionality reduction is the transformation of raw data from a high-dimensional space into a low-dimensional space such that the low-dimensional representation preserves some meaningful properties of the raw data, ideally similar to its intrinsic dimension. The objective of dimensionality reduction methods is to learn the inherent latent structure in the data, however, in this case in an unsupervised way to summarize or describe the data using less information. This can be useful to visualize dimensional data or to simplify data, which can then be used for supervised learning methods. Many dimensionality reduction methods can be leveraged in classification and regression tasks. **Principal Component Analysis (PCA)** is the most used dimensionality reduction technique in the context of mental disorder assessment using social media data, such as the representation of social data of depressed users [139], schizophrenia users [91], and suicidal users [76].

5.2 Deep Learning-based Detection

In the last few years, deep learning methods have become the mainstream detection technique for mental distress from social networking data. Focused on building larger and more complex

neural networks, deep learning detection schemes are trained with very large datasets of labeled ground-truth data, such as linguistic features, social media images, music streaming audio, and video.

5.2.1 Convolutional Neural Network (CNN). It is a type of deep neural network, mostly applied to analyzing images. CNNs are a special case of **multilayer perceptrons (MLP)**, which are typically fully connected networks, implying every neuron in one layer is connected to all neurons in the next layer, and thus over-fitting the data. This problem is usually solved by regularizing these networks by adding magnitude measurement of weights to the loss function. CNNs use a different method for regularization: they leverage the hierarchical pattern in data and form more complex patterns using smaller and simpler data patterns, making them less connected and complex yet effective. CNNs achieve good results, especially when applied to image classification, and also yield good accuracy for data with a grid-like structure. Many studies have proven that CNNs can be utilized effectively for text classification tasks. Given the user-generated social media data, CNNs can be employed using only a single layer with many filters for sentence classification [152], or using a more complex multilayer modeled as a sequence of layers such as an embedding layer, convolutional layer, dense layers, max-pooling layer, and the output [70].

5.2.2 Recurrent Neural Networks (RNN). It is a class of artificial neural networks in which the connections between the nodes is in the form of a directed graph along a temporal sequence that allows to model temporal dynamic behavior. Inspired by feedforward neural networks, RNNs utilize their internal state to process variable-length sequences of inputs, which makes them suitable for tasks such as unsegmented and connected series of data. The most famous variants of RNNs is the **Long Short-term Memory (LSTM)** and **Gated Recurrent Units (GRU)**. LSTM is augmented by recurrent gates, also known as forget gates, which stop back-propagated errors from exploding or vanishing. Instead, errors can flow backwards over various virtual layers unfolded in space. This enables LSTM to learn tasks that require memories of events that took place thousands or even millions of discrete time steps earlier. LSTM and its variants such as **Bidirectional LSTM (Bi-LSTM)** are the dominant deep-learning models in the literature on social media data analysis for mental distress assessment. Some works opted to use only LSTM or its variants for mental disorders assessment [2, 58, 64, 134], while others combine more than one deep learning model, such as LSTM+CNN [146, 164] or GRU+CNN [120].

5.2.3 Transfer Learning. Transfer learning is a technique used to store the knowledge obtained from a machine learning or deep learning model that was developed for a task and reused to tackle a different but related task. One of the widely applied transfer learning techniques is the usage of pre-trained deep-learning models as the starting point on NLP and computer vision tasks, which significantly reduces the amount of data required to train these model on the target task. This is because the neural networks have already been pre-trained on a large dataset [20, 157]. In the context of mental health analysis for mental disorder detection, pre-trained classifiers that were trained on a general-purpose NLP task can be leveraged to identify the language indicative of mental disorders [5]. For instance, in Reference [123], a transfer learning classifier is used to identify mental health-related language on Twitter, such as anxiety, depression, stress, and suicidal ideation.

5.3 Statistical Analysis

Besides mental distress detection, some other works have focused on the causal relationship between social media usage activities, and physical and mental well-being [54]. The correlation analysis allows us to capture the relationships between the extracted feature patterns and abnormal habits related to psychological well-being. In this class, statistical analysis is widely used to infer

any causal correlation or dependence and find the statistical relationship between the user's social media data and his mental well-being. For example, correlation analyses were performed in Reference [107] to identify depressive symptom-related features from Facebook users and **Center for Epidemiological Studies-Depression (CES-D)** scale scores.

6 SOCIAL MEDIA MINING FOR MENTAL HEALTHCARE

In this section, we will review the recently published works on the topic, with a special focus on works that have been conducted in the context of the COVID-19 pandemic.

6.1 Depression Detection

Depression is a common yet serious mental disorder. According to the WHO, more than 264 million people suffer from depression around the world [163]. Symptoms of depression include feeling low self-esteem, hopeless or unhappy, and finding no pleasure in activities one used to enjoy. Depression is different from daily mood swings and short-lasting emotional reactions to everyday life events. Various factors can cause depression such as continuous stressful life events, work pressure, personality disorders, family history and giving birth. Depression can cause a serious health condition, especially with long-lasting moderate or severe intensity. In such cases, depression might influence the subject to suffer from acute low mood and being less productive at work or school, or even leading to suicide [43]. Moreover, lockdown measures caused by the spread of COVID-19 have rubbed salt into such a wound. Following the economic crisis and job loss, add to that the fear of the virus that leads people not to seek mental assessment even when they genuinely needed it, not to mention the feelings of worry, fear, and stress, at the individual as well as at community level. The number of depressed people has increased sharply since the emergence of COVID-19 [1]. Although there exist various effective treatments for depression and other mental disorders, the majority of people in middle-income and low-income countries have limited access to mental healthcare [125]. Challenges to effective mental healthcare include a lack of resources, low training quality of healthcare workers, and social stigma related to mental disorders [131]. Another challenge for effective mental healthcare is inaccurate assessment. Depressed patients can be misdiagnosed, while others with no depression can be mistakenly prescribed antidepressant medication [10].

6.1.1 Depression Assessment. There exist many depression assessment methods, among which the clinical interviews administrated by practising clinicians or psychologists are the most reliable method of depression assessment. However, it is extremely difficult to conduct such interviews on a large population [142]. Psychometric self-report questionnaires for depression are generally considered valid and reliable assessment methods. Some of the most commonly used depression self-report surveys are the **Patient Health Questionnaire (PHQ-9)** [73], CES-D [115], and **Beck Depression Inventory (BDI)** [13]. To the best of our knowledge, only one social media-based depression detection study has used clinical interviews to measure the detection accuracy [158]. Some studies have asked the users to answer self-report depression questionnaires and used their answers as the ground truth to measure the accuracy of their schemes [31, 117, 118]. Alternatively, most of the works used less certain ground-truth data, such as the user participating in communities about depression in Reddit [48, 82, 90] and LiveJournal [101]. Another method is to mine the shared content on social media to establish ground truth, for example, searching the user content to find specific expressions, such as "I was diagnosed with depression" [16], or searching for self-report anti-depression medication usage evidence.

6.1.2 Depression Detection using Textual Markers. Textual features have been used in many works as depression markers. The textual content of social media posts may also help reveal some

signs of depression. For instance, a post that says, "I am feeling down," would be considered as expressing a depressive sign. In Reference [95], depressive symptoms from Facebook posts of randomly selected university students were analyzed. The findings concluded that there was a significant positive correlation between depressive signs expressed in the subject's Facebook posts and their score on the PHQ-9 depression scale. Similarly, the study in Reference [104] concluded that adolescents who publicly expressed daily life stress in their Facebook posts had a higher BDI score than those who did not. In the same vein, social media post of adolescent girls was analyzed in Reference [42]. It was found that posts containing somatic complaints, negative effect, and call for support are correlated with depression symptoms, while no depressive peers posts were more likely to contain an offer of help and support. The authors of Reference [133] collected 67,796 Reddit posts from 365 fathers, over a 6-month period around the birth of their child. They have used a list of "at-risk" keywords that were suggested by a perinatal mental health expert. Postpartum depression was detected by monitoring the change in fathers' use of words indicating depressive symptomatology after their childbirth. In the context of depression related to the COVID-19 pandemic, English Twitter depression dataset was created in Reference [173] containing 2,575 distinctly identified depressed users with their past tweets, and three transformer-based depression classification models were trained based on the collected dataset. Linguistic features have been proven to be good markers of depression.

Many studies have used psychological dictionaries to analyze the language usage and word choice and their association with depression. The authors of Reference [134] proposed an unsupervised depression detection algorithm based on RNN, computed a vector of LIWC features for each post text, and feed them to the LSTM autoencoder along with network-based features modeling how users connect in the forum. The results on detecting depressed users show that psycho-linguistic features derived from the users' social media posts are good predictors of the their depression severity. In Reference [97], the 14 psychological attributes were used in LIWC to classify the post into emotions, and assign weights to each word from happy to unhappy after LIWC classification. Machine learning classifiers were trained to distinguish the users into three depression levels—High, Medium, and Low. A depression detection method was developed in Reference [154] that combines LIWC with another 39 attribute sets and 252 depression-related words with temporal and linguistic styles. The authors of Reference [145] analyzed Reddit user posts to detect depression from textual features and found that combining more textual feature types can yield very high depression detection accuracy. They combined LIWC, LDA and Bigram features, and fed them to the MLP classifier resulting in a performance of 91% accuracy and 0.93 F1 scores for depression detection. A depression predictive model was proposed in Reference [119] that makes use of linguistic features, sentiment scores, number of likes, average comment length, and emoji sentiment features.

6.1.3 Depression Detection using Multimedia Markers. Depression signs are strongly correlated with the users' multimedia features such as profile photo choice, shared photos and videos, and music-listening activities. Compared to other psychological distress and mental disorders, the depression has a strong correlation with multimedia features. In Reference [117], the authors used CES-D scale as ground-truth information, and collected 43,950 Instagram photos of 166 users. They applied colour analysis, metadata components, and face detection on the collected photos and investigated the presence of a correlation between the photo properties and depression scores. Specifically, face detection was used to count the number of human faces in the photo as an indicator of the participants' social activity levels. Pixel-level averages were calculated using (HSV) values with three color properties widely used in photo analysis. The authors concluded that human ratings of photo attributes (happy, sad, etc.) are weaker predictors of depression, and no correlation was found with the computationally generated features. Similarly, various textual features such as

LIWC, BoW, FastText embedding, and ELMo embedding were combined in Reference [88], along with the color of images as visual features. Specifically, the authors extracted the HSV features by taking the average of the pixels in the image, and also counted the number of faces in each image using a deep learning face detection model, resulting in a total of 12 visual features and 64 textual features. In Reference [87] is proposed a multi-modal depression prediction model that leverages text, images, and videos shared by the user to implement a joint representation. The model uses word2vec to extract textual features, VGG-16 to extract image visual features, and faster R-CNN to extract video visual features. Finally, these features are utilized to obtain a weighted average score, which is used for making the final prediction using the Softmax prediction layer.

6.1.4 Depression Detection using Behavioural Markers. Online activities can be rich information to study the preferences and behaviors, as well as the differences among users. In Reference [136], it was observed that the sudden increases of Facebook posting activity are positively correlated with depression symptoms. The author in Reference [99] found that depressive symptoms are negatively correlated with the number of Facebook friends, and concluded that the connectedness of social networks (i.e., the number of mutual friends) is positively correlated with fewer depression signs, and depressed people tend to use the location tag function and the "like" button less often. In Reference [118], it was shown that the average word count per tweet was negatively correlated with depression; however, unlike Reference [136], no significant correlation between the depression symptoms and the frequency of tweets was observed. The relationship between the depression signs and the Instagram usage time and the number of strangers followed by the user, was studied in Reference [84]. The authors found a marginal positive correlation between the Instagram hourly usage time and depressive symptoms, and the number of strangers followed slightly moderated this relationship. In other words, the depression signs increased with the Instagram usage time if the user followed a large number of strangers, but if the user followed fewer strangers, the Instagram usage time and depression signs were unrelated. Table 3 summarizes the recent social media-based depression detection schemes.

6.2 Suicide Ideation Detection

Mental disorders are a major risk factor for suicide. According to a WHO report, approximately 800,000 people take their lives every year [162]. The U.S. National Institute of Mental Health classifies suicide into three levels: suicidal ideation, suicide attempts, and completed suicide. Suicidal ideation is the desire to commit suicide with no real attempt yet, which is a vital step for suicide risk assessment [36]. Conventional suicidal ideation detection methods mainly depend on clinical assessment or self-reported questionnaires. Many suicidal ideation scales and evaluation tools have been developed, such as Suicide Probability Scale [29], Adult Suicide Ideation Questionnaire, and Suicidal Affect-Behavior-Cognition Scale [57]. These questionnaires are effective and easy to conduct, but are prone to false negatives due to the participants' deliberate concealment, not to mention, the high cost and difficulties to conduct these questionnaires on a large scale and over a long-time. The prevalence of social media has presented a unique opportunity for a new method for detecting suicidal ideation and suicidal cause analysis. People with psychological distress usually do not trust traditional mental health methods and services. Research showed that people with psychological distress tend to look for help from informal resources, such as social networking platforms instead of seeking psychological expert help [9].

Many researchers have confirmed the effectiveness of suicidal ideation detection by analyzing users' online activities and the generated social data. The importance of emotions in Twitter content in detecting suicide risks is discussed in Reference [126], which analyzed the features of Twitter users' emotions and behavior responses (sadness, fear, anger, joy, positive, and negative)

Table 3. Recent Depression Detection Schemes

Work	Textual Feature	Visual Feature	Behavioral Feature	Detection Technique	Social Network
[145]	LIWC+LDA+bigram	N/A	N/A	MLP	Reddit
[134]	LIWC	N/A	Network (follow)	LSTM	ReachOut.com
[46]	Word embeddings	N/A	N/A	CNN-BiLSTM	Twitter
[97]	LIWC+ TF-IDF	N/A	N/A	Various classifiers	Twitter
[133]	Key-words	N/A	Engagement+ community affiliation	SVM	Reddit
[90]	TF/IDF	N/A	N/A	Logistic Regression	Reddit
[151]	BoW+TF/IDF	N/A	N/A	SVM	Reddit
[4]	Word embedding (Glove+ Word2Vect)	N/A	N/A	LSTM	Twitter
[2]	Plain text	N/A	N/A	BiLSTM	Twitter
[154]	LIWC+keywords	N/A	N/A	Statistical analysis	Facebook
[172]	N/A	Aesthetic	N/A	Logistic regression	WeChat
[88]	LIWC, BoW, FastText, ELMo	HSV	N/A	LSTM	Instagram
[117]	N/A	HSV	N/A		Instagram
[6]	Word embedding (Word2Vec) N-gram (TF/IDF)	N/A	N/A	Various classifiers	Online forums
[79]	Raw text	Raw images	N/A	CNN (image) BERT (text)	Twitter
[87]	word2vec	VGG-16 Faster-RCNN	N/A	Ensemble (word2vec+ VGG16+ Faster-RCNN)	N/A
[148]	Psychological dictionary (Empath+ g Textblo)	N/A	N/A	PCA	Twitter
[68]	Sentiment, emotions, personal pronoun, absolutist words, negative words	N/A	N/A	LSTM	Twitter
[132]	GloveEmbed, Word2VecEmbed, Fastext and LIWC	N/A	N/A	BiLSTM	Reddit

using SentiStrength and **NRC Affect Intensity Lexicon (NRC-AIL)** classification. A semi-supervised learning scheme based on **Yet Another Two-stage Idea (YATSI)** classifier was used to identify suicide-related tweets. The authors of Reference [130] studied the possibility of automatic suicide notes identification from social media posts as document-level classification tasks. Specifically, they extracted suicide-related LIWC features from suicide notes and trained a Dilated LSTM suicide risk detection model. The n-gram analysis is evaluated in Reference [146] to prove that phrases related to suicidal tendencies and low social engagement are often present in suicide-related forums. The authors studied the transition toward social ideation related to different psychological states such as heightened self-focused attention, frustration, anxiety, hopelessness, or loneliness. Furthermore, they have extracted various textual features (e.g., TF-IDF) and compared the performance of CNN, LSTM, and LSTM-CNN combined models, in addition to various machine learning classifiers including SVM, XGBoost, RF, and NB. They found that LSTM-CNN combined model significantly outperforms other models achieving 93.8% accuracy and 92.8% F1 score. Table 4 summarizes the recent social media-based suicide detection schemes.

Table 4. Recent Suicide Ideation Detection Schemes

Work	Textual Feature	Visual Feature	Behavioral Feature	Detection Technique	Social Network
[130]	LIWC	N/A	N/A	Dilated LSTM	Reddit
[126]	SentiStrength, NRC-AIL	N/A	N/A	YATSI	Twitter
[146]	N-gram, TF-IDF, BoW	N/A	N/A	LSTM-CNN	Reddit
[21]	BERT	Raw images	Posting time	knowledge graph	Weibo Reddit
[56]	BERT, Sentence-BERT, GUSE	N/A	N/A	DNN	Reddit
[127]	Sentence-BERT	N/A	N/A	T-LSTM	Twitter
[83]	LDA, NMF	N/A	N/A	CNN	Twitter
[105]	Deep Contextualized Word Embedding (CWE)	N/A	N/A	ANN	Facebook
[116]	keywords, VADER sentiment	N/A	N/A	Radio Frequency (RF)	Twitter
[121]	keywords	N/A	N/A	Neural network	Twitter
[114]	TF-IDF, BoW	N/A	N/A	Various classifiers	Twitter
[129]	N/A	Image tags	N/A	Classifier	Instagram
[86]	Word embeddings	visual representation (ResNet)	N/A	GRU	Weibo
[75]	SentiWordNet, POS	N/A	Posting time	Statistical analysis	Twitter
[98]	TF-IDF, Word2Vec	N/A	N/A	Various classifiers	Vkontakte

6.3 Loneliness Detection

The outbreak of COVID-19 has resulted in distressing and unexpected social isolation for many people. Fear from the virus and social distancing rules affected people's mental health, which negatively impacted their feelings, mood, daily habits, and social relationships, which are essential elements of human mental well-being. Specifically, restrictions due to social distancing and quarantines increased feelings of loneliness and social anxiety [111]. Many works have leveraged textual features for loneliness detection from social media content. The authors of Reference [72] analyzed the patterns of loneliness expression in Twitter during the COVID-19 pandemic and pointed out key areas of loneliness expression across various communities. Specifically, they searched Twitter feeds for tweets that contain "COVID-19" and "loneliness" posted between May 1, 2020 and July 1, 2020. Following that, they applied topic modeling to extract topics discussed by lonely users, and used Hierarchical Modeling to distinguish overarching topics. Variations in the prevalence of these topics were analyzed over time and across the number of followers of Twitter users. In Reference [53], the authors selected users whose Twitter posts contained the words "lonely" or "alone" and compared these users to a control group selected by gender, age and time of posting. They also filtered the topics and studied patterns of users' posts and their relation with linguistic features of mental health, and studied the effect of language on the prediction of social media manifested loneliness. Similarly, Reference [7] proposed a loneliness detection system named LonelyText that applied SVM classifier to LDA topics that were extracted from Facebook dataset. Multimedia features are also useful in the context of loneliness and social anxiety detection.

The relationships between user loneliness and the color features of their Instagram photos were analyzed in Reference [71]. The analysis considers 25,394 Instagram photos in terms of color

diversity, colorfulness, and color harmony. The results suggest that the color diversity is negatively correlated with user loneliness, in particular romantic loneliness. Behavioural features have proven to carry rich information regarding lonely and social anxious users. The authors of Reference [169] studied the relationship between user loneliness and the type of social media they use. The study involved 155 Japanese university students who were divided into four groups based on their responses to loneliness questionnaires: Twitter users group, Twitter and Facebook users group, Twitter and Instagram users group, and users of all three social media sites. Following that, the effects of social media usage time, and the usage type on loneliness and well-being for each group were analyzed. The finding is that no social media usage has effects on loneliness or mental health associated with students who used only Twitter or both Twitter and Instagram. For students using both Twitter and Facebook, loneliness was reduced when they use Twitter and Facebook more frequently, but was increased when the students posted more tweets. Students of all three social media were lonelier and had lower levels of mental well-being when they used Facebook via computer longer. However, their access time of Facebook using mobile phones helped them decrease loneliness and improve their levels of mental well-being.

6.4 Anxiety Detection

A growing number of studies suggested that the number of people with anxiety increased during stay-at-home COVID-19 orders [44]. The study in Reference [25] investigated users suffering from social anxiety disorder using behavioral and social-network topological features. The study first collected ground-truth data from various social media sources and identified anxious users with the help of mental healthcare professionals. Next, multiple features were used, such as TF-IDF word scores, negative self-disclosure, sentiment score, parasocial relationship, and social event attending. Finally, an SVM classifier was applied to identify users with social anxiety disorder. Similarly, the study in Reference [41] observed that the application of SVM classifier to behavior and interaction features can predict the user's social anxiety disorder status with 79% accuracy and 84% area under the receiver-operating characteristic curve. In the same vein, the authors of Reference [47] applied random forest and XGBoost classifier on YouTube comments for anxiety detection during the COVID-19 pandemic. The relationship between various mental disorders and online activities is investigated in Reference [167]. This study analyzed Flickr dataset by applying a multi-modal detection technique using multi-modal features, namely, textual features, visual features (e.g., color distribution, presence of faces and objects), and metadata features and their relation to mental well-being. It was found that users suffering from mental distress. such as social anxiety, have more posting activity during the afternoon and evening as compared to healthy users.

6.5 Stress Detection

The impact of COVID-19 lockdown measures on the individual's psychological states in Italy and China was investigated in Reference [143]. First the social media data was extracted from Twitter users in Lombardy, Italy and from Weibo users in Wuhan, China—both using the geo-location filter. Next all users' posted content was extracted two weeks before and after the lockdown for each region (Lombardy and Wuhan). Following that, the psycholinguistic features of these posts were extracted using the Italian version and Simplified Chinese version of LIWC. Finally, various Wilcoxon tests were performed to study the changes in the psycholinguistic features of the posts before and after the lockdown in Lombardy and Wuhan, respectively. The findings suggest that the users focused more on "home" and showed a higher level of cognitive process after a lockdown in both Lombardy and Wuhan. After the lockdown, the stress level decreased, the focus on leisure increased in Lombardy, while the focus on group, religion, and emotions became more common in Wuhan.

A three-level framework for stress detection is proposed in Reference [159]. It learns the personalized stress representations following increasingly detailed processing, i.e., from the generic mass level to the group level and finally to the individual level. The mass level is dedicated to mining the generic stress features extracted from people's linguistic and visual posts using a two-layer attention mechanism. In this layer the authors of Reference [159] used the GRU model, which inputs the embedding vector of each word and outputs the hidden representation. The group level leverages a Graph Neural Network to learn the social media group affiliation features. The individual level incorporates the user's personality traits into the proposed stress detection framework. The performance study on 1,324,121 posts collected from the social media accounts of 2,059 Weibo users shows that the proposed framework can achieve over 90% accuracy detecting the stressed users. In Reference [67], LDA is applied to non-stressed related tweets from Twitter to detect stress among the users, categorizing the tweets into two user groups – stressed and non-stressed. The results suggested that applying LDA to stress detection yields better performance than SVM-based detection. A hybrid ontology is proposed in Reference [147] for stress detection that captures users' keyword-matching search process used in social media to identify stress-related messages. In Reference [93], The CNN classifier is applied to linguistic, visual and interaction features extracted from Twitter data to identify stressed individuals.

6.6 Other Mental Disorders

The consequences of COVID-19 pandemic not only caused mental distress to previously healthy people but also may have aggravated and worsened the situation for people with previous mental issues, such as Schizophrenia and PTSD. The authors of Reference [77] analyzed 19,224 schizophrenia-related Weibo posts and extracted psycholinguistic features of the Simplified Chinese version of LIWC from each post. They applied SVM, NB, and logistic model trees to identify schizophrenic users. Similarly, in Reference [17], the authors collected 3,404,959 Facebook messages and 142,390 images of 223 participants with schizophrenia spectrum disorders. They analyzed linguistic and visual features uploaded up to 18 months before the first hospitalization using machine learning and built classifiers that identify schizophrenic users from healthy users. Specifically, they used LIWC for linguistic features; and for visual features, they extracted ten, nine, two, and one features related to hue, saturation, sharpness, and contrast, respectively, in addition to the number of pixels, width, height, and brightness. They applied logistic regression to these features to determine schizophrenic users. How to leverage user's personality traits to improve the recommendation system of social networks, is explored in References [32, 35, 103]. The study in Reference [66] identified a lexicon of terms that are more common among veterans with PTSD prone to **Angry Outburst (AOB)**-specific pre-crisis data in social media posts. This study collected tweet datasets of the general population and veterans with PTSD; the PTSD-related topics are searched by searching specific hashtags such as "#PTSD" and analyzed tweets posted by 6,000 veterans. The study in Reference [96] crawled 17,159 Reddit posts during the COVID-19 pandemic to identify users with PTSD by analyzing unstructured user data using RoBERTa, which has a similar architecture to BERT with an improved pre-training procedure.

7 FINDINGS AND FUTURE DIRECTIONS

Social media analysis is an effective mechanism for mental health assessment. As shown in this survey article, there exist various OSNs that contain different types of data used to extract various features by exploiting different detection techniques. Twitter and Weibo are the most used OSNs, due to the short length of microblogging posts that make them suitable for text processing. They contain an extract of the users' thoughts without any additional wordiness, unlike Facebook posts for instance. The psychology literature provides strong evidence of the correlation

between language usage and psychological distress. While most of the surveyed works in this article adopted textual features as markers for psychological distress, there is a clear trend that visual and behavioural features are increasingly used to supplement textual information for more accurate detection of psychological states. Unlike textual features that are relatively easy to process and transform into machine-readable format, visual and behavioral features usually require huge data and expensive computation. Nevertheless, they offer the capabilities to extract users' fine-grained level features, thus allowing inference of nuance changes of psychological states. Machine learning detection techniques usually yield good performance with a small set of data collected from the participants' social media accounts, while deep learning techniques have the upper hand when dealing with a large dataset of the general population generally collected by keywords or hashtag crawling. Specifically, LSTM and its variants are widely used for textual feature-based mental distress detection and CNN is mostly used for processing visual features (e.g., detecting the number of faces in Instagram photos) as an indicator of social exposure. Few works have used multi-modal data from various OSNs, however, they used a relatively small dataset, and therefore the benefits of such an approach are not achieved yet. Developing a multi-modal framework that incorporates various user activities is a promising future direction as different individuals tend to manifest mental status differently in their behaviors. For example, some users may manifest their mental status verbally or aesthetically, i.e., by linguistic and language usage changes, photo colors, and filters preferences, while others may display more passive symptoms, which can be monitored and derived through the music listening history and online activities.

8 OPEN ISSUES AND CHALLENGES

Leveraging people's social media content as a mental healthcare data source for assessing disorders and intervention confers various benefits, such as reduced recall bias, cost efficiency, and large-scale population-level assessment. However, relying on social media as a data source will pose significant challenges and ethical dilemmas that must be addressed to ensure that such technology is ready for population-level exploration. This section discusses these limitations and challenges.

8.1 Privacy Concerns

From both research and deployment perspectives, the subjects' privacy is one of the most challenging issues when dealing with social media data, particularly given that social media apps are widely available on wearable devices [110]. From a research point of view, the studied users' privacy might be affected throughout the data processing stages. Although the studied datasets are publicly available most of the time, problems can arise when users' personal attributes can be predicted, and the identity of the users can be revealed. Various jurisdictions require certain conditions for research to avoid compromising users' privacy. The most common procedure is that the researchers must acquire ethical approval or exemption from their Institutional Review Board prior to the study. Additionally, they must obtain informed consent from the users, when possible, and protect and anonymize sensitive data during the research stages. Moreover, they need to be careful when linking the data across sites is necessary. Finally, when sharing their data, they need to make sure that other researchers also adhere to the same privacy guidelines [14]. Researchers can rely on public social media datasets for mental healthcare research if they ensure the preservation of users' confidentiality. The privacy concern is more challenging when applying such mental healthcare solutions on the existing social media at a large-scale. Most of the OSNs store the users' data on their servers, obtaining the users' consent to process their data for service enhancement. However, additional consent is still required to use their data for mental health purposes.

8.2 Detection Certainty

Social media-based mental healthcare heavily depends on NLP, ML, and deep learning. Unfortunately, these computing techniques are not fully mature for deployment in a large scale with sensitive applications like mental assessment. Inaccurate assessment of users with a mental disorder may lead to undesirable consequences, such as false positives where the system falsely detects nonexistent mental distress, and false-negative where the system misses the detection of mentally distressed users. Although many researchers have achieved high accuracy for detecting various psychological conditions (e.g., 91% accuracy for depression detection [145], 93.8% accuracy of detecting suicide ideation [146], and 90% accuracy of stress detection [159]), the experiments yielding these results are limited; good accuracy is not guaranteed when these technologies are applied to a large-scale population. Besides the technological challenges to increasing the detection accuracy of these systems, another problem is the credibility of the user-generated content. In many cases, the users' portrayed emotions and behaviors on social media are not necessarily a reflection of their actual emotional and psychological status.

8.3 Public Acceptability

In addition to the privacy concern, the lack of public support and acceptability of using their social media data for monitoring their mental health is yet another challenge. Previous studies show that the users have worries that these technologies can be used against them. For example, many users have expressed concerns that mental assessment using their public social media data may negatively affect their credit card or insurance applications, or influence their employment career [102]. This fear is backed by the fact that only a small portion of the population suffer from mental disorders. However, with the recent lifestyle changes brought along with the spread of COVID-19, people have started to release the importance of mental healthcare. Now is the best time to convince the public that social media-based mental healthcare is needed more than ever before, and to show them the advantages of such an approach compared to conventional mental healthcare, such as non-invasive assessment.

9 CONCLUSION

In this article, we have provided a comprehensive overview of the state-of-the-art research in the field of using social media for mental distress assessment. We have extended the scope of previous surveys to incorporate the latest social media analysis techniques and their applications to monitoring and detection of multiple mental health conditions. We have particularly focused on the relevant studies that emerged during the COVID-19 pandemic. Specifically, we have classified psychological-related features that can be extracted from the social media content and reviewed mental distress detection techniques, including machine learning- and deep learning-based mental distress prediction models. We have reviewed recent studies and categorized these works according to their feature extraction and detection techniques. This survey also highlights the challenges of mental disorder detection using social media data, including privacy and ethical concerns, as well as the technical challenges of scaling and deploying such systems, and discusses the lessons learned over the past few years.

REFERENCES

[1] Alison Abbott. 2021. COVID's mental-health toll: Scientists track surge in depression. *Nature* 590, 7845 (2021), 194–195.
[2] Hussain Ahmad, Muhammad Zubair Asghar, Fahad M. Alotaibi, and Ibrahim A. Hameed. 2020. Applying deep learning technique for depression classification in social media text. *J. Med. Imag. Health Inform.* 10, 10 (2020), 2446–2451.

[3] Ravinder Ahuja, Aakarsha Chug, Shruti Kohli, Shaurya Gupta, and Pratyush Ahuja. 2019. The impact of features extraction on the sentiment analysis. *Procedia Comput. Sci.* 152 (2019), 341–348. DOI : https://doi.org/10.1016/j.procs.2019.05.008

[4] Eatedal Alabdulkreem. 2021. Prediction of depressed Arab women using their tweets. *J. Decis. Syst.* (2021), 1–16.

[5] Sreejith Alathur, Naganna Chetty, Rajesh R. Pai, Vishal Kumar, and Sahraoui Dhelim. 2022. Hate and false metaphors: Implications to emerging e-participation environment. *Future Internet* 14, 11 (Oct. 2022), 314. DOI : https://doi.org/10.3390/fi14110314

[6] Norah Saleh Alghamdi, Hanan A. Hosni Mahmoud, Ajith Abraham, Samar Awadh Alanazi, and Laura Garcia-Hernandez. 2020. Predicting depression symptoms in an Arabic psychological forum. *IEEE Access* 8 (2020), 57317–57334.

[7] Mawulolo K. Ameko, Sonia Baee, and Laura E. Barnes. 2021. LonelyText: A short messaging-based classification of loneliness. Retrieved from https://arXiv:2101.09138.

[8] Anietie Andy, Garrick Sherman, and Sharath Chandra Guntuku. 2022. Understanding the expression of loneliness on Twitter across age groups and genders. *Plos One* 17, 9 (2022), e0273636.

[9] Randy P. Auerbach, Jordi Alonso, William G. Axinn, Pim Cuijpers, David D. Ebert, Jennifer Greif Green, Irving Hwang, Ronald C. Kessler, Howard Liu, Philippe Mortier, et al. 2016. Mental disorders among college students in the World Health Organization world mental health surveys. *Psychol. Med.* 46, 14 (2016), 2955–2970.

[10] Rahn Kennedy Bailey, Holly L. Blackmon, and Francis L. Stevens. 2009. Major depressive disorder in the African American population: Meeting the challenges of stigma, misdiagnosis, and treatment disparities. *J. Natl. Med. Assoc.* 101, 11 (2009), 1084–1089.

[11] Anil Bandhakavi, Nirmalie Wiratunga, Deepak Padmanabhan, and Stewart Massie. 2017. Lexicon based feature extraction for emotion text classification. *Pattern Recogn. Lett.* 93 (2017), 133–142.

[12] Leah K. Barr, Jeffrey H. Kahn, and W. Joel Schneider. 2008. Individual differences in emotion expression: Hierarchical structure and relations with psychological distress. *J. Soc. Clin. Psychol.* 27, 10 (2008), 1045–1077.

[13] Aaron T. Beck, Robert A. Steer, and Gregory Brown. 1996. Beck depression inventory–II. *Psychol. Assess.* (1996). https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft00742-000.

[14] Adrian Benton, Glen Coppersmith, and Mark Dredze. 2017. Ethical research protocols for social media health research. In *Proceedings of the 1st ACL Workshop on Ethics in Natural Language Processing*. 94–102.

[15] Uffaq Bilal and Farhan Hassan Khan. 2019. An analysis of depression detection techniques from online social networks. In *Proceedings of the International Conference on Intelligent Technologies and Applications*. 296–308.

[16] Michael L. Birnbaum, Sindhu Kiranmai Ernala, Asra F. Rizvi, Munmun De Choudhury, and John M. Kane. 2017. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *J. Med. Internet Res.* 19, 8 (2017), e289.

[17] Michael L. Birnbaum, Raquel Norel, Anna Van Meter, Asra F. Ali, Elizabeth Arenare, Elif Eyigoz, Carla Agurto, Nicole Germano, John M. Kane, and Guillermo A. Cecchi. 2020. Identifying signals associated with psychiatric illness utilizing language and images posted to Facebook. *npj Schizophrenia* 6, 1 (Dec. 2020), 38. DOI : https://doi.org/10.1038/s41537-020-00125-0

[18] David E. Bloom, Elizabeth Cafiero, Eva Jané-Llopis, Shafika Abrahams-Gessel, Lakshmi Reddy Bloom, Sana Fathima, Andrea B. Feigl, Tom Gaziano, Ali Hamandi, Mona Mowafi, O'Farrell Danny, and Emre. 2012. *The Global Economic Burden of Noncommunicable Diseases.* Program on the Global Demography of Aging, PGDA Working Papers. https://ideas.repec.org/p/gdm/wpaper/8712.html.

[19] Sergio G. Burdisso, Marcelo Errecalde, and Manuel Montes-y Gómez. 2019. A text classification framework for simple and effective early depression detection over social media streams. *Expert Syst. Appl.* 133 (2019), 182–197.

[20] Xiaofeng Cai, Huansheng Ning, Sahraoui Dhelim, Rongjuan Zhou, Tao Zhang, Yang Xu, and Yueliang Wan. 2020. Robot and its living space A roadmap for robot development based on the view of living space. *Dig. Commun. Netw.* (Dec. 2020). DOI : https://doi.org/10.1016/j.dcan.2020.12.001

[21] Lei Cao, Huijun Zhang, and Ling Feng. 2020. Building and using personal knowledge graph to improve suicidal ideation detection on social media. *IEEE Trans. Multimedia* 24 (2020), 87–102.

[22] Gema Castillo-Sánchez, Gonçalo Marques, Enrique Dorronzoro, Octavio Rivera-Romero, Manuel Franco-Martín, and Isabel De la Torre-Díez. 2020. Suicide risk assessment using machine learning and social networks: A scoping review. *J. Med. Syst.* 44, 12 (2020), 205. DOI : https://doi.org/10.1007/s10916-020-01669-5

[23] Patricia A. Cavazos-Rehg, Melissa J. Krauss, Shaina Sowles, Sarah Connolly, Carlos Rosas, Meghana Bharadwaj, and Laura J. Bierut. 2016. A content analysis of depression-related tweets. *Comput. Hum. Behav.* 54 (Jan. 2016), 351–357. DOI : https://doi.org/10.1016/j.chb.2015.08.023

[24] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: A critical review. *NPJ Dig. Med.* 3, 1 (2020), 1–11.

[25] Ming-Yi Chang and Chih-Ying Tseng. 2020. Detecting social anxiety with online social network data. In *Proceedings of the 21st IEEE International Conference on Mobile Data Management (MDM'20)*. 333–336.

[26] Hongguang Chen and Konglai Zhang. 2020. Insight into the psychological problems on the epidemic of COVID-19 in China by online searching behaviors. *J. Affect. Disorders* 276 (2020), 1093.

[27] Qijin Cheng, Tim M. H. Li, Chi-Leung Kwok, Tingshao Zhu, and Paul S. F. Yip. 2017. Assessing suicide risk and emotional distress in Chinese social media: A text mining and machine learning study. *J. Med. Internet Res.* 19, 7 (2017), e243.

[28] Rion Brattig Correia, Ian B. Wood, Johan Bollen, and Luis M. Rocha. 2020. Mining social media data for biomedical signals and health-related behavior. *Annu. Rev. Biomed. Data Sci.* 3 (2020), 433–458.

[29] John G. Cull and Wayne S. Gill. 1988. Suicide probability scale (SPS). *Journal of Consulting and Clinical Psychology* (1988). https://doi.org/10.1037/t01198-000

[30] Emma Curran, Michael Rosato, Janine Cooper, Christine A. McGarrigle, and Gerard Leavey. 2019. Symptom profiles of late-life anxiety and depression: The influence of migration, religion and loneliness. *Depress. Anxiety* 36, 9 (2019), 824–833.

[31] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 7.

[32] Sahraoui Dhelim, Nyothiri Aung, Mohammed Amine Bouras, Huansheng Ning, and Erik Cambria. 2021. A survey on personality-aware recommendation systems. *Artific. Intell. Rev.* (Sep. 2021). DOI : https://doi.org/10.1007/s10462-021-10063-7

[33] Sahraoui Dhelim, Nyothiri Aung, Tahar Kechadi, Huansheng Ning, Liming Chen, and Abderrahmane Lakas. 2022. Trust2Vec: Large-scale IoT trust management system based on signed network embeddings. *IEEE Internet Things J.* 10, 1 (2022), 553–562.

[34] Sahraoui Dhelim, Nyothiri Aung, and Huansheng Ning. 2020. Mining user interest based on personality-aware hybrid filtering in social networks. *Knowl.-Based Syst.* 206 (Oct. 2020), 106227. DOI : https://doi.org/10.1016/j.knosys.2020.106227

[35] Sahraoui Dhelim, Liming Chen, Nyothiri Aung, Wenyin Zhang, and Huansheng Ning. 2022. A hybrid personality-aware recommendation system based on personality traits and types models. *J. Ambient Intell. Human. Comput.* (July 2022), 1–14. DOI : https://doi.org/10.1007/s12652-022-04200-5

[36] Sahraoui Dhelim, Liming Chen, Huansheng Ning, and Chris Nugent. 2022. Artificial intelligence for suicide assessment using Audiovisual Cues: A review. *Artific. Intell. Rev.* (Nov. 2022), 1–28. DOI : https://doi.org/10.1007/s10462-022-10290-6

[37] Sahraoui Dhelim, Huansheng Ning, and Nyothiri Aung. 2020. ComPath: User interest mining in heterogeneous signed social networks for Internet of people. *IEEE Internet Things J.* 8, 8 (2020), 7024–7035. DOI : https://doi.org/10.1109/JIOT.2020.3037109

[38] Sahraoui Dhelim, Huansheng Ning, Nyothiri Aung, Runhe Huang, and Jianhua Ma. 2021. Personality-aware product recommendation system based on user interests mining and metapath discovery. *IEEE Trans. Comput. Soc. Syst.* 8, 1 (Feb. 2021), 86–98. DOI : https://doi.org/10.1109/TCSS.2020.3037040

[39] Tiance Dong, Chenxi Liang, and Xu He. 2017. Social media and internet public events. *Telemat. Informat.* 34, 3 (2017), 726–739.

[40] David Donovan. 2016. Mood, emotions, and emojis: Conversations about health with young people. *Mental Health Pract.* 20, 2 (Oct. 2016), 23–26. DOI : https://doi.org/10.7748/mhp.2016.e1143

[41] Sarmistha Dutta and Munmun De Choudhury. 2020. Characterizing anxiety disorders with online social and interactional networks. In *Proceedings of the HCI International 2020—Late Breaking Papers: Interaction, Knowledge and Social Media*, Constantine Stephanidis, Gavriel Salvendy, June Wei, Sakae Yamamoto, Hirohiko Mori, Gabriele Meiselwitz, Fiona Fui-Hoon Nah, and Keng Siau (Eds.). Springer International Publishing, Cham, 249–264.

[42] Samuel E. Ehrenreich and Marion K. Underwood. 2016. Adolescents' internalizing symptoms as predictors of the content of their Facebook communication and responses received from peers. *Translat. Iss. Psychol. Sci.* 2, 3 (2016), 227.

[43] Daniel Eisenberg, Sarah E. Gollust, Ezra Golberstein, and Jennifer L. Hefner. 2007. Prevalence and correlates of depression, anxiety, and suicidality among university students. *Amer. J. Orthopsych.* 77, 4 (2007), 534–542.

[44] Susanna Every-Palmer, Matthew Jenkins, Philip Gendall, Janet Hoek, Ben Beaglehole, Caroline Bell, Jonathan Williman, Charlene Rapsey, and James Stanley. 2020. Psychological distress, anxiety, family violence, suicidality, and wellbeing in New Zealand during the COVID-19 lockdown: A cross-sectional study. *PLoS One* 15, 11 (2020), e0241658.

[45] Weiguo Fan and Michael D. Gordon. 2014. The power of social media analytics. *Commun. ACM* 57, 6 (2014), 74–81.

[46] Boumahdi Fatima, Madani Amina, Rezoug Nachida, and Hentabli Hamza. 2020. A mixed deep learning based model to early detection of depression. *J. Web Eng.* (2020), 429–456.

[47] Yuli Fauziah, Shoffan Saifullah, and Agus Sasmito Aribowo. 2020. Design text mining for anxiety detection using machine learning based-on social media data during COVID-19 pandemic. In *Proceedings of the LPPM UPN "Veteran" Yogyakarta Conference Series 2020—Engineering and Science Series*, Vol. 1. 253–261.

[48] Johannes Feldhege, Markus Moessner, and Stephanie Bauer. 2020. Who says what? Content and participation characteristics in an online depression community. *J. Affect. Disorders* 263 (Feb. 2020), 521–527. DOI:https://doi.org/10.1016/j.jad.2019.11.007

[49] Alize J. Ferrari, Rosana E. Norman, Greg Freedman, Amanda J. Baxter, Jane E. Pirkis, Meredith G. Harris, Andrew Page, Emily Carnahan, Louisa Degenhardt, Theo Vos, et al. 2014. The burden attributable to mental and substance use disorders as risk factors for suicide: Findings from the Global Burden of Disease Study 2010. *PloS One* 9, 4 (2014), e91936.

[50] Finola Ferry, Brendan Bunting, Michael Rosato, Emma Curran, and Gerard Leavey. 2021. The impact of reduced working on mental health in the early months of the COVID-19 pandemic: Results from the understanding society COVID-19 study. *J. Affect. Disorders* 287 (2021), 308–315.

[51] Felipe T. Giuntini, Mirela T. Cazzolato, Maria de Jesus Dutra dos Reis, Andrew T. Campbell, Agma J M Traina, and Jo Ueyama. 2020. A review on recognizing depression in social networks: Challenges and opportunities. *J. Amb. Intell. Human. Comput.* 11, 11 (2020), 4713–4729.

[52] Sharath Chandra Guntuku, Daniel Preotiuc-Pietro, Johannes C. Eichstaedt, and Lyle H. Ungar. 2019. What Twitter profile and posted images reveal about depression and anxiety. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 236–246.

[53] Sharath Chandra Guntuku, Rachelle Schneider, Arthur Pelullo, Jami Young, Vivien Wong, Lyle Ungar, Daniel Polsky, Kevin G. Volpp, and Raina Merchant. 2019. Studying expressions of loneliness in individuals using Twitter: An observational study. *BMJ Open* 9, 11 (2019), e030355.

[54] Sharath Chandra Guntuku, H. Andrew Schwartz, Adarsh Kashyap, Jessica S. Gaulton, Daniel C. Stokes, David A. Asch, Lyle H. Ungar, and Raina M. Merchant. 2020. Variability in language used on social media prior to hospital visits. *Sci. Rep.* 10, 1 (2020), 1–9.

[55] Sharath Chandra Guntuku, Garrick Sherman, Daniel C. Stokes, Anish K. Agarwal, Emily Seltzer, Raina M. Merchant, and Lyle H. Ungar. 2020. Tracking mental health and symptom mentions on Twitter during COVID-19. *J. Gen. Internal Med.* 35, 9 (2020), 2798–2800.

[56] Ayaan Haque, Viraaj Reddi, and Tyler Giallanza. 2021. Deep learning for suicide and depression identification with unsupervised label correction. Retrieved from https://arXiv:2102.09427.

[57] Keith M. Harris, Jia-Jia Syu, Owen D. Lello, Y. L. Eileen Chew, Christopher H. Willcox, and Roger H. M. Ho. 2015. The ABC's of suicide risk assessment: Applying a tripartite approach to individual evaluations. *PLoS One* 10, 6 (2015), e0127442.

[58] Saeed Hassanpour, Naofumi Tomita, Timothy DeLise, Benjamin Crosier, and Lisa A. Marsch. 2019. Identifying substance use risk based on deep neural networks and Instagram social media data. *Neuropsychopharmacology* 44, 3 (2019), 487–494.

[59] Nicholas S. Holtzman et al. 2017. A meta-analysis of correlations between depression and first person singular pronoun use. *J. Res. Personal.* 68 (2017), 63–68.

[60] Zhiyuan Hou, Fanxing Du, Hao Jiang, Xinyu Zhou, and Leesa Lin. 2020. Assessment of public attention, risk perception, emotional and behavioural responses to the COVID-19 outbreak: social media surveillance in China. Cold Spring Harbor Laboratory Press. https://www.medrxiv.org/content/early/2020/03/17/2020.03.14.20035956.

[61] Clayton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 8.

[62] Md Rafiqul Islam, Abu Raihan M. Kamal, Naznin Sultana, Robiul Islam, Mohammad Ali Moni, et al. 2018. Detecting depression using k-nearest neighbors classification technique. In *Proceedings of the International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2'18)*. 1–4.

[63] Nicholas C. Jacobson, Damien Lekkas, George Price, Michael V. Heinz, Minkeun Song, A. James O'Malley, and Paul J. Barr. 2020. Flattening the mental health curve: COVID-19 stay-at-home orders are associated with alterations in mental health search behavior in the United States. *JMIR Mental Health* 7, 6 (2020), e19347.

[64] Hamed Jelodar, Yongli Wang, Rita Orji, and Shucheng Huang. 2020. Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. *IEEE J. Biomed. Health Inform.* 24, 10 (2020), 2733–2742.

[65] Shaoxiong Ji, Shirui Pan, Xue Li, Erik Cambria, Guodong Long, and Zi Huang. 2020. Suicidal ideation detection: A review of machine learning methods and applications. *IEEE Trans. Comput. Soc. Syst.* 8, 1 (2020), 214–226.

[66] Nadiyah Johnson, Joseph Coelho, Md Fitrat Hossain, Thomas Kissane, Wylie Frydrychowi, Madiraju Praveen, Zeno Franco, Katinka Hooyer, Priyanka Annapureddy, and Sheikh Iqbal Ahamed. 2020. Understanding veterans expression of anger using social media analysis. In *Proceedings of the IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC'20)*. 1689–1694.

[67] Aysha Khan and Rashid Ali. 2020. Stress detection from Twitter posts using LDA. *Int. J. High Perform. Comput. Netw.* 16, 2-3 (2020), 137–147.

[68] Binti Kholifah, Iwan Syarif, and Tessy Badriyah. 2020. Mental disorder detection via social media mining using deep learning. *Kinetik: Game Technol., Info. Syst., Comput. Netw., Comput., Electro., Control* 5, 4 (Nov. 2020), 309–316. DOI : https://doi.org/10.22219/kinetik.v5i4.1120

[69] Jiin Kim, Zara Ahmad, Yena Lee, Flora Nasri, Hartej Gill, Roger Mclntyre, Lee Phan, and Leanna Lui. 2021. Systematic review of the validity of screening depression through Facebook, Twitter, and Instagram. *J. Affect. Disord.* (Feb. 2021). DOI : https://doi.org/10.1016/j.jad.2020.08.091

[70] Jina Kim, Jieon Lee, Eunil Park, and Jinyoung Han. 2020. A deep learning model for detecting mental illness from user content on social media. *Sci. Rep.* 10, 1 (Dec. 2020), 11846. DOI : https://doi.org/10.1038/s41598-020-68764-y

[71] Jang Hyun Kim and Yunhwan Kim. 2019. Instagram user characteristics and the color of their photos: Colorfulness, color diversity, and color harmony. *Info. Process. Manage.* 56, 4 (July 2019), 1494–1505. DOI : https://doi.org/10.1016/j.ipm.2018.10.018

[72] Jing Xuan Koh and Tau Ming Liew. 2022. How loneliness is talked about in social media during COVID-19 pandemic: Text mining of 4,492 Twitter feeds. *J. Psych. Res.* 145 (2022), 317–324. https://www.sciencedirect.com/science/article/pii/S0022395620310748.

[73] Kurt Kroenke, Robert L. Spitzer, and Janet B. W. Williams. 2001. The PHQ-9: Validity of a brief depression severity measure. *J. Gen. Internal Med.* 16, 9 (2001), 606–613.

[74] E. Megan Lachmar, Andrea K. Wittenborn, Katherine W. Bogen, and Heather L. McCauley. 2017. # MyDepression-LooksLike: Examining public discourse about depression on Twitter. *JMIR Mental Health* 4, 4 (2017), e43.

[75] C. Lalrinmawii, Somen Debnath, et al. 2020. Analysis of post centric suicidal expressions and classification on the social media post: Twitter. In *Proceedings of the 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT'20)*. 1–5.

[76] Victor Leiva and Ana Freire. 2017. Towards suicide prevention: Early detection of depression on social media. In *Proceedings of the International Conference on Internet Science*. 428–436.

[77] Ang Li, Dongdong Jiao, Xiaoqian Liu, and Tingshao Zhu. 2020. A comparison of the psycholinguistic styles of schizophrenia-related stigma and depression-related stigma on social media: Content analysis. *J. Med. Internet Res.* 22, 4 (2020), e16470.

[78] Sijia Li, Yilin Wang, Jia Xue, Nan Zhao, and Tingshao Zhu. 2020. The impact of COVID-19 epidemic declaration on psychological consequences: A study on active weibo users. *Int. J. Environ. Res. Public Health* 17, 6 (Mar. 2020), 2032. DOI : https://doi.org/10.3390/ijerph17062032

[79] Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li, Jing Mei, Jie Zhou, and Henry Leung. 2020. SenseMood: Depression detection on social media. In *Proceedings of the International Conference on Multimedia Retrieval*. ACM, New York, NY, 407–411. DOI : https://doi.org/10.1145/3372278.3391932

[80] Dexi Liu, Qi Fu, Changxuan Wan, Xiping Liu, Tengjiao Jiang, Guoqiong Liao, Xiangqing Qiu, and Renqiang Liu. 2020. Suicidal ideation cause extraction from social texts. *IEEE Access* 8 (2020), 169333–169351.

[81] Tingting Liu, Lyle H. Ungar, Brenda Curtis, Garrick Sherman, Kenna Yadeta, Louis Tay, Johannes C. Eichstaedt, and Sharath Chandra Guntuku. 2022. Head versus heart: Social media reveals differential language of loneliness from depression. *npj Mental Health Res.* 1, 1 (2022), 1–8.

[82] Daniel M. Low, Laurie Rumker, Tanya Talkar, John Torous, Guillermo Cecchi, and Satrajit S. Ghosh. 2020. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: Observational study. *J. Med. Internet Res.* 22, 10 (Oct. 2020), e22635. DOI : https://doi.org/10.2196/22635

[83] Jianhong Luo, Jingcheng Du, Cui Tao, Hua Xu, and Yaoyun Zhang. 2020. Exploring temporal suicidal behavior patterns on social media: Insight from Twitter analytics. *Health Info. J.* 26, 2 (2020), 738–752.

[84] Katerina Lup, Leora Trub, and Lisa Rosenthal. 2015. Instagram# instasad?: Exploring associations among instagram use, depressive symptoms, negative social comparison, and strangers followed. *Cyberpsychol., Behav. Soc. Netw.* 18, 5 (2015), 247–252.

[85] Meizhen Lv, Ang Li, Tianli Liu, and Tingshao Zhu. 2015. Creating a Chinese suicide dictionary for identifying suicide risk on social media. *PeerJ* 3 (2015), e1455.

[86] Yihua Ma and Youliang Cao. 2020. Dual attention based suicide risk detection on social media. In *Proceedings of the IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA'20)*. 637–640.

[87] Anshu Malhotra and Rajni Jindal. 2020. Multimodal deep learning based framework for detecting depression and suicidal behaviour by affective analysis of social media posts. *EAI Endorsed Trans. Pervas. Health Technol.* 6, 21 (2020), e1.

[88] Paulo Mann, Aline Paes, and Elton H. Matsushima. 2020. See and read: Detecting depression symptoms in higher education students using multimodal social media data. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14. 440–451.

[89] Davide Marengo, Michele Settanni, and Fabrizia Giannotta. 2019. Development and preliminary validation of an image-based instrument to assess depressive symptoms. *Psychiatry Res.* 279 (Sep. 2019), 180–185. DOI : https://doi.org/10.1016/j.psychres.2019.02.059

[90]   Rodrigo Martínez-Castaño, Juan C. Pichel, and David E. Losada. 2020. A big data platform for real time analysis of signs of depression in social media. *Int. J. Environ. Res. Public Health* 17, 13 (July 2020), 4752. DOI:https://doi.org/10.3390/ijerph17134752

[91]   Kimberly McManus, Emily K. Mallory, Rachel L. Goldfeder, Winston A. Haynes, and Jonathan D. Tatum. 2015. Mining Twitter data to improve detection of schizophrenia. *AMIA Summits Translat. Sci. Proc.* 2015 (2015), 122.

[92]   Yiwen Meng, William Farran Speier, Michael Ong, and Corey W. Arnold. 2020. HCET: Hierarchical clinical embedding with topic modeling on electronic health record for predicting depression. *IEEE J. Biomed. Health Inform.* 25, 4 (2020), 1265–1272. DOI:https://doi.org/10.1109/JBHI.2020.3004072

[93]   Shweta Meshram, Rajesh Babu, and Jayanth Adhikari. 2020. Detecting psychological stress using machine learning over social media interaction. In *Proceedings of the 5th International Conference on Communication and Electronics Systems (ICCES'20)*. 646–649.

[94]   David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, et al. 2010. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Int. J. Surg.* 8, 5 (2010), 336–341.

[95]   Megan Andreas Moreno, Dimitri A. Christakis, Katie G. Egan, Lauren A. Jelenchick, Elizabeth Cox, Henry Young, Hope Villiard, and Tara Becker. 2012. A pilot evaluation of associations between displayed depression references on Facebook and self-reported depression using a clinical scale. *J. Behav. Health Serv. Res.* 39, 3 (2012), 295–304.

[96]   Ankit Murarka, Balaji Radhakrishnan, and Sushma Ravichandran. 2020. Detection and classification of mental illnesses on social media using RoBERTa. Retrieved from https://arXiv:2011.11226.

[97]   Raza Ul Mustafa, Noman Ashraf, Fahad Shabbir Ahmed, Javed Ferzund, Basit Shahzad, and Alexander Gelbukh. 2020. A multiclass depression detection in social media based on sentiment analysis. In *Proceedings of the 17th International Conference on Information Technology–New Generations (ITNG'20)*. 659–662.

[98]   Sergazy Narynov, Kanat Kozhakhmet, Daniyar Mukhtarkhanuly, Aizhan Sambetbayeva, and Batyrkhan Omarov. 2020. Artificial intelligence in detecting suicidal content on russian-language social networks. In *Proceedings of the International Conference on Computational Collective Intelligence*. 811–820.

[99]   Sonya Negriff. 2019. Depressive symptoms predict characteristics of online social networks. *J. Adolesc. Health* 65, 1 (2019), 101–106.

[100] Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y. Ng. 2011. Multimodal deep learning. In *Proceedings of the International Conference on Machine Learning (ICML'11)*. 689–696.

[101] Thin Nguyen, Dinh Phung, Bo Dao, Svetha Venkatesh, and Michael Berk. 2014. Affective and content analysis of online depression communities. *IEEE Trans. Affect. Comput.* 5, 3 (July 2014), 217–226. DOI:https://doi.org/10.1109/TAFFC.2014.2315623

[102] Jennifer Nicholas, Sandersan Onie, and Mark E. Larsen. 2020. Ethics and privacy in social media research for mental health. *Curr. Psychiatry Rep.* 22, 12 (2020), 1–7.

[103] Huansheng Ning, Sahraoui Dhelim, and Nyothiri Aung. 2019. PersoNet: Friend recommendation system based on big-five personality traits and hybrid filtering. *IEEE Trans. Comput. Soc. Syst.* 6, 3 (June 2019), 394–402. DOI:https://doi.org/10.1109/TCSS.2019.2903857

[104] Yaakov Ophir, Christa S. C. Asterhan, and Baruch B. Schwarz. 2019. The digital footprints of adolescent depression, social rejection and victimization of bullying on Facebook. *Comput. Hum. Behav.* 91 (2019), 62–71.

[105] Yaakov Ophir, Refael Tikochinski, Christa S. C. Asterhan, Itay Sisso, and Roi Reichart. 2020. Deep neural networks detect suicide risk from textual Facebook posts. *Sci. Rep.* 10, 1 (2020), 1–10.

[106] Roger O'Sullivan, Brian Lawlor, Annette Burns, and Gerry Leavey. 2021. Will the pandemic reframe loneliness and social isolation?*Lancet Healthy Longev.* 2, 2 (2021), e54–e55.

[107] Sungkyu Park, Sang Won Lee, Jinah Kwak, Meeyoung Cha, and Bumseok Jeong. 2013. Activities on Facebook reveal the depressive state of users. *J. Med. Internet Res.* 15, 10 (Oct. 2013), e217. DOI:https://doi.org/10.2196/jmir.2718

[108] Zhichao Peng, Qinghua Hu, and Jianwu Dang. 2019. Multi-kernel SVM based depression recognition using social media data. *Int. J. Mach. Learn. Cybernet.* 10, 1 (2019), 43–57.

[109] Matthew Pittman and Brandon Reich. 2016. Social media and loneliness: Why an Instagram picture may be worth more than a thousand Twitter words. *Comput. Hum. Behav.* 62 (2016), 155–167.

[110] Ismini Psychoula, Liming Chen, and Oliver Amft. 2020. Privacy risk awareness in wearables and the internet of things. *IEEE Pervas. Comput.* 19, 3 (Aug. 2020), 60–66. DOI:https://doi.org/10.1109/MPRV.2020.2997616

[111] Malik Muhammad Qirtas, Evi Zafeiridi, Dirk Pesch, and Eleanor Bantry White. 2022. Loneliness and social isolation detection using passive sensing techniques: Scoping review. *JMIR mHealth uHealth* 10, 4 (2022), e34638.

[112] Darren Quinn, Liming Chen, and Maurice Mulvenna. 2012. Social network analysis: A survey. *Int. J. Ambient Comput. Intell.* 4, 3 (July 2012), 46–58. DOI:https://doi.org/10.4018/jaci.2012070104

[113] Darren Quinn, Liming Chen, Maurice Mulvenna, and Raymond Bond. 2016. Exploring the relationship between online social network site usage and the impact on quality of life for older and younger users: An interaction analysis. *J. Med. Internet Res.* 18, 9 (Sep. 2016), e245. DOI:https://doi.org/10.2196/jmir.5377

[114]   Syed Tanzeel Rabani, Qamar Rayees Khan, and Akib Mohi Ud Din Khanday. 2020. Detection of suicidal ideation on Twitter using machine learning & ensemble approaches. *Baghdad Sci. J.* 17, 4 (2020), 1328.

[115]   Lenore Sawyer Radloff. 1977. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Measure.* 1, 3 (1977), 385–401.

[116]   E. Rajesh Kumar, K. V. S. N. Rama Rao, Soumya Ranjan Nayak, and Ramesh Chandra. 2020. Suicidal ideation prediction in Twitter data using machine learning techniques. *J. Interdisc. Math.* 23, 1 (2020), 117–125.

[117]   Andrew G. Reece and Christopher M. Danforth. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Sci.* 6 (2017), 1–12.

[118]   Andrew G. Reece, Andrew J. Reagan, Katharina L. M. Lix, Peter Sheridan Dodds, Christopher M. Danforth, and Ellen J. Langer. 2017. Forecasting the onset and course of mental illness with Twitter data. *Sci. Rep.* 7, 1 (Dec. 2017), 13006. DOI: https://doi.org/10.1038/s41598-017-12961-9

[119]   Benjamin J. Ricard, Lisa A. Marsch, Benjamin Crosier, and Saeed Hassanpour. 2018. Exploring the utility of community-generated social media content for detecting depression: An analytical study on Instagram. *J. Med. Internet Res.* 20, 12 (2018), e11817.

[120]   Mariana Rodrigues Makiuchi, Tifani Warnita, Kuniaki Uto, and Koichi Shinoda. 2019. Multimodal fusion of BERT-CNN and gated CNN representations for depression detection. In *Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop.* 55–63.

[121]   Arunima Roy, Katerina Nikolitch, Rachel McGinn, Safiya Jinah, William Klement, and Zachary A. Kaminsky. 2020. A machine learning approach predicts future risk to suicidal ideation from social media data. *NPJ Dig. Med.* 3, 1 (2020), 1–12.

[122]   Budhaditya Saha, Thin Nguyen, Dinh Phung, and Svetha Venkatesh. 2016. A framework for classifying online mental health-related communities with an Interest in depression. *IEEE J. Biomed. Health Inform.* 20, 4 (Jul. 2016), 1008–1015. DOI: https://doi.org/10.1109/JBHI.2016.2543741

[123]   Koustuv Saha, John Torous, Eric D. Caine, Munmun De Choudhury, et al. 2020. Psychosocial effects of the COVID-19 pandemic: Large-scale quasi-experimental study on social media. *J. Med. Internet Res.* 22, 11 (2020), e22600.

[124]   Flora Sakketou and Nicholas Ampazis. 2020. A constrained optimization algorithm for learning GloVe embeddings with semantic lexicons. *Knowl.-Based Syst.* 195 (2020), 105628.

[125]   Yaser Sarikhani, Peivand Bastani, Mohammad Rafiee, Zahra Kavosi, and Ramin Ravangard. 2021. Key barriers to the provision and utilization of mental health services in low-and middle-income countries: A scope study. *Commun. Mental Health J.* 57 (2021), 836–852.

[126]   Samer Muthana Sarsam, Hosam Al-Samarraie, Ahmed Ibrahim Alzahrani, Waleed Alnumay, and Andrew Paul Smith. 2021. A lexicon-based approach to detecting suicide-related messages on Twitter. *Biomed. Signal Process. Control* 65 (2021), 102355.

[127]   Ramit Sawhney, Harshit Joshi, Saumya Gandhi, and Rajiv Shah. 2020. A time-aware transformer based model for suicide ideation detection on social media. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP'20).* 7685–7697.

[128]   Shekhar Saxena, Michelle Funk, and Dan Chisholm. 2013. World health assembly adopts comprehensive mental health action plan 2013–2020. *Lancet* 381, 9882 (2013), 1970–1971.

[129]   Sebastian Scherr, Florian Arendt, Thomas Frissen, and José Oramas. 2020. Detecting intentional self-harm on Instagram: Development, testing, and validation of an automatic image-recognition algorithm to discover cutting-related posts. *Soc. Sci. Comput. Rev.* 38, 6 (2020), 673–685.

[130]   Annika Marie Schoene, Alexander P. Turner, Geeth De Mel, and Nina Dethlefs. 2023. Hierarchical multiscale recurrent neural networks for detecting suicide notes. *IEEE Trans. Affect. Comput.* 14, 1 (2023), 153–164. DOI: https://doi.org/10.1109/TAFFC.2021.3057105

[131]   M Semrau, S Evans-Lacko, M Koschorke, L Ashenafi, and G Thornicroft. 2015. Stigma and discrimination related to mental illness in low-and middle-income countries. *Epidemiol. Psych. Sci.* 24, 5 (2015), 382–394.

[132]   Faisal Muhammad Shah, Farzad Ahmed, Sajib Kumar Saha Joy, Sifat Ahmed, Samir Sadek, Rimon Shil, and Md. Hasanul Kabir. 2020. Early depression detection from social network using deep learning techniques. In *Proceedings of the IEEE Region 10 Symposium (TENSYMP'20).* IEEE, 823–826. DOI: https://doi.org/10.1109/TENSYMP50017.2020.9231008

[133]   Adrian B. R. Shatte, Delyse M. Hutchinson, Matthew Fuller-Tyszkiewicz, and Samantha J. Teague. 2020. Social media markers to identify fathers at risk of postpartum depression: A machine learning approach. *Cyberpsychol., Behav. Soc. Netw.* 23, 9 (Sep. 2020), 611–618. DOI: https://doi.org/10.1089/cyber.2019.0746

[134]   Anu Shrestha, Edoardo Serra, and Francesca Spezzano. 2020. Multi-modal social and psycho-linguistic embedding via recurrent neural networks to identify depressed users in online forums. *Netw. Model. Anal. Health Inform. Bioinform.* 9, 1 (2020), 1–11.

[135] Ruba Skaik and Diana Inkpen. 2020. Using social media for mental health surveillance: A review. *ACM Comput. Surveys* 53, 6 (2020), 1–31.

[136] Robert J. Smith, Patrick Crutchley, H. Andrew Schwartz, Lyle Ungar, Frances Shofer, Kevin A. Padrez, and Raina M. Merchant. 2017. Variations in Facebook posting patterns across validated patient health conditions: A prospective cohort study. *J. Med. Internet Res.* 19, 1 (2017), e7.

[137] Anke R. Sonnenschein, Stefan G. Hofmann, Tobias Ziegelmayer, and Wolfgang Lutz. 2018. Linguistic analysis of patients with mood and anxiety disorders during cognitive behavioral therapy. *Cogn. Behav. Ther.* 47, 4 (2018), 315–327.

[138] Lauren Southwick, Sharath Chandra Guntuku, Elissa V. Klinger, Arthur Pelullo, Haley McCalpin, and Raina M. Merchant. 2021. The role of digital health technologies in COVID-19 surveillance and recovery: A specific case of long haulers. *Int. Rev. Psych.* 33, 4 (2021), 412–423.

[139] Maxim Stankevich, Ivan Smirnov, Natalia Kiselnikova, and Anastasia Ushakova. 2019. Depression detection from social media profiles. In *Proceedings of the International Conference on Data Analytics and Management in Data Intensive Domains.* 181–194.

[140] Stefan Stieglitz and Linh Dang-Xuan. 2013. Social media and political communication: A social media analytics framework. *Soc. Netw. Anal. Min.* 3, 4 (2013), 1277–1291.

[141] Daniel C. Stokes, Anietie Andy, Sharath Chandra Guntuku, Lyle H. Ungar, and Raina M. Merchant. 2020. Public priorities and concerns regarding COVID-19 in an online discussion forum: longitudinal topic modeling. *J. Gen. Internal Med.* 35, 7 (2020), 2244–2247.

[142] Amanda L. Stuart, Julie A. Pasco, Felice N. Jacka, Sharon L. Brennan, Michael Berk, and Lana J. Williams. 2014. Comparison of self-report and structured clinical interview in the identification of depression. *Comp. Psych.* 55, 4 (2014), 866–869.

[143] Yue Su, Jia Xue, Xiaoqian Liu, Peijing Wu, Junxiang Chen, Chen Chen, Tianli Liu, Weigang Gong, and Tingshao Zhu. 2020. Examining the impact of COVID-19 lockdown in Wuhan and Lombardy: A psycholinguistic analysis on Weibo and Twitter. *Int. J. Environ. Res. Public Health* 17, 12 (2020), 4552.

[144] Jessie Sun, H. Andrew Schwartz, Youngseo Son, Margaret L. Kern, and Simine Vazire. 2020. The language of well-being: Tracking fluctuations in emotion experience through everyday speech. *J. Personal. Soc. Psychol.* 118, 2 (2020), 364.

[145] Michael M. Tadesse, Hongfei Lin, Bo Xu, and Liang Yang. 2019. Detection of depression-related posts in reddit social media forum. *IEEE Access* 7 (2019), 44883–44893. DOI:https://doi.org/10.1109/ACCESS.2019.2909180

[146] Michael Mesfin Tadesse, Hongfei Lin, Bo Xu, and Liang Yang. 2020. Detection of suicide ideation in social media forums using deep learning. *Algorithms* 13, 1 (2020), 7.

[147] Mohd Tajuddin, M. Kabeer, and Mohammed Misbahuddin. 2020. Analysis of social media for psychological stress detection using ontologies. In *Proceedings of the 4th International Conference on Inventive Systems and Control (ICISC'20).* 181–185.

[148] M. L. Tlachac and Elke Rundensteiner. 2020. Screening for depression with retrospectively harvested private versus public text. *IEEE J. Biomed. Health Inform.* 24, 11 (2020), 3326–3332.

[149] Zahra Toosinezhad, Mohamad Mohamadpoor, and Hadi Tabatabaee Malazi. 2019. Dynamic windowing mechanism to combine sentiment and N-gram analysis in detecting events from social media. *Knowl. Info. Syst.* 60, 1 (2019), 179–196.

[150] Tuan Tran, Dong Nguyen, Anh Nguyen, and Erik Golen. 2018. Sentiment analysis of marijuana content via Facebook emoji-based reactions. In *Proceedings of the IEEE International Conference on Communications (ICC'18).* IEEE, 1–6. DOI:https://doi.org/10.1109/ICC.2018.8422104

[151] Alina Trifan, Dave Semeraro, Justin Drake, Radek Bukowski, and Jose Luis Oliveira. 2020. Social media mining for postpartum depression prediction. *Studies Health Technol. Inform.* 270 (2020), 1391–1392.

[152] Marcel Trotzek, Sven Koitka, and Christoph M. Friedrich. 2018. Utilizing neural networks and linguistic meta-data for early detection of depression indications in text sequences. *IEEE Trans. Knowl. Data Eng.* 32, 3 (2018), 588–601.

[153] Danny Valdez, Marijn Ten Thij, Krishna Bathina, Lauren A. Rutter, Johan Bollen, et al. 2020. Social media insights into U.S. mental health during the COVID-19 pandemic: Longitudinal analysis of Twitter data. *J. Med. Internet Res.* 22, 12 (2020), e21418.

[154] R. Vanlalawmpuia and Mr Lalhmingliana. 2020. Prediction of depression in social network sites using data mining. In *Proceedings of the 4th International Conference on Intelligent Computing and Control Systems (ICICCS'20).* IEEE, 489–495. DOI:https://doi.org/10.1109/ICICCS48265.2020.9120899

[155] D. Vivek and P. Balasubramanie. 2020. An ensemble learning model to predict mental depression disorder using Tweets. *J. Med. Imag. Health Inform.* 10, 1 (Jan. 2020), 143–151. DOI:https://doi.org/10.1166/jmihi.2020.2871

[156] Wenyu Wan, Jiumo Sun, Jiehan Liu, Shu-Wen Yang, Mingming Liu, Jia Xue, Dongdong Jiao, and Xiaoqian Liu. 2019. Using social media to explore the linguistic features in female adults with childhood sexual abuse by Linguistic Inquiry and Word Count. *Hum. Behav. Emerg. Technol.* 1, 3 (July 2019), 181–189. DOI : https://doi.org/10.1002/hbe2.162

[157] Wenxi Wang, Huansheng Ning, Feifei Shi, Sahraoui Dhelim, Weishan Zhang, and Liming Chen. 2022. A survey of hybrid human-artificial intelligence for social computing. *IEEE Trans. Hum.-Mach. Syst.* 52, 3 (2022), 468–480. DOI : 10.1109/THMS.2021.3131683

[158] Xinyu Wang, Chunhong Zhang, Yang Ji, Li Sun, Leijia Wu, and Zhana Bao. 2013. A depression detection model based on sentiment analysis in micro-blog social network. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*. 201–213.

[159] Xin Wang, Huijun Zhang, Lei Cao, and Ling Feng. 2020. Leverage social media for personalized stress detection. In *Proceedings of the 28th ACM International Conference on Multimedia*. 2710–2718.

[160] Zheng Wang, Guang Yu, Xianyun Tian, Jingyun Tang, and Xiangbin Yan. 2018. A study of users with suicidal ideation on Sina Weibo. *Telemed. E-Health* 24, 9 (2018), 702–709.

[161] J. T. Wolohan, Misato Hiraga, Atreyee Mukherjee, Zeeshan Ali Sayyed, and Matthew Millard. 2018. Detecting linguistic traces of depression in topic-restricted text: Attending to self-stigmatized depression with NLP. In *Proceedings of the 1st International Workshop on Language Cognition and Computational Models*. 11–21.

[162] World Health Organization. 2019. Suicide Key Facts. Retrieved from https://www.who.int/news-room/fact-sheets/detail/suicide.

[163] World Health Organization. 2023. Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression.

[164] Jheng-Long Wu, Yuanye He, Liang-Chih Yu, and K. Robert Lai. 2020. Identifying emotion labels from psychiatric social texts using a bi-directional LSTM-CNN model. *IEEE Access* 8 (2020), 66638–66646.

[165] Mengqi Xing, Jacklynn M. Fitzgerald, and Heide Klumpp. 2020. Classification of social anxiety disorder with support vector machine analysis using neural correlates of social signals of threat. *Front. Psych.* 11 (2020), 144.

[166] Guixian Xu, Yueting Meng, Xiaoyu Qiu, Ziheng Yu, and Xu Wu. 2019. Sentiment analysis of comment texts based on BiLSTM. *IEEE Access* 7 (2019), 51522–51532. DOI : https://doi.org/10.1109/ACCESS.2019.2909919

[167] Zhentao Xu, Verónica Pérez-Rosas, and Rada Mihalcea. 2020. Inferring social media users' mental health status from multimodal information. In *Proceedings of the 12th Language Resources and Evaluation Conference*. 6292–6299.

[168] Yang Yang, Keqiao Liu, Siqi Li, and Man Shu. 2020. Social media activities, emotion regulation strategies, and their interactions on people's mental health in COVID-19 pandemic. *Int. J. Environ. Res. Public Health* 17, 23 (Dec. 2020), 8931. DOI : https://doi.org/10.3390/ijerph17238931

[169] Shaoyu Ye, Kevin K. W. Ho, and Andre Zerbe. 2021. The effects of social media usage on loneliness and well-being: Analysing friendship connections of Facebook, Twitter, and Instagram. *Info. Discov. Deliv.* 49, 2 (2021), 136–150.

[170] Hui Yin, Shuiqiao Yang, and Jianxin Li. 2020. Detecting topic and sentiment dynamics due to Covid-19 pandemic using social media. In *Proceedings of the International Conference on Advanced Data Mining and Applications*. 610–623.

[171] Senqi Zhang, Li Sun, Daiwei Zhang, Pin Li, Yue Liu, Ajay Anand, Zidian Xie, and Dongmei Li. 2022. The COVID-19 pandemic and mental health concerns on Twitter in the United States. *Health Data Sci.* 2022 (2022). Retrieved from https://pubmed.ncbi.nlm.nih.gov/36408202/.

[172] Weina Zhang, Lu Liu, Qijin Cheng, Yan Chen, Dong Xu, and Wenjie Gong. 2020. The relationship between images posted by new mothers on WeChat moments and postpartum depression: Cohort study. *J. Med. Internet Res.* 22, 11 (Nov. 2020), e23575. DOI : https://doi.org/10.2196/23575

[173] Yipeng Zhang, Hanjia Lyu, Yubao Liu, Xiyang Zhang, Yu Wang, and Jiebo Luo. 2020. Monitoring depression trend on Twitter during the COVID-19 pandemic. Retrieved from https://arXiv:2007.00228.