



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

Dhelim, S., Chen, L., Das, S. K., Ning, H., Nugent, C., Leavey, G., Pesch, D., Bantry-White, E., & Burns, D. (2023). Detecting Mental Distresses Using Social Behavior Analysis in the Context of COVID-19: A Survey. *ACM Computing Surveys*. <https://doi.org/10.1145/3589784>

[Link to publication record in Ulster University Research Portal](#)

Published in:
ACM Computing Surveys

Publication Status:
Published online: 20/03/2023

DOI:
<https://doi.org/10.1145/3589784>

Document Version
Author Accepted version

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

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

Sahraoui Dhelim¹, Liming Luke Chen², Huansheng Ning¹, Sajal K Das³, Chris Nugent², Devin Burns⁴, Gerard Leavey⁵, Dirk Pesch⁶ and Eleanor Bantry-White⁷.

¹School of Computer Science, University College Dublin, Ireland.

²School of Computing, Ulster University, UK.

³Department of Computer Science, Missouri University of Science and Technology, USA.

⁴Department of Psychological Science, Missouri University of Science and Technology, USA.

⁵Bamford Centre for Mental Health & Wellbeing, Ulster University, UK.

⁶School of Computer Science and IT, University College Cork, Ireland

⁷School of Applied Social Studies, University College Cork, Ireland.

ABSTRACT

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 have been increasingly used to express opinions, views and moods due to the restrictions to physical activities, leading to a significant amount of diverse user-generated social media content. This offers an unique opportunity to examine how COVID-19 changed global behaviours in terms of its ramifications on mental wellbeing. In this paper, we surveyed the literature of social media analysis for detection of mental distress, with a special emphasis being placed on the studies published since the COVID-19 outbreak. By analyzing relevant research and its characteristics, we propose new approaches to organizing the huge amount of studies arising from this emerging research area, thus drawing new views, insights and knowledge for related communities. Specifically, we first classify the studies in terms of feature extraction types, language usage patterns, aesthetic preferences and online behaviors. We then explored detection methods used for detecting mental health problems, including machine learning and deep learning detection methods. Built upon the in-depth review, we present our findings and discuss the future research directions and niche areas in detecting mental health problems using social media data. We also elaborate the challenges of this fast-growing research area, including technical scaling and deploying such systems at scale as well as privacy and ethical concerns.

CCS CONCEPTS • Applied computing → Law, social and behavioral sciences → Psychology • Computing methodologies → Artificial intelligence → Natural language processing • Computing methodologies → Machine learning → Machine learning algorithms.

Additional Keywords and Phrases: Social media analysis, mental disorder detection, COVID-19, mental health.

1. INTRODUCTION

Mental illness is a major contributor to the global burden of diseases [46]. The economic cost of mental disorders treatment was estimated to be US \$2.5 trillion in 2010, and is expected to double by 2030 [17]. 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 wellbeing, including increasing capacity for population health diagnosis [124]. 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 misuse appeared to have increased in the general population. Grief, loneliness and isolation will play a role in all of these conditions [103]. 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 [47]. 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 [34]. 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 connection to vulnerable people.

Online social networks are well established as a data source for public opinion mining [137], business analytics [42], events detection [36] and population health monitoring [32]; and are increasingly being used for mental health applications, at both population-level and individual-level health. Social media analysis is especially promising for mental healthcare, as Online Social Networks (OSN) 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 wellbeing. 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 information source for studying mental illnesses, such as depression and loneliness. In this paper, we survey the existing research literature on social media analysis for mental distress detection, with a special focus on the works in the context of COVID-19 pandemic.

During and in the aftermath of COVID-19, the number of papers that focused on social media analysis for mental disorders detection has increased sharply, including several survey and review articles. For example, authors in [132] reviewed the literature on social media analysis for depression and suicide detection, but they limited the review only for text-based social media platforms such as Twitter, Reddit, and Weibo. In [68], 15 papers were summarized to discuss the depressed detection from social media data prior to August 20, 2019. In [79], the authors analyzed the potential causes of suicidal ideation using text-based social media data. Similarly, the authors in [63] 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 the COVID-19 pandemic, numerous studies have appeared to address concerns for mental wellbeing and many of them have examined social media analysis on users' mental well-being related to COVID-19 pandemic. These works have not been fully covered in previous reviews. In addition, most of the recent reviews ([132][79][63]) focus only on text-based social media. The survey presented in this paper contributes to the following four aspects. In terms of the scope of covered mental wellbeing, it covers more mental health conditions, e.g., loneliness, anxiety, stress and post-traumatic stress disorder (PTSD). In terms of the scope of covered studies, it includes most recent research based on visual-oriented social media analysis such as Instagram, and it is the first survey to the best of our knowledge to cover the most recent research and findings regarding the COVID-19 impact on mental wellbeing. Most importantly, we provide new approaches to organizing the literature in terms of feature classifications and detection methods, thus offering new insights and visions to related research communities. Last but not the least, we draw observations, derive findings and point out research gaps, directions and challenges which are most valuable to guide researchers and practitioners for their future research.

Table 1 Recent reviews and surveys on the topic (2020-2021)

Reference	Scope	Limitation	Difference with current work
[132]	Covers only text-based social media Focuses on depression and suicide ideation detection.	Text-only	Current survey focuses more on COVID-19 related works

[68]	Covers 15 works prior to August 20 2019 Reviews only depression detection.	Not COVID-related	Current survey covers text-based as well as visual-based social media analysis In addition to depression and suicide ideation, the current survey also covers loneliness, anxiety, stress and post-traumatic stress disorder (PTSD), as well as other mental disorders.
[79]	Covers suicidal ideation causes analyses from text-based social media data	Text-only	
[63]	Reviews suicidal ideation detection methods using clinical data, also covers few works of text-based social networks data	Not COVID-related	
[22]	Reviews papers published between 2013 and 2018		
[20]	Reviewed 16 suicide detection techniques using social media data	Text-only	

The rest of the paper is organized as follows: In Section 2, we define the scope of the survey and the methodologies considered. Section 3 presents briefly the widely used OSNs and their basic functionalities as preliminaries. In Section 4, we describe features, their discrimination capabilities and feature extraction methods of social media content. Section 5 presents mental distress detection techniques. Section 6 reviews the recent mental distress detection literature works and classifies these works according to their feature extraction and detection techniques. In Section 7, we summarize our finding and outline possible future directions. Section 8 presents the open issues and challenges of mental distress detection from social media. Finally, Section 9 concludes the survey.

2. SURVEY SCOPE AND METHODOLOGY

The focus of the current review is mental disorders, also known as psychological disorders; these two terms are used interchangeably throughout the rest of the paper. 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 mental disorders detection. We focus on works published between 2020, 2021 and 2022, with a special focus on the works

conducted in the context of COVID-19 pandemic. However, we also cover recent key works, such as highly influential works in the field of mental disorders detection using social media analysis.

We used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework [94] guidelines 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 research terms: “mental disorder”, “psychological disorder”, “mental distress”, “psychological distress”, “social media”, “social networks”, “depression”, “covid-19”, “stress”, “anxiety”, “loneliness”, “PTSD”, “suicide”, “schizophrenia”.

The searches were limited to articles written in English. 955 additional articles were identified as related works in the field published between January 2014 and December 2019, these articles were found by following the citations map of the searched articles. After removing duplicated papers, a total of 1516 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, and based on wearables and smartphones passive sensors, or they focus on non-mental healthcare, such as diabetes, obesity or outbreaks predictions. 416 articles were excluded in the eligibility phase after full-text reading. Finally, 120 articles were qualified for final inclusion.

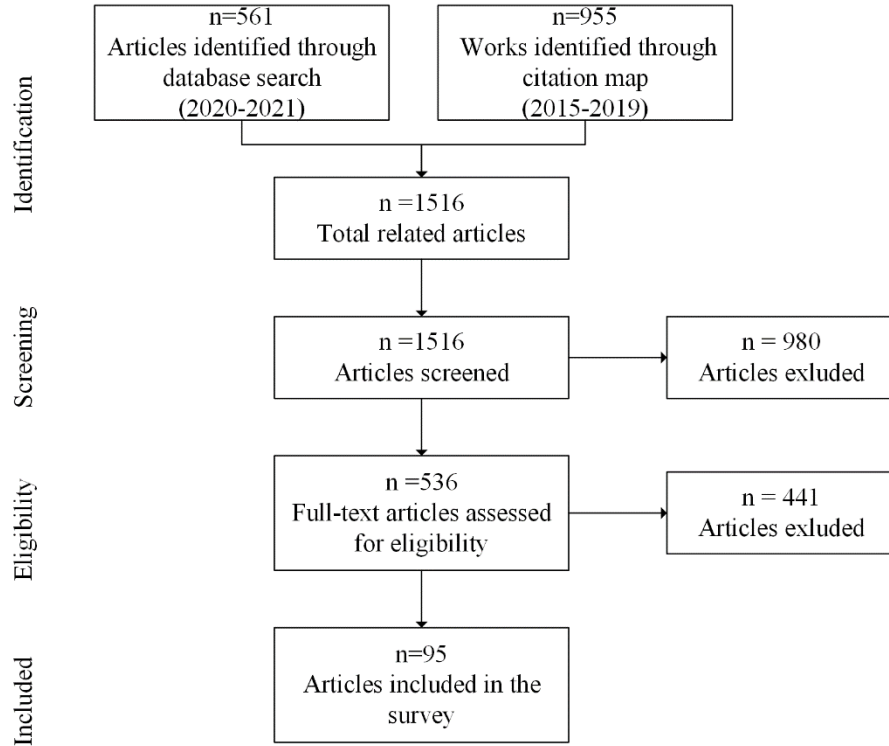


Figure 1 PRISMA flowchart of the review phases

3. ONLINE SOCIAL NETWORKS (OSN)

The popularity of OSN had increased sharply during the last few years. Following the travel restrictions and lockdown measures imposed during COVID-19 pandemic, people have increasingly relied on social media. OSN differ in many aspects, some are text-based such as Twitter and Sina Weibo, while other are photo-based such as Instagram and Flickr. The data extracted from these OSN differ in type and features. Table 2 present the most used OSN for mental disorder detection literature. In this section, we introduce the most used OSN in the literature of mental disorder detection, and their data extraction methods.

Table 2 Most used OSN for mental disorder detection

OSN	Active users in millions (2021)	Main functionality
Twitter	353	Microblogging
Reddit	430	Forum and blogging

Sina	511	Microblogging
Weibo		
Facebook	2740	Social networking
Instagram	1221	Photo and video-sharing

3.1. Twitter

Despite its relatively low number of active users (353 million active users by 2021) compared to other social networks, Twitter is by far the most popular OSN in the researcher 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 data source. Tweets are publicly accessible and can be extracted and analyzed. In addition to the tweet's text body, tweets also include user metadata, such as user's geographical information and the tweet's date and time, and user networking information. 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 using keywords, hashtags, or other queries, the search can be limited to specific locations or time intervals. In the context of mental distress detection, roughly there are three ways to find and extract the data of mentally distressed users. Some works recruit participants for the experiment and directly extract their Twitter data ([26]), and analyze the extracted data and compare it with their self-assisted 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 [52], 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 [168][31]. Another way is to search for a specific hashtag to identify tweets related to a particular mental illness, such as #depression or #MyDepressionLooksLike [73].

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 430M active monthly users by 2021. Reddit users can participate in a community by sending text- or media-based posts, and by commenting on other posts or comments. 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 users share their daily life, express their opinions and feelings, and some may offer mental support to others, or COVID-19 related topics [135]. By March 2021, /r/depression subreddit had 736K members, r/SuicideWatch had 236k members and r/Anxiety had 437k 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 concentrate in one place, unlike twitter data that require extensive searching to identify such mentally distressed users. Figure 2 shows a recent post of Reddit user asking for help in r/SuicideWatch subreddit.



Figure 2 A Reddit user asking for help in r/SuicideWatch subreddit, posted 3 March 2021

3.3. Sina Weibo

Sina Weibo is the largest microblogging website in China (simply known as Weibo), its main functionality is very similar to Twitter. Many studies in the literature of mental distress detection were conducted using Weibo datasets. With more than 511 million active users that post about their daily life struggle, Weibo offers a rich source of user mental health information. Various works have collected datasets from Weibo, and these datasets are used for various mental illness detection tasks. Wang et al. [156] extracted the published data of 1M users with more than 390M posts, and employed a keyword-based method to mark users at suicide risk. Following that, three mental health experts manually labeled users at suicide risk. They could identify 114 users with suicide ideation and leveraged linguistic analysis to explore behavioral and demographic characteristics.

3.4. Facebook

Facebook by far is the largest social network in terms of the number of active users. Facebook offers the possibility to create and share text posts, as well as photo and videos, users can also join various groups and follow Facebook pages. Unlike microblogging OSN such as Twitter and Weibo, Facebook textual posts are not limited by characters limit (tweets have 280 character limit). In the context of mental distress detection, user-generated data can be used for various purposes, specifically, we can extract textual features, visual features and also behavioral features, see Section 4 for more about the different type 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 need 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 photo can be represented by various characteristics: Are there people in the photos? Is the photo setting in outdoor or indoors? The photo was taken night or day? photo 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 received any comments? How many 'Likes' did the photo received? Finally, behavioral features, such as usage and posting frequency, may also give some hints about the user's mental state.

4. MENTAL FEATURES EXTRACTION

Social network data contains various features associated with mental distress and could reveal the psychological state of the users. These features are either extracted from a single modalities (e.g., text, images or audio), or fused from a variety of data sources (multimodal) [100]. Multimodal feature processing generally achieves better mental distress detection accuracy compared to single modal schemes, as individuals tend to manifest their inner psychological states using different expression mediums [10]. Hence, multimodal features can alleviate the generalization error caused by individual differences, such as personality type, ethnicity, age and gender. There is a plethora of social networks. Users share their daily life activities; express their opinions and feelings on different social networking platforms with services using different data formats. For instance, text-based microblogging social networks, e.g. Twitter and Weibo, are mainly used for

posting short textual messages. While, visual-based social networks, e.g. Facebook and Instagram, are primarily used for sharing photos and videos. Different features are extracted from social network data depending on the type of shared content. We proposed the taxonomy presented in Figure 5 to classify the previous works on mental distress detection using social network data.

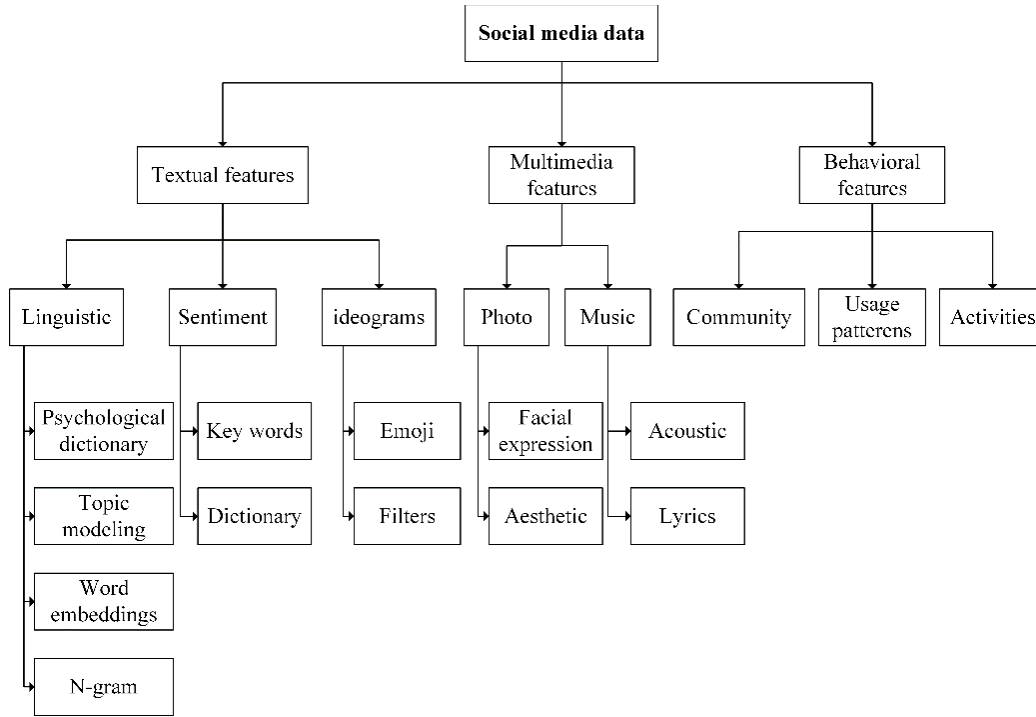


Figure 3 Psychological features classification

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 3, textual features can be divided into three classes, namely linguistic, sentiment, and ideogram features. Multimodal feature fusing can be applied within the same textual features class by combining various features from the same class (e.g. linguistic-linguistic [142]), or by combining features from various feature classes (e.g. linguistic-sentiment [115]).

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) [108]. Which is widely used text analysis technique in psychology, and it 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. In the past 30 years, LIWC had been used in different research 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 [56], emotions and LIWC emotion-related categories [141], LIWC positive emotions category word usage and anxiety [134]. In the context of mental distress detection using social media analysis, LIWC is by far the most used psychological dictionary-based feature extraction method ([80][77][57][158][91] [153]). Other works used other psychological dictionaries to extract linguistic features, such as Affective Norms for English Words (ANEW) [35]. Another effective linguistic feature extraction method is topic modeling, which is the process of extracting the 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, hidden topics such as topics connected with psychological distress can be extracted from the text corpus.

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-labeled 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) [16]. In LDA, a topic is represented as a multinomial distribution associated to 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. Most of the work leverage the original unsupervised LDA algorithm for topic extraction ([92][45][118] [150][169]), some works reported better results using supervised or semi-supervised LDA, that guide the topic modeling process by specifying mental-related lexicon of which are likely to appear in the generated content of distressed individuals ([7][138]).

Another widely used linguistic features for mental distress detection are the word frequency features that represent the original text corpus into representational language model. The most basic models in this category is the Bag of Words (BoW) model [171] and term frequency–inverse document frequency (TF-IDF) [161]. 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. While TF-IDF is a term-weighting scheme that measure 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 with that we lost important information about the temporal dimension of the user psychological state. The N-gram model [146] solves this problem, 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 model are used as linguistic features that represent the probability of co-occurrence of each input sentence as a unigram and bigram.

Word embedding is a deep learning technique that is used for text representation for the recognition of comparatively important words. This modeling 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 embeddings techniques is Global Vector for word representation (GloVe) [121]. Glove searches for similar words in the whole context of user-generated text. Another widely used word embedding is word2vec [29]. Generally, linguistic features fusing from more than class can increase the mental distress detection accuracy compared to single class linguistic features, e.g. LIWC+N-gram [157], LIWC+LDA+N-gram [142][72].

4.1.2. *Sentiment features*

Sentiment analysis is concerned with extracting emotions, opinions, affects and from user-generated texts on social media, it is crucial that knowledge from this area can also be very useful to find the psychological state of the user. Especially the sentiments that users describe towards their personal situation could be an important marker. Besides the linguistic features described above, the user-generated text on social media could contain semantic features related to mental distress that are difficult to extract using linguistic features extraction, 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 sentimental features is the key-word search,

in which the system looks for specific pre-defined key-words in the users' posts, e.g. 'depressed', 'feeling lonely' or 'happy' [21] [166]. The drawback of key-word approach is the lack of context of using these words, as user may use these keywords in a sarcastic way for instance. A more selective method is to compute the frequency of used words that belong to LIWC emotion categories ([97][48]), or LDA topics that represent basic emotions. Another way to measure the importance of emotion-related words is to measure the occurrence frequency of these words using TF-IDF, which is often used in sentiment analysis tasks ([164][3]). In addition to this, the NRC Emotion Lexicon [9] and VADER Sentiment Lexicon [58] are used for user sentiment and emotion recognition .

4.1.3. *Ideogram features*

Most of 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 emoji and stickers might be associated with stress. Emoji usage choices have been associated with depression ([152][88][13]), mental distress [37] and emotions ([147]). One of the main advantages of 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 researches have 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, profile photos and posted photos. Generally speaking, profile photos convey less information about the mental state of the user, as most of the user choose socially attractive profile photos, which is known as self-presentation biases, for instance even depressed users express positive emotions in their profile photos [49]. As we have shown in Figure 3, there are three types of

features that can be extracted from user-generated photos, facial expression features, object class features and aesthetic features. Facial expression features of the studied users, as well as people present in the photo. Facial expression features not only used to identify the user's sentiments but also information about his 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 and psychological state [170]. On the other hand, aesthetic features represent the photo characteristics such as the color and filter choices. For example, Photo might be analyzed at the pixel level averages to extract Hue, Saturation, and Value (HSV), three-color properties widely used in image feature extraction. Hue represents the photo's coloring on the light spectrum (varying from red to purple/blue). Lower hue values reflect in more red, and higher hue values result to photo tending more to blue. Saturation represents the vividness of the photo. Low saturation results the photo to be more grey and faded. Value is the image brightness. Lower brightness scores reflect a darker photo. Previous researches associated depression with the posted images being grayscale and low aesthetic cohesion across a variety of image features, and users suffering from social anxiety also tend to post grayscale and low aesthetic cohesion photos but less than depressed users [49], see Figure 4.



Figure 4 The right photo has higher Hue (bluer), lower Saturation (grayer), and lower Brightness (darker) than left photograph. Instagram photos posted by depressed individuals had HSV values shifted towards those in the right photograph, compared with photos posted by healthy individuals [114].

4.2.2. Audio features

Various features can be extracted from music listening activities on 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, timbral texture, pitch, spectral features. Lyrics features are extracted by applying textual feature extraction techniques 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 live activities and their interests.

4.3.1. Community affiliation

Social networks offer the user the ability to join social network groups and communities (e.g. Facebook groups, Reddit communities). Analyzing the type of groups and communities that the user has joined might help to detect his topics of interest. Mental distress-related groups and communities are specifically helpful to identify users with mental health issues, many previous studies have shown that topics and psycholinguistic features were found to be accurate predictors of mental disorders like depression [118], bipolar [69] and suicide ideation [126].

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. For instance, Smith et al. [133] found that the sudden increases of Facebook postings activity is positively correlated with depression symptoms. Lup et al. [83] found that there was a marginal positive correlation between Instagram hourly usage time and depressive symptoms

4.3.3. Online activities

Besides the usage time and community affiliations, the interaction activities with posted content, such as likes, comments and follows, might be correlated with certain psychological markers, for example, Negriff [99] found that depressive symptoms is negatively correlated with the number of Facebook friends. Lup et al. [83] investigated the amount 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 psychological problem is by analysing keyword searching behaviors [61]. For instance, Chen et al [24] analyzed Baidu search data, and noticed that searches related to psychological problems increased during COVID-19 pandemic, keywords like depressed, sad, panic, fear, insomnia, obsessive-compulsive disorder, psychological counselling.

5. 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 the various learning models used in the literature of social media mental distress detection. In Figure 5, we present the classification of the detection methods.

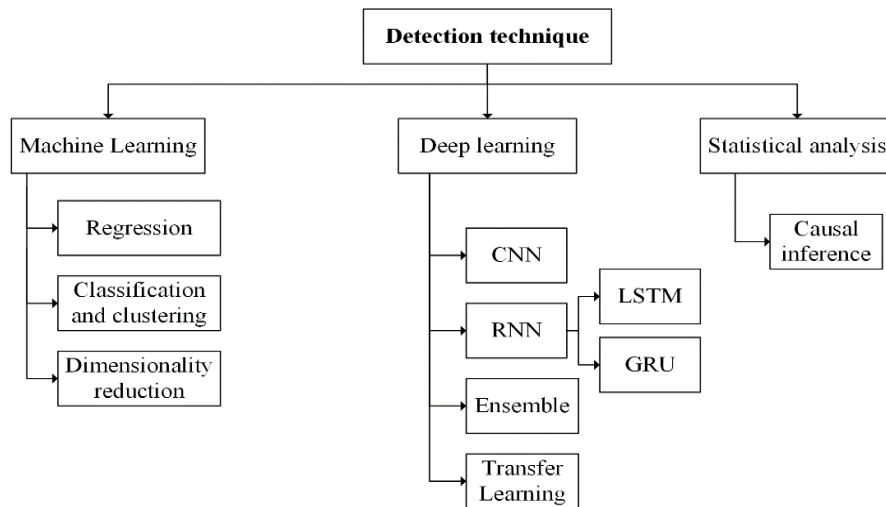


Figure 5 Mental distress detection classification

5.1. Machine learning detection

Machine learning detection schemes are by far the most used techniques in the literature of mental distress detection.

5.1.1. *Classification and clustering*

In machine learning, classification is the problem of finding to which of a group/class/category a new observation belongs, based on existing knowledge of a training set of data that includes instances whose class affiliation is already known. For example, given three sets of users grouped based on their depression level, low-group, medium-group and high-group, 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). Support-vector machines (SVM) is one of the most effective machine learning classifiers, SVM has been proven to perform well with short informal text, with promising results when applied to mental health classification tasks [26]. SVM is widely used in mental disorders detection using social networks data, including: depression detection ([142][107][26]), social anxiety ([23][38][163]), suicide risk assessment ([143][25][84]). Other widely used classifiers for mental health assessments include K-Nearest Neighbor (KNN) [59] and Naïve Bayes (NB) classifier [18].

5.1.2. *Regression methods*

Regression is a commonly used statistical procedure of iteratively refining a measure of the error in the prediction made by the model to find the relationship between variables. It can be used to predict the risk of mental distress based on extracted social media features of the user, specifically, logistic regression model can calculate the probability of a categorical variable (depressed / non-depressed) from a set of predicted features. Linear regression is also used in various mental disorders detection methods. In linear regression, the relationships between the mental disorder and user data are modeled using linear predictor functions whose unknown model parameters are estimated from the user-generated social media data. For example, in [115] to construct linear regression depression predictive model, linguistic features were obtained from Instagram comments and post captions and, including emoji sentiment analysis results, multiple sentiment scores and meta-features such as likes count and average comment length.

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 in order 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 later for supervised learning method. Many of 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 disorders assessment using social media data, such as the representation of social data of depressed users [136] and users with schizophrenia [90] and suicidal users [75].

5.2. Deep learning detection

In the last few years, Deep-learning methods have become the mainstream detection techniques for mental distress from social networking data. They are 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 image, music streaming audio, and video.

5.2.1. Convolutional Neural Network

Convolutional neural network (CNN) is a type of deep neural networks, mostly applied to analyzing images. CNNs are special case of multilayer perceptrons (MLP). Multilayer perceptrons typically mean fully connected networks, where every neuron in one layer is connected to all other neurons in the next layer. The "total-connectedness" of MLP makes them prone to overfitting data. This problem is typically solved by regularizing these networks by adding magnitude measurement of weights to the loss function. CNN uses 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 in image classification and they generally yield good accuracy for data with a grid-like structure. However, many studies have proven that CNNs can also 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 [149], or more complex multilayer modeled as a sequence of layers that includes an embedding layer, convolutional layer, dense layers, max-pooling layer and the output [69].

5.2.2. Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are a class of artificial neural networks in which connections between nodes is in the form of a directed graph along a temporal sequence. Which allows it 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, connected series of data. The most famous variants of RNNs is 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 backward 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) is the dominant deep-learning model in the literature of social media data analysis for mental distress assessment. Some of works opted to use only LSTM or its variants for mental disorders assessment ([62][131][2][55]), while others combine more than one deep learning model such as LSTM+CNN ([162][143]) or GRU+CNN [116].

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 second different but related task. One of the widely applied transfer learning technique 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 been already pre-trained on a large dataset. In the context of mental health analysis for mental disorder detection, pre-trained classifiers that were trained on general purpose NLP task can be leveraged to identify the language indicative of mental disorders. For instance, Saha et al. [119] used a transfer learning classifier to identify mental health related language on twitter (anxiety, depression, stress, and suicidal ideation). Sharma et al [129] used deep reinforcement learning to transform low-empathy conversational posts to higher empathy.

5.3. Statistical analysis

Besides mental distress detection, some other works have focused on the causal relationship between social media usage activities, physical and mental wellbeing [51]. The correlation analysis allows us to capture the relationships between extracted feature patterns and abnormal

habits related to psychological wellbeing. 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 wellbeing. For example, correlation analyses were performed in [106] to identify depressive symptom-related features from users of Facebook and Center for Epidemiological Studies-Depression (CES-D) scale scores. De Choudhury et al. analyzed social media data related to suicide ideation risk [27,28,128].

6. SOCIAL MEDIA MINING FOR MENTAL HEALTHCARE

In this section, we will review the recent published works on the topic, with a special focus on works that have been conducted in the context of 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 [160]. Symptoms of depression include feeling low self-esteem, hopeless or unhappy, and finding no pleasure in activities you 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 could cause a serious health condition, especially with long-lasting moderate or severe intensity. In such case, depression might influence the subject to suffer from acute low mood and being less productive at work or school, or even leads to suicide [40]. Moreover, lockdown measures cause by the spread of COVID-19 has rubbed salt in a wound. Following the economic crises and job lost, add to that the fear of virus that leads the people not to seek mental assessment even when they genuinely needed it, not to mention the feelings of worry, fear, and stress, at individual as well as at community level. The numbers of depressed people had increased sharply since the emerging of COVID-19 [1]. Although there are various effective treatments for depression and other mental disorders, the majority of people in middle-income and low-income countries have no access to mental healthcare. Challenges of effective mental healthcare include a lack of resources, low training quality of healthcare providers and social stigma related to mental disorders. Another challenge to effective mantel healthcare is inaccurate assessment. Even in high-income countries, depressed people usually are not accurately diagnosed, and others with no depression are often misdiagnosed and prescribed antidepressant medication.

6.1.1. *Depression assessment*

There many depression assessment methods, clinical interviews administrated by practicing clinicians or psychologists are the most reliable method of depression assessment; however, it is extremely difficult to conduct such interviews on a large population [139]. Psychometric self-report questionnaires for depression are generally considered as valid and reliable assessment method. Some of the most commonly used depression self-report surveys are the Patient Health Questionnaire (PHQ-9) [71], CES-D [111] and Beck Depression Inventory (BDI) [11]. To the best of our knowledge, only one social media based depression detection study has used clinical interviews to measure the detection accuracy [155]. 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 ([26][113][114]). Alternatively, most of the works used less certain ground truth data, such as the user participating in communities about depression in Reddit ([45][81][60][89]) 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” ([31][30][15]), or search 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. Monreno et al. [95] analyzed depressive symptoms from Facebook posts of randomly selected university students. They concluded that there was a significant positive correlation between depressive signs expressed in the subject’s Facebook post and their score on the PHQ-9 depression scale. Similarly, Ophir et al. [104] study 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, Ehrenreich et al. [39] analyzed social media post of adolescent girls, and found that posts contain somatic complaints, negative effect and call for supports are correlated with depression symptoms, while no depressive peers posts were more likely to contain offer of help and support. Shatte et al. [130] have 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 COVID-19 pandemic, Zhang et al. [172] created an English Twitter depression dataset containing 2,575 distinct identified depressed users with their past tweets, and trained three transformer-based depression classification models based on the collected dataset. Linguistic features have been proved to be good markers of depression.

Many studies have used psychological dictionaries to analyze language usage and word choice and its association with depression. Shrestha et al. [131] proposed an unsupervised depression detection algorithm based on RNN, they computed a vector of LIWC features for each post text, and feed it to LSTM autoencoder along with network-based features modeling how users connect in the forum. their results on detecting depressed users show that psycho-linguistic features derived from the user's social media posts are good predictors of the user's depression severity. Mustafa et al. [97] used the 14 psychological attributes in LIWC to classify the post into emotions, and assigned weights to each word from happy to unhappy after LIWC classification and trained machine learning classifiers to distinguish the users into three depression levels, High, Medium, and Low. Vanlalawmpuia et al. [151] developed a depression detection method that combines LIWC with another 39 attribute sets and 252 depression-related words with temporal and linguistic style. Tadesse et al. [142] analyzed Reddit users' posts to detect depression from textual features, and found that combining more than textual feature types can yield very high depression detection accuracy. They have combined LIWC, LDA and bigram features, and feed them to MLP classifier resulting in a performance for depression detection of 91% accuracy and 0.93 F1 scores. Ricard et al. [115] have built a depression predictive model that make 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 user's multimedia features such as profile photo choose, shared photos and videos, and music listening activities. Compared to other psychological distress and mental disorders, depression has a strong correlation with multimedia features. Reece et al. [114] used CES-D scale as ground truth information, and collected 43,950 Instagram photos of 166 users. They applied color analysis, metadata components, and face detection on the collected photos, and investigated the presence of 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 for participants' social activity levels. Pixel-level averages were calculated using HSV values, three color properties widely used in photo analysis. They 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, Mann et al. [87] combined various textual features, LIWC, BoW, FastText embedding and ELMO embedding, along with the color of images as visual features. Specifically, they extracted HSV features by taking the average of the pixels in the image, and also counted the number of faces in each image using deep-learning face detection model, resulting in a total of 12 visual features and 64 textual features. Malhotra et al. [86] proposed a multimodal 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 weighted average score, which is used for making the final prediction using the Softmax prediction layer.

6.1.4. Depression detection using behavioral markers

Online activities can be rich information to study the preferences and behaviors, as well as the differences among users. Smith et al. [133] found that the sudden increases of Facebook postings activity is positively correlated with depression symptoms. Negri [99] found that depressive symptoms is negatively correlated with the number of Facebook friends. Negri also concluded that the connectedness of social network (i.e., number of mutual friends) is positively correlated with less depression signs, and depressed people tend to use the location tag function and the 'like' button less often. Reece et al. [113] found that the average word count per tweet was negatively correlated with depression. Nonetheless, unlike [133], they did not find any significant correlation between depression symptoms and the frequency of tweets. Lup et al. [83] studied the relationship between depression signs and Instagram usage time, and amount of strangers the user follows. They found a marginal positive correlation between Instagram hourly usage time and depressive symptoms, and the number of strangers followed slightly moderated this relationship. In other words, depression signs increased with Instagram usage time if the user followed a high number of strangers, but if the user followed fewer strangers, Instagram usage time and depression signs were unrelated.

Table 3 Recent depression detection schemes

Work	Textual feature	Visual feature	Behavioral feature	Detection technique	Social network
[142]	LIWC+LDA+bigram	N/A	N/A	MLP	Reddit

[131]	LIWC	N/A	Network (follow)	LSTM	ReachOut.co m
[43]	Word embeddings	N/A	N/A	CNN-BiLSTM	Twitter
[97]	LIWC+ TF-IDF	N/A	N/A	Various classifiers	Twitter
[130]	Key-words	N/A	Engagement+ community affiliation	SVM	Reddit
[89]	TF/IDF	N/A	N/A	Logistic Regression	Reddit
[148]	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
[151]	LIWC+keywords	N/A	N/A	Statistical analysis	Facebook
[170]	N/A	Aesthetic	N/A	Logistic regression	WeChat
[87]	LIWC, BoW, FastText, ELMo	HSV	N/A	LSTM	Instagram
[114]	N/A	HSV	N/A		Instagram
[5]	Word embedding (Word2Vec) N-gram (TF/IDF)	N/A	N/A	Various classifiers	Online forums
[78]	Raw text	Raw images	N/A	CNN (image) BERT (text)	Twitter
[86]	word2vec	VGG-16 Faster- RCNN	N/A	Ensemble (word2vec+ VGG16+ Faster- RCNN)	N/A
[145]	Psychological dictionary (Empath+ g Textblo)	N/A	N/A	PCA	Twitter
[66]	Sentiment, emotions, personal pronoun,	N/A	N/A	LSTM	Twitter

	absolutist words, negative words				
[127]	GloveEmbed, Word2VecEmbed, Fastext and LIWC	N/A	N/A	BiLSTM	Reddit

6.2. Suicide ideation detection

Mental disorders are major risk factor of suicide, according to a WHO report, approximately 800,000 people take their lives every year [159]. The US 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 attempted yet, which is a vital step for suicide risk assessment. 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 [33], Adult Suicide Ideation Questionnaire, Suicidal Affect-Behavior-Cognition Scale [54]. 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 a new method for detection suicidal ideation and suicidal cause analysis. People with psychological distress usually do not trust traditional mental health methods and services. Researches 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 [8]. Many researchers have confirmed the effectiveness of suicidal ideation detection by analyzing the user's online activities and the generated social data. Sarsam et al. [122] discussed the importance of emotions in Twitter content in detecting suicide risks. They analyzed the features of Twitter users' emotions and behavior responses (sadness, fear, anger, joy, positive, and negative) using SentiStrength and NRC Affect Intensity Lexicon (NRC-AIL) classification. A semi-supervised learning scheme based on Yet Another Two-Stage Idea (YATSI) classifier to identify suicide-related tweets. Schoene et al. [126] studied the possibility of automatic suicide notes identification from social media posts as document-level classification tasks. Specifically, they have extracted suicide related LIWC features extracted from suicide notes and trained a Dilated LSTM suicide risk detection model. Tadesse et al. [143] evaluated the n-gram analysis to prove that the phrases related to suicidal tendencies and low social engagements are often present in

suicide-related forums. They studied the transition towards the social ideation related with different psychological states such as heightened self-focused attention, frustration, anxiety, hopelessness, or loneliness. Furthermore, they have extracted various textual features such as 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 Recent Suicide ideation detection schemes

Work	Textual feature	Visual feature	Behavioral feature	Detection technique	Social network
[126]	LIWC	N/A	N/A	Dilated LSTM	Reddit
[122]	SentiStrength, NRC-AIL	N/A	N/A	YATSI	Twitter
[143]	N-gram, TF-IDF, BoW	N/A	N/A	LSTM-CNN	Reddit
[19]	BERT	Raw images	Posting time	knowledge graph	Weibo Reddit
[53]	BERT, Sentence-BERT, GUSE	N/A	N/A	DNN	Reddit
[123]	SentenceBERT	N/A	N/A	T-LSTM	Twitter
[82]	LDA, NMF	N/A	N/A	CNN	Twitter
[105]	Deep Contextualized Word Embeddings (CWE)	N/A	N/A	ANN	Facebook
[112]	keywords, VADER sentiment	N/A	N/A	RF	Twitter
[117]	keywords	N/A	N/A	neural network	Twitter
[110]	TF-IDF, BoW	N/A	N/A	Various classifiers	Twitter
[125]	N/A	Image tags	N/A	classifier	Instagram
[85]	Word embeddings	visual representation (ResNet)	N/A	GRU	Weibo

[74]	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 [120]. Many works have leveraged textual features for loneliness detect from social media content. Xuan and Liew [70] analyzed the patterns of loneliness expression in Twitter during 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 apply 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. Guntuku et al. [50] 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. Following that, they 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, Ameko et al. [6] proposed a loneliness detection system named LonelyText, specifically, they applied SVM classifier on LDA topics that were extracted from Facebook dataset. Multimedia features are also useful in the context of loneliness and social anxiety detection. Kim et al. [67] analyzed the relationships between user loneliness and color features of their Instagram photos. They analyzed 25,394 Instagram photos in terms of color diversity, colorfulness, and color harmony. Their results suggest that color diversity is negatively correlated with user loneliness, in particular romantic loneliness. Behavioral features have proven to carry rich information regarding lonely and social anxious users. Ye et al. [167] studied the relationship between user loneliness and the type of social media they use. 155 Japanese university students involved in the study were divided into four groups based on their response 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 usage type on loneliness and well-being for each group were analyzed. They found that, No social media usage has effects on loneliness or mental health were associated for students who used only Twitter or both Twitter/Instagram. Students that use both Twitter/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 use Facebook via computer longer; however, their access time of Facebook using mobile phone helped them decrease loneliness and improve their levels of mental well-being.

6.4. Anxiety detection

A growing number of studies suggesting the number of people with anxiety increased during stay-at-home COVID-19 orders [41]. Chang et al. [23] investigated users suffering from social anxiety disorder using behavioral and social-network topological features. They first collected ground truth data from various social media sources and identify anxious users with the help of mental healthcare professionals. After that, they used multiple features, mainly TF-IDF word scores, negative self-disclosure, sentiment score, parasocial relationship and social event attending. Finally, they applied SVM classifier to identify users with social anxiety disorder. Similarly, Dutta et al. [38] found that when SVM classifier is applied 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, Fauziah et al. [44] have applied random forest and xgboost classifier on YouTube comments for anxiety detection during COVID -19 pandemics. Xu et al. [165] investigated the relationship between various mental disorders and online activities, they have analyzed Flickr dataset by applying multimodal detection technique that uses multimodal features, namely textual features, visual features such as color distribution and the presence of faces and objects, and metadata features and their relation to mental wellbeing. They 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

Su et al. [140] investigated the impact of COVID -19 lockdown measures on individuals' psychological states in Italy and China. They firstly extracted data of Wuhan, China Weibo users using the geo-location filter, and the social media data of Lombardy, Italy Twitter users also using geo-location filter. Following that, they extracted all users' posted content with two weeks before and after the lockdown for each region (Wuhan and Lombardy lockdowns). After that, they

extracted psycholinguistic features of these posts using the Simplified Chinese version and Italian version of LIWC. Finally, they performed various Wilcoxon tests to study the changes in the psycholinguistic features of the posts before and after the lockdown in Lombardy and Wuhan respectively. Their finding suggests that 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, and the focus on group, religion, and emotions became more common in Wuhan after the lockdown. Wang et al. [154] proposed a three-leveled framework for stress detection. The three-leveled framework learns the personalized stress representations following increasingly detailed processing, i.e., from the generic mass level, group level, to the final individual level. The first mass-level is dedicated on mining the generic stress features extracted from people’s linguistic and visual posts using a two-layer attention mechanism, in this layer they used GRU model, which inputs the embedding vector of each word and outputs the hidden representation. The second group-level leverages a Graph Neural Network (GNN) to learn the social media group affiliation features. The third 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 in detection stressed users. Khan et al. [65] applied LDA on non-stressed related tweets from Twitter to detect stress among the users from their tweets and categorized the tweets into two classes stressed and non-stressed user groups. Their results suggested that applying LDA for stress detection yield better performance than SVM-based detection. Tajuddin et al. [144] proposed a hybrid ontology for stress detection that captures the users’ keyword matching search process used in social media to identify the stress-related messages. Meshram et al. [93] applied CNN classifier on linguistic, visual and interaction features extracted from twitter data to identify stressed individuals.

6.6. Other disorder

The consequences of COVID -19 pandemic caused mental distress to previously healthy people, but may also have aggravated and worsened the situation for people with previous mental issues, such as Schizophrenia and PTSD. Li et al. [76] analyzed 19,224 schizophrenia-related Weibo posts and extracted psycholinguistic features of the Simplified Chinese version of LIWC from each post. Following this, they applied SVM, NB and logistic model trees (LMT) to identify schizophrenic users. Similarly, Birnbaum et al. [14] collected 3,404,959 Facebook messages and 142,390 images of 223 participants with schizophrenia spectrum disorders (SSD). 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 have extracted ten features related to hue, nine features related to saturation and value, two measures related to sharpness, one for contrast, in addition to the number of pixels, width, height, and brightness. They applied logistic regression on these features to determine schizophrenic users. Johnson et al. [64] have 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. They have collected tweets dataset of general population and veteran with PTSD; they extracted PTSD related topics by searching specific hashtags such as “#PTSD”, and analyzed tweets posted by 6,000 veterans. Murarka et al. [96] crawled 17159 Reddit posts during 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. OBSERVATIONS, FINDINGS AND FUTURE DIRECTIONS

Social media analysis is an effective mental health assessment medium. However, as we have shown in this work, there are various online social networks (OSN) that contain different types of data used to extract various features and processed with different detection techniques. Twitter and Weibo are the most used OSN, due to the short length of microblogging posts that make them suitable for text processing, as they contain an extract of the user's thoughts without any additional wordiness, unlike Facebook posts for instance. Most of the surveyed works used textual features as markers for psychological distress, as textual features are relatively easy to process and transform to machine-readable format. Also, because the psychology literature provides strong evidence about the correlation between language usage and psychological distress. Unlike visual and behavioral features that require huge data input to capture the user's preferences and relate them to psychological patterns, and the relatively little evidence of the correlation between e.g. color preferences and psychological distress. Machine-learning detection techniques usually yield good performance with a small set of data collected from participants' social media accounts, while deep-learning techniques have the upper hand when dealing with 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 visual features processing, such as detecting the number of faces in Instagram photos as an indicator of social exposure. Few works have used multimodal data from various OSN, however, they used a relatively small dataset, and therefore the benefits of such an approach are not achieved yet. Developing a multimodal framework that

incorporates various user activities is a promising future direction. As different individuals tend to manifest mental status differently in their behaviors, some may manifest their mental status verbally (e.g. linguistic and language usage changes), while other may manifest that aesthetically (e.g. photo colors and filters preferences), and some other may manifest more passive symptoms that can be captured through monitoring their music listening history and online activities.

8. OPEN ISSUES AND CHALLENGES

Leveraging people's social media content as a mental healthcare data source for disorders assessments and intervention confers various benefits (i.e. reduced recall bias, cost efficiency, 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

The subjects' privacy is by far the most challenging issue when dealing with social media data, from a research perspective as well as a deployment perspective. From a research point of view, the studied users' privacy might be affected throughout the data processing stages. Although most of the time the studied datasets are publicly available, problems can arise when users' personal attributes can be predicted, and the identity of 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 an ethical approval or exemption from their Institutional Review Board (IRB) prior to the study. In addition to that, they must obtain informed consent from users when possible, and protect and anonymize sensitive data during all research stages. Moreover, they need to be careful when linking 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 [12]. Researchers can rely on public social media datasets for mental healthcare research if they ensure the preserving of the confidentiality of users. The privacy concern is more challenging when applying such mental healthcare solutions on the existing social media in large-scale level. Most of the OSNs store the users' data on their servers, and already obtained the users' consents to process their data for service enhancement, however, an additional consent to use their data for mental health purposes is still required.

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 inexistent mental distress, and false-negative where the system misses the detection of mentally distressed users. Although many researches have achieved high accuracy for detecting various psychological conditions (e.g. 91% accuracy detection depression [142] and 93.8% accuracy detecting suicide ideation [143], 90% accuracy detection stress [154]), the experiments that yielded these results are limited, and good accuracy is not guaranteed when these technologies are applied in a large-scale population. Besides the technological challenges to increase 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 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 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 approach compared to the conventional mental healthcare, such as non-invasive assessment.

9. CONCLUSION

In this work, 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 scopes of previous surveys to incorporate the latest social media analysis techniques and their application to the monitoring and detection of multiple mental health conditions. We have particularly focused

on relevant studies 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 mental distress prediction models. We have reviewed recent studies and developed a new structure to categorise 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 the privacy and ethical concerns, as well as the technical challenges of scaling and deploying such systems at large scales, and discusses the lessons learned over the last few years.

REFERENCES

- [1] Alison Abbott. 2021. COVID’s mental-health toll: how scientists are tracking a surge in depression. *Nature* (2021).
- [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. Imaging Heal. Informatics* 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] Norah Saleh Alghamdi, Hanan A Hosni Mahmoud, Ajith Abraham, Samar Awadh Alanazi, and Laura García-Hernández. 2020. Predicting depression symptoms in an Arabic psychological forum. *IEEE Access* 8, (2020), 57317–57334.
- [6] Mawulolo K Ameko, Sonia Baee, and Laura E Barnes. 2021. LonelyText: A Short Messaging Based Classification of Loneliness. *arXiv Prepr. arXiv2101.09138* (2021).
- [7] 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.
- [8] 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, and others. 2016. Mental disorders among college students in the World Health Organization world mental health surveys. *Psychol. Med.* 46, 14 (2016), 2955–2970.
- [9] Anil Bandhakavi, Nirmalie Wiratunga, Deepak Padmanabhan, and Stewart Massie. 2017.

- Lexicon based feature extraction for emotion text classification. *Pattern Recognit. Lett.* 93, (2017), 133–142.
- [10] 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.
 - [11] Aaron T Beck, Robert A Steer, and Gregory Brown. 1996. Beck depression inventory--II. *Psychol. Assess.* (1996).
 - [12] Adrian Benton, Glen Coppersmith, and Mark Dredze. 2017. Ethical research protocols for social media health research. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, 94–102.
 - [13] Uffaq Bilal and Farhan Hassan Khan. 2019. An Analysis of Depression Detection Techniques from Online Social Networks. In *International Conference on Intelligent Technologies and Applications*, 296–308.
 - [14] 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 Schizophr.* 6, 1 (December 2020), 38.
DOI:<https://doi.org/10.1038/s41537-020-00125-0>
 - [15] 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.
 - [16] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, Jan (2003), 993–1022.
 - [17] 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, and others. 2012. *The global economic burden of noncommunicable diseases*.
 - [18] 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.
 - [19] Lei Cao, Huijun Zhang, and Ling Feng. 2020. Building and Using Personal Knowledge Graph to Improve Suicidal Ideation Detection on Social Media. *IEEE Trans. Multimed.*

- (2020).
- [20] Gema Castillo-Sánchez, Gonçalo Marques, Enrique Dorrnoro, 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>
 - [21] 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. Human Behav.* 54, (January 2016), 351–357. DOI:<https://doi.org/10.1016/j.chb.2015.08.023>
 - [22] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ Digit. Med.* 3, 1 (2020), 1–11.
 - [23] Ming-Yi Chang and Chih-Ying Tseng. 2020. Detecting Social Anxiety with Online Social Network Data. In *2020 21st IEEE International Conference on Mobile Data Management (MDM)*, 333–336.
 - [24] 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. Disord.* 276, (2020), 1093.
 - [25] 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.
 - [26] 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*.
 - [27] Munmun De Choudhury and Emre Kiciman. 2017. The language of social support in social media and its effect on suicidal ideation risk. In *Eleventh International AAAI Conference on Web and Social Media*.
 - [28] Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2098–2110.
 - [29] Kenneth Ward Church. 2017. Word2Vec. *Nat. Lang. Eng.* 23, 1 (2017), 155–162.

- [30] Glen Coppersmith, Mark Dredze, and Craig Harman. 2014. Quantifying mental health signals in Twitter. In *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality*, 51–60.
- [31] Glen Coppersmith, Craig Harman, and Mark Dredze. 2014. Measuring post traumatic stress disorder in Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*.
- [32] 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.* (2020).
- [33] John G Cull and Wayne S Gill. 1989. Suicide probability scale (SPS). *Los Angeles West. Psychol. Serv.* (1989).
- [34] Emma Curran, Michael Rosato, Janine Cooper, Christine A Mc Garrigle, 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.
- [35] Peter Sheridan Dodds, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth. 2011. Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. *PLoS One* 6, 12 (December 2011), e26752. DOI:<https://doi.org/10.1371/journal.pone.0026752>
- [36] Tiance Dong, Chenxi Liang, and Xu He. 2017. Social media and internet public events. *Telemat. Informatics* 34, 3 (2017), 726–739.
- [37] David Donovan. 2016. Mood, emotions and emojis: conversations about health with young people. *Ment. Heal. Pract.* 20, 2 (October 2016), 23–26. DOI:<https://doi.org/10.7748/mhp.2016.e1143>
- [38] Sarmistha Dutta and Munmun De Choudhury. 2020. Characterizing Anxiety Disorders with Online Social and Interactional Networks. In *HCI International 2020 -- Late Breaking Papers: Interaction, Knowledge and Social Media*, Springer International Publishing, Cham, 249–264.
- [39] 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. *Transl. issues Psychol. Sci.* 2, 3 (2016), 227.
- [40] Daniel Eisenberg, Sarah E Gollust, Ezra Golberstein, and Jennifer L Hefner. 2007. Prevalence and correlates of depression, anxiety, and suicidality among university

- students. *Am. J. Orthopsychiatry* 77, 4 (2007), 534–542.
- [41] 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.
 - [42] Weiguo Fan and Michael D Gordon. 2014. The power of social media analytics. *Commun. ACM* 57, 6 (2014), 74–81.
 - [43] 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.
 - [44] 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 *Proceeding of LPPM UPN “Veteran” Yogyakarta Conference Series 2020--Engineering and Science Series*, 253–261.
 - [45] Johannes Feldhege, Markus Moessner, and Stephanie Bauer. 2020. Who says what? Content and participation characteristics in an online depression community. *J. Affect. Disord.* 263, (February 2020), 521–527. DOI:<https://doi.org/10.1016/j.jad.2019.11.007>
 - [46] 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, and others. 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.
 - [47] 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. Disord.* (2021).
 - [48] 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. Ambient Intell. Humaniz. Comput.* 11, 11 (2020), 4713–4729.
 - [49] 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*, 236–246.
- [50] 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.
 - [51] 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.
 - [52] 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. Intern. Med.* 35, 9 (2020), 2798–2800.
 - [53] Ayaan Haque, Viraj Reddi, and Tyler Giallanza. 2021. Deep Learning for Suicide and Depression Identification with Unsupervised Label Correction. *arXiv Prepr. arXiv2102.09427* (2021).
 - [54] 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.
 - [55] 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.
 - [56] Nicholas S Holtzman and others. 2017. A meta-analysis of correlations between depression and first person singular pronoun use. *J. Res. Pers.* 68, (2017), 63–68.
 - [57] 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. *SSRN Electron. J.* (2020). DOI:<https://doi.org/10.2139/ssrn.3551338>
 - [58] 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*.
 - [59] Md Rafiqul Islam, Abu Raihan M Kamal, Naznin Sultana, Robiul Islam, Mohammad Ali

- Moni, and others. 2018. Detecting depression using k-nearest neighbors (knn) classification technique. In *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, 1–4.
- [60] Julia Ive, George Gkotsis, Rina Dutta, Robert Stewart, and Sumithra Velupillai. 2018. Hierarchical neural model with attention mechanisms for the classification of social media text related to mental health. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, 69–77.
- [61] 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 Ment. Heal.* 7, 6 (2020), e19347.
- [62] 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. Heal. Informatics* 24, 10 (2020), 2733–2742.
- [63] 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.* (2020).
- [64] 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 *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, 1689–1694.
- [65] 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.
- [66] Binti Kholifah, Iwan Syarif, and Tessy Badriyah. 2020. Mental Disorder Detection via Social Media Mining using Deep Learning. *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control* (November 2020), 309–316.
DOI:<https://doi.org/10.22219/kinetik.v5i4.1120>
- [67] Jang Hyun Kim and Yunhwan Kim. 2019. Instagram user characteristics and the color of their photos: Colorfulness, color diversity, and color harmony. *Inf. Process. Manag.* 56, 4 (July 2019), 1494–1505. DOI:<https://doi.org/10.1016/j.ipm.2018.10.018>
- [68] Jiin Kim, Zara Ahmad, Yena Lee, Flora Nasri, Hartej Gill, Roger McIntyre, Lee Phan, and

- Leanna Lui. 2021. Systematic Review of the Validity of Screening Depression through Facebook, Twitter, and Instagram. *J. Affect. Disord.* (February 2021). DOI:<https://doi.org/10.1016/j.jad.2020.08.091>
- [69] 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 (December 2020), 11846. DOI:<https://doi.org/10.1038/s41598-020-68764-y>
- [70] Jing Xuan Koh and Tau Ming Liew. 2020. How loneliness is talked about in social media during COVID-19 pandemic: text mining of 4,492 Twitter feeds. *J. Psychiatr. Res.* (2020).
- [71] Kurt Kroenke, Robert L Spitzer, and Janet B W Williams. 2001. The PHQ-9: validity of a brief depression severity measure. *J. Gen. Intern. Med.* 16, 9 (2001), 606–613.
- [72] Mrinal Kumar, Mark Dredze, Glen Coppersmith, and Munmun De Choudhury. 2015. Detecting Changes in Suicide Content Manifested in Social Media Following Celebrity Suicides. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT '15*, ACM Press, New York, New York, USA, 85–94. DOI:<https://doi.org/10.1145/2700171.2791026>
- [73] E Megan Lachmar, Andrea K Wittenborn, Katherine W Bogen, and Heather L McCauley. 2017. # MyDepressionLooksLike: Examining public discourse about depression on Twitter. *JMIR Ment. Heal.* 4, 4 (2017), e43.
- [74] C Lalrinmawii, Somen Debnath, and others. 2020. Analysis of Post centric suicidal expressions and classification on the Social Media Post: Twitter. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–5.
- [75] Victor Leiva and Ana Freire. 2017. Towards suicide prevention: early detection of depression on social media. In *International Conference on Internet Science*, 428–436.
- [76] 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.
- [77] 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 (March 2020), 2032. DOI:<https://doi.org/10.3390/ijerph17062032>
- [78] 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 2020 International Conference on Multimedia Retrieval*, ACM, New York, NY, USA, 407–411. DOI:<https://doi.org/10.1145/3372278.3391932>
- [79] 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.
- [80] 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 Ment. Heal. Res.* 1, 1 (2022), 1–8.
- [81] 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 (October 2020), e22635. DOI:<https://doi.org/10.2196/22635>
- [82] 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 Informatics J.* 26, 2 (2020), 738–752.
- [83] Katerina Lup, Leora Trub, and Lisa Rosenthal. 2015. Instagram# instasad?: exploring associations among instagram use, depressive symptoms, negative social comparison, and strangers followed. *Cyberpsychology, Behav. Soc. Netw.* 18, 5 (2015), 247–252.
- [84] 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.
- [85] Yihua Ma and Youliang Cao. 2020. Dual Attention based Suicide Risk Detection on Social Media. In *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 637–640.
- [86] 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. Pervasive Heal. Technol.* 6, 21 (2020), e1.
- [87] 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*, 440–451.

- [88] Davide Marengo, Michele Settanni, and Fabrizia Giannotta. 2019. “Development and preliminary validation of an image-based instrument to assess depressive symptoms.” *Psychiatry Res.* 279, (September 2019), 180–185.
DOI:<https://doi.org/10.1016/j.psychres.2019.02.059>
- [89] 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>
- [90] 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 Transl. Sci. Proc.* 2015, (2015), 122.
- [91] Sanjana Mendu, Mehdi Boukhechba, Anna Baglione, Sonia Baee, Congyu Wu, and Laura Barnes. 2019. SocialText: A Framework for Understanding the Relationship Between Digital Communication Patterns and Mental Health. In *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, IEEE, 428–433.
DOI:<https://doi.org/10.1109/ICOSC.2019.8665567>
- [92] Yiwen Meng, William Farran Speier, Michael Ong, and Corey Arnold. 2020. HCET: Hierarchical Clinical Embedding with Topic Modeling on Electronic Health Record for Predicting Depression. *IEEE J. Biomed. Heal. Informatics* (2020), 1–1.
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 *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, 646–649.
- [94] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, and others. 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. *arXiv Prepr. arXiv2011.11226* (2020).
- [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 *17th International Conference on Information Technology--New Generations (ITNG 2020)*, 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 *International Conference on Computational Collective Intelligence*, 811–820.
- [99] Sonya Negriff. 2019. Depressive symptoms predict characteristics of online social networks. *J. Adolesc. Heal.* 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 *ICML*, 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] Roger O’Sullivan, Brian Lawlor, Annette Burns, and Gerry Leavey. 2021. Will the pandemic reframe loneliness and social isolation? *Lancet Heal. Longev.* 2, 2 (2021), e54–e55.
- [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. Human 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] 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 (October 2013), e217. DOI:<https://doi.org/10.2196/jmir.2718>
- [107] Zhichao Peng, Qinghua Hu, and Jianwu Dang. 2019. Multi-kernel SVM based depression recognition using social media data. *Int. J. Mach. Learn. Cybern.* 10, 1 (2019), 43–57.
- [108] James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahw. Lawrence Erlbaum Assoc.* 71, 2001 (2001), 2001.

- [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. Human Behav.* 62, (2016), 155–167.
- [110] 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.
- [111] Lenore Sawyer Radloff. 1977. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 3 (1977), 385–401.
- [112] E Rajesh Kumar, KVS N Rama Rao, Soumya Ranjan Nayak, and Ramesh Chandra. 2020. Suicidal ideation prediction in twitter data using machine learning techniques. *J. Interdiscip. Math.* 23, 1 (2020), 117–125.
- [113] 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 (December 2017), 13006. DOI:<https://doi.org/10.1038/s41598-017-12961-9>
- [114] Andrew G Reece and Christopher M Danforth. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Sci.* 6, (2017), 1–12.
- [115] 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.
- [116] 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.
- [117] 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 Digit. Med.* 3, 1 (2020), 1–12.
- [118] 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. Heal. Informatics* 20, 4 (July 2016), 1008–1015. DOI:<https://doi.org/10.1109/JBHI.2016.2543741>

- [119] Koustuv Saha, John Torous, Eric D Caine, Munmun De Choudhury, and others. 2020. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. *J. Med. internet Res.* 22, 11 (2020), e22600.
- [120] Koustuv Saha, Asra Yousuf, Ryan Boyd, James W Pennebaker, and Munmun De Choudhury. 2021. Mental Health Consultations on College Campuses: Examining the Predictive Ability of Social Media. *Available SSRN 3774189* (2021).
- [121] Flora Sakketou and Nicholas Ampazis. 2020. A constrained optimization algorithm for learning GloVe embeddings with semantic lexicons. *Knowledge-Based Syst.* 195, (2020), 105628.
- [122] 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.
- [123] 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 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 7685–7697.
- [124] 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.
- [125] Sebastian Scherr, Florian Arendt, Thomas Frissen, and José Oramas M. 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.
- [126] Annika Marie Schoene, Alexander Turner, Geeth Ranmal De Mel, and Nina Dethlefs. 2021. Hierarchical Multiscale Recurrent Neural Networks for Detecting Suicide Notes. *IEEE Trans. Affect. Comput.* (2021), 1–1.
DOI:<https://doi.org/10.1109/TAFFC.2021.3057105>
- [127] 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 *2020 IEEE Region 10 Symposium (TENSYP)*, IEEE, 823–826.
DOI:<https://doi.org/10.1109/TENSYP50017.2020.9231008>
- [128] Ashish Sharma, Monojit Choudhury, Tim Althoff, and Amit Sharma. 2020. Engagement

- patterns of peer-to-peer interactions on mental health platforms. In *Proceedings of the International AAAI Conference on Web and Social Media*, 614–625.
- [129] Ashish Sharma, Inna W Lin, Adam S Miner, David C Atkins, and Tim Althoff. 2021. Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. In *Proceedings of the Web Conference 2021*, 194–205.
 - [130] 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. *Cyberpsychology, Behav. Soc. Netw.* 23, 9 (September 2020), 611–618. DOI:<https://doi.org/10.1089/cyber.2019.0746>
 - [131] 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. Heal. Informatics Bioinforma.* 9, 1 (2020), 1–11.
 - [132] Ruba Skaik and Diana Inkpen. 2020. Using Social Media for Mental Health Surveillance: A Review. *ACM Comput. Surv.* 53, 6 (2020), 1–31.
 - [133] 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.
 - [134] 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.
 - [135] 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. Psychiatry* 33, 4 (2021), 412–423.
 - [136] Maxim Stankevich, Ivan Smirnov, Natalia Kiselnikova, and Anastasia Ushakova. 2019. Depression detection from social media profiles. In *International Conference on Data Analytics and Management in Data Intensive Domains*, 181–194.
 - [137] 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.
 - [138] 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. Intern. Med.* 35, 7 (2020), 2244–2247.
- [139] 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. *Compr. Psychiatry* 55, 4 (2014), 866–869.
- [140] 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.
- [141] 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. Pers. Soc. Psychol.* 118, 2 (2020), 364.
- [142] 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>
- [143] 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.
- [144] Mohd Tajuddin, M Kabeer, and Mohammed Misbahuddin. 2020. Analysis of Social Media for Psychological Stress Detection using Ontologies. In *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*, 181–185.
- [145] M L Tlachac and Elke Rundensteiner. 2020. Screening for depression with retrospectively harvested private versus public text. *IEEE J. Biomed. Heal. informatics* 24, 11 (2020), 3326–3332.
- [146] 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. Inf. Syst.* 60, 1 (2019), 179–196.
- [147] Tuan Tran, Dong Nguyen, Anh Nguyen, and Erik Golen. 2018. Sentiment Analysis of Marijuana Content via Facebook Emoji-Based Reactions. In *2018 IEEE International Conference on Communications (ICC)*, IEEE, 1–6. DOI:<https://doi.org/10.1109/ICC.2018.8422104>
- [148] Alina Trifan, Dave Semeraro, Justin Drake, Radek Bukowski, and José Luís Oliveira. 2020. Social Media Mining for Postpartum Depression Prediction. *Stud. Health Technol.*

- Inform.* 270, (2020), 1391–1392.
- [149] Marcel Trotzek, Sven Koitka, and Christoph M Friedrich. 2018. Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences. *IEEE Trans. Knowl. Data Eng.* 32, 3 (2018), 588–601.
 - [150] Danny Valdez, Marijn Ten Thij, Krishna Bathina, Lauren A Rutter, Johan Bollen, and others. 2020. Social media insights into US mental health during the COVID-19 pandemic: Longitudinal analysis of Twitter data. *J. Med. Internet Res.* 22, 12 (2020), e21418.
 - [151] R. Vanlalawmpuia and Mr Lalhmingliana. 2020. Prediction of Depression in Social Network Sites Using Data Mining. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, 489–495.
DOI:<https://doi.org/10.1109/ICICCS48265.2020.9120899>
 - [152] D. Vivek and P. Balasubramanie. 2020. An Ensemble Learning Model to Predict Mental Depression Disorder Using Tweets. *J. Med. Imaging Heal. Informatics* 10, 1 (January 2020), 143–151. DOI:<https://doi.org/10.1166/jmihi.2020.2871>
 - [153] 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>
 - [154] 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.
 - [155] 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 *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 201–213.
 - [156] 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.
 - [157] 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 First International Workshop on Language Cognition and Computational Models*, 11–21.
 - [158] Akkapon Wongkoblaph, Miguel A Vadillo, and Vasa Curcin. 2019. Modeling Depression

- Symptoms from Social Network Data through Multiple Instance Learning. *AMIA Jt. Summits Transl. Sci. proceedings. AMIA Jt. Summits Transl. Sci.* 2019, (May 2019), 44–53. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/31258955>
- [159] World Health Organization. 2019. Suicide key facts. Retrieved April 14, 2021 from <https://www.who.int/news-room/fact-sheets/detail/suicide>
- [160] World Health Organization. 2020. Depression fact sheet.
- [161] Ho Chung Wu, Robert Wing Pong Luk, Kam Fai Wong, and Kui Lam Kwok. 2008. Interpreting tf-idf term weights as making relevance decisions. *ACM Trans. Inf. Syst.* 26, 3 (2008), 1–37.
- [162] 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.
- [163] 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. psychiatry* 11, (2020), 144.
- [164] 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>
- [165] 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.
- [166] 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 (December 2020), 8931. DOI:<https://doi.org/10.3390/ijerph17238931>
- [167] 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. *Inf. Discov. Deliv.* (2021).
- [168] Hui Yin, Shuiqiao Yang, and Jianxin Li. 2020. Detecting topic and sentiment dynamics due to Covid-19 pandemic using social media. In *International Conference on Advanced Data Mining and Applications*, 610–623.
- [169] 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. *Heal. Data Sci.* 2022, (2022).
- [170] 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 (November 2020), e23575. DOI:<https://doi.org/10.2196/23575>
- [171] Yin Zhang, Rong Jin, and Zhi-Hua Zhou. 2010. Understanding bag-of-words model: a statistical framework. *Int. J. Mach. Learn. Cybern.* 1, 1–4 (2010), 43–52.
- [172] Yipeng Zhang, Hanjia Lyu, Yubao Liu, Xiyang Zhang, Yu Wang, and Jiebo Luo. 2020. Monitoring Depression Trend on Twitter during the COVID-19 Pandemic. *arXiv Prepr. arXiv2007.00228* (2020).