SOCIAL NETWORK SENTIMENT ANALYSIS TOOLS FOR THE DETECTION OF MENTAL ILLNESS: A SYSTEMATIC MAPPING

Ariel Altamirano¹, Daniel Avila¹ and Graciela Guerrero^{1,2(0000-0002-0903-734X)}

¹ *Universidad de las Fuerzas Armadas ESPE, Avenida General Rumiñahui, 171103 Sangolquí, Ecuador*

² *Universidad de Castilla – La Mancha*

Departamento de Sistemas Informáticos, campus universitario s/n 02071, Albacete, España

ABSTRACT

The growing demand of users on the Internet and their dependence on it has caused websites such as Twitter or Facebook to become a source of personal information for each user, which is why, within the psychological context, one can take advantage of it by analyzing the publications of each of these users; on the other hand, the importance of detecting mental disorders in a person has grown exponentially, so that in recent years several analysis tools have been developed to detect a person's mental state. This systematic mapping work details relevant information about the analysis tools and their importance to society, as well as describing the methodology used in this article, its results, conclusions and proposals for future work.

KEYWORDS

Feelings Analysis, Social Network Analysis, Mental Illness, Social Networks, Anxiety

1. INTRODUCTION

According to (Wongkoblap, et al., 2017), mental illness is becoming a serious global health problem, with a growing number of patients suffering from depression, anxiety and other disorders. This issue should be highlighted and given importance because out of 350 million patients worldwide suffering from mental disorders, it also impacts the world economy at a cost of US \$2.5 trillion in 2010 and an estimated expenditure of US \$6 trillion by 2030.

Anxiety is the most common mental health condition in children and adolescents, with prevalence rates ranging from 5% to 10% in children and up to 25% in adolescents (Glover & Fritsch, 2018). Adolescents are the most likely to suffer from anxiety disorders which, according to (Gerez, et al., 2016) are the most pervasive mental health problems and despite the devastating effects of anxiety disorders (ADS) worldwide, debates about causes and remedies have not solved the psychologist's riddle: should they be treated and how? Anxiety disorders have consequences for an individual's mental health, which leads to a certain insecurity and inferiority when it comes to offering professional advice. According to (Deepali, et al., 2018) unlike physical illness, in mental illness the symptoms are not evident in the early stages, for this reason psychologists look for different ways to understand the patient in treatment, so, according to (Jia, 2018) with the popularity of social networks it is feasible to take advantage of these data for stress and detection of depression and anxiety.

Over the past 10 years, social networks have become an important component of young people's lives, with the use of a peak in 16-24 years (Glover & Fritsch, 2018). Today, the use of social networks and other popular microblogging tools has become an everyday activity since in addition to being a means of communication it is a source of information and according to (Miks & Aurel, 2019) they allow individuals to express their opinions, feelings and thoughts on a variety of topics in the form of short text messages or tweets. When published on social network platforms, the day-to-day feelings or thoughts of each user become valuable information and according to (Manikonda & De Choudhury, 2017) is self-revelation by the user. Although there is a debate about the excessive use of social networks, whether they are for personal benefit or harm, since it is not possible to control what type of information is true within social networks or whether personal

information is completely secure; according to (Park, et al., 2015) there are recent studies that have used data from online social networks to predict the well-being of the individual.

Several published articles have studied mental illness and recommended various free tools for the analysis of feelings, however, no systematic mapping articles like the one presented in this article have been found. According to (Cooper, 2016) systematic mappings are critical, but they do not discuss the findings. They are based on the concept that published articles not only represent findings, but, indirectly, represent the activity related to the finding. Therefore, this article will provide an overview to help psychologists and software developers to have a clear and comprehensive idea of sentiment analysis and social network tools.

This article is structured as follows: i) Section 2: introduction of the methodology to be used, research guidelines, response to research questions and data analysis and extraction. ii) Section 3: the results of the study and its limitations are presented. iv) Section 4: the results obtained are discussed. Finally, iv) Section 5 presents the conclusions obtained.

2. METHOD

In this section, the research methodology used is detailed, as well as the interest of the research in which the reason for the proposed subject matter is stated, research guidelines in which the research questions are stated, source selection in which the research databases used are stated, formulation of the search chain in which the structure of the chain used is detailed, exclusion and inclusion criteria, data collection and the search process.

2.1 Research Methodology

2.1.1 Research Interest

There are several reasons for conducting this systematic mapping review. The first is to conduct a broad review of the literature related to sentimental analysis or social network tools for the detection of mental illness, with the aim of providing summaries of which tools have been used according to two specific criteria: sentimental analysis or social network analysis. The second reason, if possible, is to identify what kind of complications exist in the detection of mental illness when using sentimental analysis tools or social networks to provide facilities for psychologists, academics and software developers to plan future research in this area. The third reason is to identify topics and areas for future systematic reviews of the literature that will cover smaller areas of research after analyzing the literature in depth.

Figure 1. Systematic mapping review process applied at work

To guide this empirical study, the "Template for a Mapping Study Protocol" proposed by (EBSE-RG, 2009) was taken into consideration. As shown in Figure 1, the research protocol has three phases: (1) Research guidelines that define the protocol applied for this study, identify which dimensions should be analyzed, and determine which research questions should be answered throughout the work. (2) Data collection identifies relevant documents, according to the inclusion and exclusion criteria defined in the protocol. (3) The results phase leads to the mapping of the existing literature, according to the defined criteria and the respective analysis to answer the proposed research questions.

2.1.2 Research Directives

In this section, the first phase of the systematic mapping review. The process is introduced, where the protocol and research questions are defined.

The protocol includes the subject of the study (sentimental analysis tools or social networks for the detection of mental illness); its rationale (explained above); preliminary research questions; and the search strategy, selection criteria, and form of data extraction. An overview of the empirical screening approaches found in the selected articles is also provided. Finally, the protocol includes an overview of the selected articles in terms of their publication locations and years.

The three research questions (RQ) for this study and their rationale are

RQ1. How many articles have been published on sentiment analysis tools or social networks for the detection of mental illness? What is their temporal trend?

As argued in the Introduction section, mental illness can be understood as mental health conditions such as anxiety and depression, which are most diagnosed in adolescents and children. From this context, a first step might be to determine the number of articles that have been published on sentimental analysis tools or social networks for the detection of mental illness and second, whether there is any temporal trend (increasing or decreasing) in the publication of these articles.

RQ2. What types of mental illness are investigated with sentimental analysis tools or social networks? What characteristics do these analysis tools possess?

As is known, there are several types of mental illness, however, it is necessary to know which mental illnesses are investigated or sought to be detected when using any sentiment analysis or social network tool. Likewise, understanding the characteristics that these different tools have.

RQ3. What type of analysis has been used for the detection of mental illness and on what social network platforms have these tools been developed?

It is necessary to find out which type of analysis is the most used when detecting a mental illness. Similarly, to know on which social network platform these tools are applied or used more, to know the reason for this situation or to distinguish if there is any reason for privacy, number of users, security for which such an effect occurs; if it exists.

2.2 Data Collection

Data collection should lead to the inclusion of relevant documents and the exclusion of irrelevant documents. The search strategy used was to query relevant databases, create search strings and enter the selected databases, evaluate the results obtained and finally apply inclusion and exclusion criteria. All researchers were involved in the execution of data collection.

2.2.1 Source Selection and Search String

In order to avoid the loss of information, the consultation was carried out in three relevant databases: Scopus (Elsevier), ACM (Association for Computing Machinery) and IEEE (Institute of Electrical and Electronics Engineers).

From the research questions, the key words in the search string were identified. Synonyms and relationships of the keywords were identified to form the search string that is connected to OR and AND respectively. The search string entered in the different databases is as follows: (Mental Health OR Anxiety OR Depression OR Moods) AND ((Social Network Analysis OR Sentiment Analysis) AND (Application OR Tool)).

The first part of the search string consists of words related to mental illness, while the second part consists of words related to sentiment analysis and social network analysis, and finally the third part consists of tools for sentiment analysis. Each part is related to AND connectors.

2.2.2 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria for filtering the articles were

- Inclusion criteria

- I1. Articles published between 2014 and 2020
- I2. Articles that include the selected keywords within the title and are in turn within the keywords of each article.

- Exclusion criteria

- E1. Words like "Education", "Big Data" and their synonyms are within the titles and keywords of the articles.
- E2. Articles with themes and areas outside the context of the research.
- E3. Articles repeated among the three databases.

2.2.3 Search Process

As shown in Figure 2 the search string was entered in each of the above-mentioned databases and different item numbers were obtained from each of them, making a total of 4142.

Applying the inclusion criteria (E1, E2 and E3) to the 4142 articles, each researcher reading the summary of each article ruled out those that were not in the context of the investigation, in case the researcher was unsure of keeping the article decided to read the introduction and conclusion of the article and together with the opinion of the other researcher the article in question was included or discarded. At the end of the analysis, 15 articles were obtained as a result at this stage.

Figure 2. Result of the search for articles in each of the databases applying the inclusion I1 and I2

Figure 3 shows the number of articles selected among all the articles shown in figure 2. You can appreciate the importance of searching in different databases since by searching in only one you run the risk of losing information from other articles.

Figure 3. Results of the articles obtained by applying the criteria of E1, E2 and E3

3. RESULTS

Finally, 15 articles have been selected according to the process described above. Then, descriptive statistics and frequency analysis were used to answer the research questions. An overview of the articles and the answers to the three research questions are shown in the following subsections.

3.1 Summary of Selected Articles

The summary of the selected articles was presented in terms of their locations and years of publication.

3.1.1 Journal and Conference Publications

Figure 4 shows the journals where more articles have been published on the use of sentiment analysis tools or social networks for the detection of mental illness

Figure 4. Articles published in journals and conferences according to the topic of research

3.1.2 RQ1. How Many Articles have been published on Sentiment Analysis Tools or Social Networks for the Detection of Mental Illness? What is their Temporal Trend?

In the study period from January 1, 2014 to December 30, 2019, a total of 15 published articles on sentiment analysis tools or social networks for the detection of mental illness were found. An upward temporal trend was observed from 2014 to 2018 as shown in Figure 5., where it is intuited that this behavior occurs, in proportion to discarding the Big Data theme as an exclusion criterion.

Figure 5. Temporal Trend (2014 – 2019)

3.1.3 RQ2. What kind of Mental Illnesses are investigated in the Tools Of Analysis of feelings or Social Networks? What Characteristics do These Analysis Tools have?

Among the different types of mental illness or mental health states, the most studied in relation to the number of people who are diagnosed with such a mental state are anxiety, depression and stress. Figure 6 shows the type of mental illness studied and its importance and prominence within the articles selected in the period described above. The characteristics of social network analysis tools are shown in Table 1. Similarly, the characteristics of feeling analysis tools are shown in Table 2, which compares OSAT with the other feeling analysis tools (Sharma, et al., 2016).

Figure 6. Type of mental illness where tools for detection are made or linked

Table 1. Characteristics of the functioning of social network analysis tools

Software	Type	Platform	Calculation Time	Number of Nodes
NETWORKX	Library	Python	Fast	1 million
IGRAPH	Library	Biblioteca Python/R/C	Fast	1 million
GEPHI	Independent	Windows	Fast	0.15 million
PAJEK	Independent	Windows	Medium	1 million

Table 2. Characteristics of the functioning of social network analysis tools

Software	Accuracy	Precision
Sentic Net	0.59	0.94
Sentic Word Net	0.63	0.79
Senti Strength	0.81	0.78
LIWC	0.68	0.84
OSAT	0.85	0.9

3.1.4 RQ3. What Type of Analysis has been used for the Detection of Mental Illness and on what Social Network Platforms have these tools been developed?

As established in this paper, two types of analysis can be used for the detection of mental illness. Sentimental analysis is defined as the process of extracting the polarity of a subjective sentence, where it is classified into three main categories: lexical-based method, automatic learning method, hybrid method (Rathan, et al., 2017). Social network analysis is defined as the way to identify, analyze, visualize or simulate nodes and edges of various types of input, including mathematical models of social networks (Akhtar, 2014). However, it was found that when any tool is used to link it to the detection of mental illness, only sentiment analysis tools have been used, which capture three general feelings (negative, neutral and positive).

The social media platforms where this type of opinion analysis tools or social networks are most used are Facebook and Twitter; are the most mentioned platforms in all selected articles, where for both cases facilities are exposed and difficulties in obtaining data. However, Twitter is the social network platform most used for this type of opinion analysis, because unlike Facebook, Twitter users' posts are public, so you don't need permissions to access them and perform the respective analyses.

4. DISCUSSION

This article shows the literary background for the development of a future software. The articles researched between the years 2014-2019 have produced results, which have been presented in the answers to the research questions.

Figure 7 shows a representation of the citations of each article that was used for the development of this work.

Figure 7. Number of quotations of the selected articles

As for the tools researched, we found two types to differentiate, which are: tools for the analysis of feelings in social networks and tools for the analysis of social networks despite the existence of a large number of software tools, (Akhtar, 2014) defines social network analysis as to identify, analyze, visualize or simulate nodes and edges of various input types, including mathematical social media models and recommends four analysis tools that are: Gephi, NetworkX, igraph and Pajek; this tools are describe on table 1; recommends them for being tools of free use and that can handle large graphic size. On the other hand, (Afra, et al., 2018) propose NetDriller2 as a powerful tool for the analysis of social networks, explaining that this tool covers the lack of other tools by building a network of raw data using data mining techniques.

With regard to sentiment analysis; (Abhijit, et al., 2016) emphasize that sentiment analysis is much better at capturing general feelings (negative, neutral or positive) expressed in star ratings. On the other hand, (Cao, et al., 2014) define sentiment analysis as the mining of affective information from data and the recognition of the polarity of feeling contained in the information; they also stress that most studies of sentiment analysis have focused on the English language.

5. CONCLUSIONS AND FUTURE WORK

According to (Cerón Gúzman & León Gúzman, 2016), thanks to platforms such as Facebook and Twitter, profound changes have taken place in the paradigm of information generation and consumption. However, it should be noted that by using tweets as a source of data, there would be no problem with the privacy of such information, since this information is public. When using data provided by Facebook there are problems, since this information is private to the user and sometimes users decide not to lend their Facebook information.

According to (Wongkoblap, et al., 2017) social network data are a valuable source of information for predicting mental health problems. Understanding the factors associated with symptoms of mental illness can contribute to early detection and treatment of network users with poor mental health.

In the selected articles, screening for mental illness is only done with sentimental analysis tools, so a future job could be to carry out screening work for mental illness using only social network analysis tools or to use both types of analysis.

Finally, it was determined that anxiety and depression are the most researched types of mental illness in the last 5 years; related to the early detection of them.

REFERENCES

- Abhijit, J. P., Kshama, K. & Kotrappa, S., 2016. Emotweet Sentiment Analysis tool for twitter. *IEEE Xplore*, pp. 157-159.
- Afra, S., Ozyer, T., Rokne, J. & Alhajj, R., 2018. NetDriller Version 2 A Powerful Social Network Analysis Tool. *IEEE Explore*, pp. 1475-1480.
- Akhtar, N., 2014. Social Network Analysis Tools. *IEEE Explore*, pp. 388-392.
- Cao, J. y otros, 2014. Web-Based Traffic Sentiment Analysis Methods and Applications. *IEEE Xplore*, pp. 844-853.
- Cerón Gúzman, J. A. & León Gúzman, E., 2016. A Sentiment Analysis System of Spanish Tweets and Its Application in Colombia 2014 Presidential Election. *IEEE Xplore*, pp. 250-257.
- Cooper, D., 2016. What is a “mapping study?”. *JMLA*, pp. 76-78.
- Deepali, J., Makhija, J., Nabar, Y. & Nehet, N., 2018. Mental Health Analysis Using Deep Learning for Feature Extraction. *ACM*, pp. 356-359.
- EBSE-RG, 2009. Durham University Community. [Online] Available at: <http://community.dur.ac.uk/ebse/resources/templates/MappingStudyTemplate.pdf> [Last access: November,11-2019].
- Gerez, M. y otros, 2016. The crossroads of anxiety: Distinct neurophysiological maps for different symptomatic groups. *SCOPUS*, pp. 159-175.
- Glover, J. & Fritsch, S., 2018. #KidsAnxiety and Social Media: A Review. *SCOPUS*, pp. 171-182.
- Jia, J., 2018. Mental Health Computing via Harvesting Social Media Data. *ACM*, pp. 70-70.
- Manikonda, L. & De Choudhury, M., 2017. Modeling and Understanding Visual Attributes of Mental Health Disclosures in Social Media. *ACM*, pp. 170-181.
- Miks, Q. & Aurel, J., 2019. Sentiment Analysis on Tweets with Punctuations, Emoticons, and Negations. *ACM*, pp. 266-270.
- Park, S., Kim, I., Won, S. & Yoo, J., 2015. Manifestation of Depression and Loneliness on Social Networks: A Case Study of Young Adults on Facebook. *ACM*, pp. 557-570.
- Rathan, Hulipalled, V., Murugeswari & Sushmitha, 2017. Every Post Matters A Survey on Applications of Sentiment Analysis in Social Media. *IEEE Xplore*, pp. 709-714.
- Sharma, N., Sharma, C., Indu & Chugh, P., 2016. Optimized Sentiment Analysis Tool A sentiment analysis tool to study cognitive inclinations. *IEEE Xplore*, pp. 2589-2592.
- Wongkoblap, A., Vadillo, M. & Curcina, V., 2017. Detecting and Treating Mental Illness on Social Networks. *IEEE Explore*, p. 330.