IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

In Mental Health Detection Application of Machine Learning Methods

Sangita Roy

ECE Department, Narula Institute of Technology, WBUT, Agarpara, Kolkata, India.

Purba Das

Electronics & Communication Engineering Department, NIT, Agarpara, Kolkata, India.

ABSTRACT -- In online social networks (OSNs) based on Machine learning techniques, data sources and feature extraction method this paper presents a critical assessment analysis on mental health detection. By identifying its data analysis method, challenges, comparison and limitations the appropriateness of mental health detection was also investigated. Between 2017 and 2018 through keyword searches articles of this study published major databases. Before the full texts were reviewed the articles were screened based on their titles and abstracts . In accordance with data set the articles were coded (e.g. , keywords , data sources and geographical locations) , deep learning or machine learning technique , feature extraction method, classifier performance and method of data analysis . From the total of 2770 , 22 articles were selected for review. Several researchers used text analysis in early detection of mental health problem detection on a new data set removed from different OSNs sources. Using a statistical analysis or machine learning techniques removed data were analysed. Multimethod techniques also applied in several studies.Multimethod techniques have been included distributing questionnaires while requesting for the respondents consent to later access and remove information from OSNs account. In OSNs big data's are contribute on mental health problem detection. The early detection's alternative method is presented method for mental health problems rather than using traditional strategies, such as devices and sensors which are time consuming and costly.

INDEX TERMS – Deep learning, feature extraction, machine learning, mental health, online social network.

I. INTRODUCTION

Day by day health problems are increasing. Mental health problems are coming because some peoples are not able to take stress because all of the humans ability is not same .Many people's have family problems and some have financial issues and that he or she is not able to take. And from these many health problems are coming like anxiety, depression, bipolar disorder (BD), psychotic disorder, and schizophrenia. Mental health problem has defined by the World Health Organization .

Approximately 450 million of people are mentally ill with the global disease .Because of these mental health problems people die earlier than the general population. For depression people's can take suicidal attempts and suicidal ideation. People suffering from mental illness who are able to recover and live a normal person life as outlined in the Comprehensive Mental Health Action Plan (2013–2020) .

Mental health problems should be detected and addressed early.Accurate diagnosis, early detection and effective treatment can improve the suffering of people who are dealing with mental health problem of challenges. The effects of mental ill-ness can be severe on the concerned individuals and their families, and on the society as a whole. In general, the traditional methods of mental health detection normally use face-to-face interviews, self-reporting, or questionnaire distribution. Mental health problem detection's traditional methods are normally use for

face to face interviews and self reporting. Technologies have been applied from previous studies such as smartphones in health care, wearable sensors and mental health detection; these technologies are typically used for those who suffer from mental illness and who have been monitored overtime .

In online social networks a recent research presented a novel approach of mental health problem detection. In recent years OSNS have become popular and have provided a new medium to communicate and share information .OSNs are used regularly by millions of people worldwide . Users can express their feelings and thoughts through OSNS by posting different types of data (e.g. Text, images, videos, and audios) regarding their daily activities and they can also communicate with friends by commenting on the posts of others .Thus, this new research trend is related to big data research and the increasing availability of resources on the Internet through OSNs (e.g., Facebook, YouTube, Twitter, Instagram, and Sina Weibo).Increases of data overflows and these huge amount of communication data have become content generators that may be useful for further exploration and analysis .Therefore, OSNs can generate a massive amount of information that can be used to develop an approach for mental health Problem detection .

Researchers from the West and the East harvest data from OSNs, such as Twitter, Facebook, and Sina Weibo, and use them as data source for online studies and crowdsourcing . the types of mental health problems detected in OSNs - depression , psychological stress ,mental disorders and suicidal ideation .An analysis of the current mental health detection in OSNs is required to comprehend data sets, data analysis methods, feature extraction method, classifier performance (i.e., accuracy and efficiency), challenges, limitations, and future work. The purpose of this systematic review is to conduct a critical assessment analysis of mental health problem detection based on data extracted from OSNs. It intends to explore the competence of mental health problem detection in OSNs, including its challenges, limitations, and future work .

There are two common ways to analyze the data from OSNs posted by the users and by using dictionary-based and Machine learning method. Both methods have limitations. Researchers have to find alternative ways to improve the performance and efficiency of the analysis. Traditional machine learning faces common training problems, such as model inter-Pretation, overfitting and generalization . So, researcher moved to Deep learning techniques which have changed in recent years as a powerful tool. This is because machine learning can solve more complex problems, especially in health data .

The first section discusses the objectives of this paper , The second section describes the systematic review Methodology with a discussion on the types of mental health problems , The third section presents the results of each study the data sets ,feature extraction method, classifier performance, and machine learning techniques used in each study, The fourth section displays the challenges , The fifth section recognizes the limitations , The sixth section offers criticisms for Future implication and The seventh section concludes the study.

II. METHODS

In this research paper the planning phase is followed by the searching and analysis phase. Then the discussion of the relevant documents that are found will be highlighted and summarize this paper. The conclusions will be presented to conclude this review paper.

The main purpose of this paper is to explore the adequacy, Challenges, and limitations of a mental health problem detection based on OSNs data. The objective of this review paper is to conduct a critical assessment analysis on detection of mental health problems using OSNs.This analysis consists of the data source, the feature extraction method, and classifier performance in machine learning techniques. We also investigated the appropriateness of this Pre-mental health detection by identifying its data analysis Method, comparison, challenges, and limitations.

A. IDENTIFICATION AND SELECTION OF STUDIES

Criteria of Inclusion	Criteria of Exclusion
Articles in English language	Studies that unrelated to OSN data
Articles published between 2007 and 2018	Studies that unrelated to mental health problems
Articles that used user-	Studies that unfulfilled any inclusion criterion

generated data in OSNs	
Articles discuss any type of mental health problems	

OSNS data sources in their mental health problem detection were included in the selection . OSNS is enable to online social interaction to users and and facilitate information exchange, as well as diffusion .Therefore, OSNs have created a massive amount of data and offer a Distinctive chance for learning and understanding of social Communication and interaction among considerably superior populations . Mental health problems refer to psychological disorders such as depression, anxiety . OSNS data sources are used in their research for mental health problem detection.

B. CRITERIA OF INCLUSION AND EXCLUSION

Criteria of inclusion : Articles in English language , inclusion articles are published in between 2007 and 2018, user generated data's are used in OSNS and another one criteria of exclusion: unrelated to OSN data, unrelated to mental health problems, unfulfilled any inclusion criterion.

C.METHODOLOGICAL QUALITY ASSESSMENT

This systematic review assumed the Critical Appraisal Skills Program (CASP) checklist for evaluating several methodological quality articles completely .Features and limitations both are based on the extraction of data such as data source, keywords, duration, and geographical location of data extracted; the quality of data such as data set related to mental health problems; study design such as suitable methodology applied; and the results such as clear study Objectives and outcomes, were analyzed and compared to indicate the strengths and weaknesses for each of the studies. The articles were selected based on the title, objectives, outcomes, findings, data set, feature extraction method, machine learning techniques, and classifier performance.

III.Results :

A.FINDING AND SELECTING STUDIES

2705 articles were founded through selected database searching. 65 additional articles were founded through other sources. Total number of articles were 2770. Then 35 duplicates articles were removed. After the duplicate articles were removed then the total number was decreased to 2735. 2465 articles were excluded. There were 270 full-text articles . Then 248 full-text articles were excluded. 22 articles were included for comprehensive analysis. The 22 articles were published between 2007 and 2018 in English language.

A flowchart of citation from the early stage of searching still the final number of articles followed by PRISMA and the figure is assumed in figure 1. Summary and major characteristics are discussed content analysis of the articles. In table 1 we see that the analysis of the 22 selected Selected articles that is based on the criteria of inclusion and exclusion. In accordance with the data set the summary of the selected articles (data sources, keywords , duration and geographical location), study objectives, method of data analysis, machine learning techniques, feature extraction method, and classifier performances.

B.DESCRIPTION OF SELECTED STUDIES

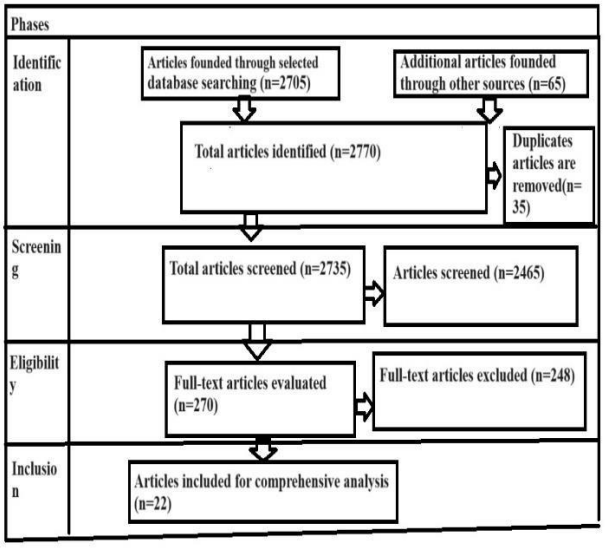

Figure 1: Flowchart of Citation from the First Stage of Identification until the Final Stage of Inclusion Articles.

C. DATA SET

All the researchers have prepared their own data sets that's are given in table 2 except for one those who are used data from another study. Data sets are provided those are based on the country where the research was conducted, the type of OSNs used and the duration of data extraction .The advantage of using an original data set is that the information is specific to the objectives of the research depending on the location . Several data were extracted directly from public posts. Three types of data were extracted. The first type included data extracted from public application programming interface services, such As Twitter or Sina Weibo . Sina Weibo is similar to Twitter but is available only in China. The second type of data was extracted from Twitter and Facebook user accounts after obtaining their consent. The users were discovered of having mental health problems based on their answers to the questionnaires distributed by the Researchers . The third type of data was removed from online sources such as web for a online communities, microblogs, crowdsourcing.

1. Geographical Location

Within their territories most of the countries have been studying about mental health diseases. With their country of origin such as Japan, China and the United States the removed data were tagged. When preparing their data sets most of the studies have been used their mother language English .

2. DATA SOURCE

Extracting data from OSNS is the main purpose of this research based on the mental health and which OSNS should be used is important determine that. Different types of OSNs were used as data sources by the previous researchers including Facebook, Twitter , Sina Weibo, microblogs, and other online sources. Data's are preferred from Twitter by most of the researchers because it is reachable to the public .Data of Twitter are rich in information, such as user ID, user name , screen name, biography, account creation data, URL and tweet texts .Other researchers have been used Facebook as their data source . Researchers who had already been identified had to request approval from the users as having mental Health problems to extract data from their Facebook or Twitter accounts . OSNs is similar to Twitter because in China, Sina Weibo has become the choice of many researchers . From microblogs other Researchers used extracted data.For data extraction various online resources were used . In the future other public resources from different types of OSNs will be possibly used for data extraction in various types of research.

3. DATA EXTRACTION DURATION

Several researchers provided their duration of data extraction for their original data sets. The longest period of data extraction indicated by the several researchers to be approximately three years . For less than one year other researchers removed data for only one week . The duration of data extraction plays a role in mental health problem detection no other researchers stated that. The longer data extraction duration a conclusion can be drawn and can be collected more data.

D. FEATURE EXTRACTION

Feature extraction is a process that reduces the number of resources without losing the syntactic and semantic relations between words. It generates new features from the original features and implements a considerable number of variables and complex data. Feature extraction can also reduce the amount of redundant data from the given analysis. Previous researchers performed N-Gram features (e.g., bigram, unigram and trigram) to create word tokens , bag-of-words , TF-IDF features to capture frequent and representative words and parts of speech . Feature extractions most common techniques are Bag of word (Bow), Term frequency – Inverse Document Frequency N-Gram,(TF-IDF), and Global Vectors for Word Representation (Glove),Word2Vec. Feature extraction methods is a crucial step in text classification for mental health problem detection.

E. MACHINE LEARNING TECHNIQUES

For mental health problem detection Machine learning techniques are currently popular . Classification techniques are used in previous research for detecting many types of mental health Problems, such as suicidal ideation, stress, depression and distress. For the analysis of data many approaches were developed those were used in various types of mental health problem detections. Several researchers generated new methods, such as the Social Network Mental Disorder Detection (SNMDD) model and tensiStrength . Hybrid

methods are developed by other scholars that comprised two types of machine learning techniques and performed deep learning techniques in their research . Past researchers compared different types of available machine learning techniques. The most commonly used machine learning techniques in mental health problem detection included Support Vector Machines (SVM), Naïve Bayes (NB), logistic regression (LR), random Forest (RF), Decision Tree (DT), Gaussian Process, K-means, and Artificial Neural Network (ANN). This algorithm is appropriate in performing text classification. Several of the text classification processes will go through for four phases: dimension reduction, feature extraction, evaluation and selection of machine learning techniques. Generally, once the date is filtered, a formal feature extraction method is implemented. Then, using machine learning techniques the data will be classified. Furthermore, compared all the machine learning techniques to the efficiency of each classifier based on the metrics of Recall,Precision, accuracy and F-measure. Deep learning technique has become popular in these few years and in text classification achieved surpassing results in comparison to the previous machine learning techniques. In mental health problem detection deep learning techniques were implemented in previous research such as sparse Deep Neural Network and Deep Neural Network. SVM is most commonly used technique which was applied by 13 researchers . One of the standard machine learning and data mining tool is SVM that is used to solve for two-group classification problems – nonlinear and linear. The second most used machine learning or deep learning techniques were LR and RF.This technique choosed by four researchers with third NB ranked. The used of machine learning or deep learning techniques leasts are Gaussian process, AdaBoost, Gradient-boosted DT, the hybrid technique of factor graph Model (FGM) with Convolutional Neural Network (CNN), Markov logic networks (MLNs),JRip Rule, Multinomial LR, Multinomial NB, sequential minimal optimization, Transductive SVM (TSVM) and SVM with Radial Basis Function TensiStrength.

IV. DISCUSSION AND CHALLENGES

The information available in OSNs and that provides a huge amount of data with immense potential to be explored in modern research. To understand the phenomenon millions of data can be extracted from OSNS for study. In recent studies many researchers used data from OSNs those related to pandemics and cyberbullying. In this research paper the researchers focused on mental health problem detection through OSNs.

V. LIMITATIONS

Limitations in this study to exhibits of articles because it is used only four journal databases that is IEEE Xplore, PubMed, ScienceDirect, Scopus. Articles published in English and also related to mental health problems were included. Collected articles are published in between 2007 and 2018 and In October 2018 the duration of the search for articles started. Only articles those are related and selected to data extracted from OSNS. Articles must be about mental health problem detection and detecting.

VI. FUTURE IMPLICATIONS

In future studies based on the mental health problem detection the researchers believe that the challenges those are outlined in the previous section can be addressed. Several aspects, such as methods, geographical locations, data extraction type, language, and multiple data sources, should be considered. Data were collected from specific geographical locations, such as Japan,Greece,China and the United States Means several researchers developed their new algorithms for mental health problem detection . The current researchers believe that creating a new data set and implementing a new method based on the localities and countries and in future that may be improve those are based on the mental health problem detection .

VII. CONCLUSION

In mental health problem detection the main purpose of this research paper is to conduct a critical assessment . In deep learning or machine learning techniques this analysis contains of the feature extraction method, data source and classifier performance method . By identifying its method of data analysis , challenges and limitations this research also investigated the appropriateness of pre-mental health detection . For detecting mental health problem detection OSNs can provide combination of the two approaches and complementary data and in future that can enhance research . Who are suffering from mental health problems or whose those conditions are observed to become serious without further treatment mental

health problem detection can probably reduces the number of those people . Based on face to face interviews , self reporting this study concludes that OSNs exhibit high potential as data sources of mental health problem detection but can never be exchanged with traditional mental health detection methods. Based on the mental health problem detection in future this study requires innovative algorithms, comprehensive adoption and computational linguistics for increase precision and accuracy.

REFERENCES

[1] Promoting Mental Health: Concepts, Emerging Evidence, Practice. World Health Org., Geneva, Switzerland, 2004.

[2] R. A. Rahman, K. Omar, S. A. M. Noah, and M. S. N. M. Danuri, ‘‘A survey on mental health detection in Online Social Network,’’ Int. J. Adv. Sci. Eng. Inf. Technol., vol. 8, nos. 2–4, pp. 1431–1436, 2018.

[3] Global Burden of Disease Study 2013 Collaborators, ‘‘Global, regional, And national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013:A systematic analysis for the global burden of disease study 2013,’’ Lancet, Vol. 386, no. 9995, pp. 743–800, 2015.

[4] The World Health Report 2001: Mental health: New Understanding, New Hope, World Health Org., Geneva, Switzerland, 2001.

[5] Management of Physical Health Conditions in Adults With Severe Mental Disorders: WHO Guidelines, World Health Org., Geneva, Switzerland, 2018.

[6] M. Marcus, M. T. Yasamy, M. van Ommeren, D. Chisholm, and S. Saxena, ‘‘Depression: A global public health concern,’’ World Health Org., Geneva, Switzerland, Paper Depression, 2012, pp. 6–8.

[7] Mental Health Action Plan 2013-2020, World Health Org., Geneva, Switzerland, 2013.

[8] C. L. M. Keyes, ‘‘Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health,’’ Amer. Psychologist, vol. 62, no. 2, pp. 95–108, 2007.

[9] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, and T.-S. Chua, ‘‘Detecting stress based on social interactions in social networks,’’ IEEE Trans. Knowl. Data Eng., vol. 29, no. 9, pp. 1820–1833,Sep. 2017.

[10] E. Garcia-Ceja, V. Osmani, and O. Mayora, ‘‘Automatic stress detection in working environments from Smartphones’ accelerometer data: A first step,’’ IEEE J. Biomed. Health Inform., vol. 20, no. 4, pp. 1053–1060, Jul. 2016.

[11] A. Sano and R. W. Picard, ‘‘Stress recognition using wearable sensors and Mobile phones,’’ in Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact., Sep. 2013, pp. 671–676.

[12] T. G. M. Vrijkotte, L. J. P. van Doornen, and E. J. C. de Geus, ‘‘Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability,’’ Hypertension, vol. 35, no. 4, pp. 880–886, Apr. 2000.

[13] G. Ali, A. Ali, F. Ali, U. Draz, F. Majeed, S. Yasin, T. Ali, and N. Haider, ‘‘Artificial neural network based ensemble approach for multicultural facial expressions analysis,’’ IEEE Access, vol. 8, pp. 134950–134963,2020.

[14] M. Kandias, D. Gritzalis, V. Stavrou, and K. Nikoloulis, ‘‘Stress level detection via OSN usage pattern and chronicity analysis: An OSINT threat intelligence module,’’ Comput. Secur., vol. 69, pp. 3–17, Aug. 2017.

[15] C.-H. Chang, E. Saravia, and Y.-S. Chen, ‘‘Subconscious crowdsourcing: A feasible data collection mechanism for mental disorder detection on social media,’’ in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), Aug. 2016, pp. 374–379.

[16] Q. Li, Y. Xue, L. Zhao, J. Jia, and L. Feng, ‘‘Analyzing and identifying teens’ stressful periods and stressor events from a microblog,’’ IEEE J. Biomed. Health Inform., vol. 21, no. 5, pp. 1434–1448, Sep. 2017.

[17] M. Thelwall, ‘‘TensiStrength: Stress and relaxation magnitude detection for social media texts,’’ Inf. Process. Manage., vol. 53, no. 1, pp. 106–121,Jan. 2017.

[18] G. A. Coppersmith, C. T. Harman, and M. H. Dredze, ‘‘Measuring post traumatic stress disorder in Twitter,’’ in Proc. 7th Int. Conf. Weblogs Social

Media (ICWSM), 2014, pp. 1–4.

[19] M. Park, C. Cha, and M. Cha, ‘‘Depressive moods of users portrayed in Twitter,’’ in Proc. ACM

SIGKDD Workshop Healthcare Infomart. (HI-KDD), 2012, pp. 1–8.

[20] H.-H. Shuai, C.-Y. Shen, D.-N. Yang, Y.-F.-C. Lan, W.-C. Lee, P. S. Yu, and M.-S. Chen, ‘‘A comprehensive study on social network mental disorders detection via online social media mining,’’ IEEE Trans. Knowl. Data Eng., vol. 30, no. 7, pp. 1212–1225, Jul. 2018.

[21] B. O’Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, ‘‘Detecting suicidality on Twitter,’’ Internet Intervent., vol. 2, no. 2, pp. 183–188, May 2015.

[22] H. Lin, J. Jia, Q. Guo, Y. Xue, J. Huang, L. Cai, and L. Feng, ‘‘Psychological stress detection from cross-media microblog data using deep sparse neural network,’’ in Proc. IEEE Int. Conf. Multimedia Expo (ICME),Jul. 2014, pp. 1–6.

[23] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh, and H. Ohsaki, ‘‘Recognizing depression from Twitter activity,’’ in Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst. (CHI), 2015, pp. 3187–3196.

[24] S. Saleem, R. Prasad, S. Vitaladevuni, M. Pacula, M. Crystal, B. Marx. Sloan, J. Vasterling,and T. Speroff, ‘‘Automatic detection of psychological distress indicators and severity assessment from online forum posts,’’in Proc. COLING, 2012, pp. 2375–2388.

[25] X. Huang, L. Zhang, D. Chiu, T. Liu, X. Li, and T. Zhu, ‘‘Detecting suicidal ideation in Chinese microblogs with psychological lexicons,’’ in Proc. IEEE 11th Int. Conf. Ubiquitous Intell. Comput. IEEE 11th Int.Conf. Autonomic Trusted Comput. IEEE 14th Int. Conf. Scalable Comput. Commun. Associated Workshops, Dec. 2014, pp. 844–849.

[26] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, ‘‘Predicting depression via social media,’’ in Proc. 7th Int. AAAI Conf. Weblogs Social Media (ICWSM), 2013, pp. 128–137.

[27] M. Deshpande and V. Rao, ‘‘Depression detection using emotion artificial intelligence,’’ in Proc. Int. Conf. Intell. Sustain. Syst. (ICISS), Dec. 2017, pp. 858–862.

[28] X. Wang, C. Zhang, and L. Sun, ‘‘An improved model for depression detection in micro-blog social network,’’ in Proc. IEEE 13th Int. Conf.Data Mining Workshops, Dec. 2013, pp. 80–87.

[29] Y. Xue, Q. Li, L. Jin, L. Feng, D. A. Clifton, and G. D. Clifford, ‘‘Detecting ado- escent psychological pressures from micro-blog,’’ in Proc. Int. Conf. Health Inf. Sci., 2014, pp. 83–94.

[30] B. Saha, T. Nguyen, D. Phung, and S. Venkatesh, ‘‘A framework for classifying online mental health-related communities with an interest in depression,’’ IEEE J. Biomed. Health Inform., vol. 20, no. 4, pp. 1008–1015, Jul. 2016.

Sangita Roy is an Assistant Professor at ECE Department of Narula Institute of Technology under WBUT. She has a teaching experience of more than sixteen years.

Purba Das is a student at ECE Department of Narula Institute of Technology under WBUT.