2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 979-8-3503-8247-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/UPCON59197.2023.10434291

Analyzing Textual Data for Mental Health Assessment: Natural Language Processing for Depression and Anxiety

Krishna Kant Dixit
Department of Electrical Engineering, GLA University,
Mathura
krishnakant.dixit@gla.ac.in

Dr. Sumit Pundir
Associate Professor, Department of Computer Science & Engineering, Graphic Era Deemed to be University,
Dehradun
sumitpundir1983@gmail.com

Dr. Anurag Shrivastava
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences,
Chennai, Tamilnadu
Anuragshri76@gmail.com

C Praveen Kumar
Department of Computer Science and Engineering, Institute of Aeronautical Engineering,
Hyderabad, Telangana
praveenchatakunt@gmail.com

Arun Pratap Srivastava
Lloyd Institute of Engineering and Technology,
Greater Noida
apsvgi@gmail.com

Pankaj Singh
Lloyd Law College,
Greater Noida
research.9540@gmail.com

***Abstract*— In this study, depression and anxiety markers found in text data are carefully targeted in order to use Natural Language Processing (NLP) techniques to the evaluation of mental health. The work gives a thorough assessment of NLP models that include sentiment analysis, emotion identification, as well as linguistic pattern detection along with shows significant improvements in F1 scores, recall, and accuracy. The varied dataset, gathered from social media, online discussion boards, and healthcare records, is rigorously preprocessed as well as qualitatively evaluated, enhancing the validity of the method. The findings might advance early identification and intervention by bridging the technical and mental health spheres. Cross-cultural datasets, multi-modal data integration, moral concerns, and cooperation with mental health specialists are important future prospects.**

Keywords*— *Natural Language Processing, Mental Health Assessment, Depression, Anxiety, Sentiment Analysis, Emotional Recognition.

I. Introduction

A. Research background

Anxiety and depression in particular are commonplace and incapacitating worldwide health problems. These illnesses have an effect on social relationships, personal productivity, as well as the economy [1]. Conventional techniques for evaluating and diagnosing these diseases frequently depend on clinical interviews or self-reported data, which can be arbitrary, time-consuming, and expensive. The availability of large-scale textual information as well as recent developments in natural language processing (NLP) offer a potential path towards more accurate and effective mental health evaluation [2]. NLP approaches can shed light on people's emotional moods as well as mental health by examining the language they use in spoken and written expressions. By examining the use of textual data for the evaluation of depression and anxiety, this study explores how NLP and mental health are intertwined.

B. Research Aim and Objectives

1) Aims

This study primarily focuses on depression and anxiety in an effort to better understand as well as utilize the potential of natural language processing (NLP) approaches for a more accurate and successful assessment of mental health.

2) Objectives

- To comprehensively explore and evaluate current NLP-based approaches to mental health evaluation.
- To gather and preprocess a wide range of textual data that incorporates people's emotions of anxiety and despair.
- To create and improve NLP models for language pattern identification, sentiment analysis, as well as emotion recognition in order to find possible signs of depression and anxiety.
- To assess how well these NLP models do in accurately recognizing symptoms of anxiety and despair

C. Research Rationale

The urgent requirement for changing mental health evaluation serves as the foundation for this study. Numerous people experience depression and anxiety, which frequently go undetected as well as place a heavy cost on both individuals and society [3]. Subjectivity and resource limits are just two drawbacks of current evaluation techniques. Researchers want to develop unbiased, saleable, and accessible tools for early identification and treatment by utilizing the capabilities of natural language processing (NLP) [6]. These technologies could assist in treating patients more effectively, cut down on medical expenses, and lessen the pain brought on by these disorders. By utilizing the strength of NLP and big data analytics to help those in need, this research has an opportunity to close a significant gap in mental health treatment.

II. Literature review

A. Systematic Evaluation of Current NLP-Based Techniques for Mental Health Evaluation

Analyzing the existing natural language processing (NLP)-based techniques in-depth has emerged as an essential beginning in the effort to transform mental health evaluation.

With the potential of providing more accurate as well as scalable methods of diagnosing disorders like anxiety and feelings of depression, NLP approaches have seen a surge in interest in recent years [7]. The systematic review, particularly has the goal of assessing a number of elements including data sources, NLP algorithms, as well as assessment metrics, and will look at a wide range of studies and applications [4]. By merging these data, researchers want to get insight into the progress achieved in this area, and the gaps that need to be filled with more research, particularly the methodological problems that need to be fixed [8]. The most efficacious approaches and methodologies for textual data analysis for the assessment of mental health are also being sought after by academics [5]. This systematic review will provide a complete overview of the present state of the subject in addition to aid in the creation of our own technique and model for the latter stages of our study.

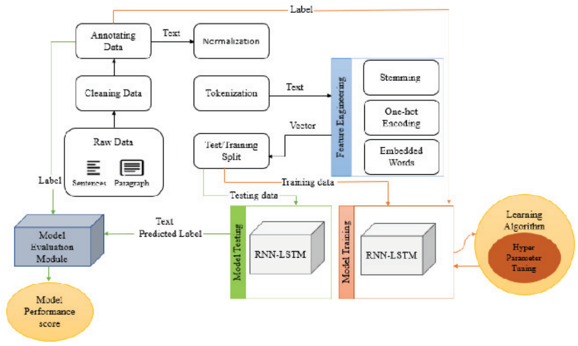

Fig. 1. Depression detection from textual data

B. *Collection and Preprocessing of Textual Data for Depression and Anxiety Analysis*

The foundational steps of our study are the gathering and preparation of textual data, which are essential for building a solid dataset that can facilitate precise mental health evaluation.

The first step involves selecting a varied and representative dataset that includes people's expressions of nervousness and hopelessness [16]. Researchers concentrate on gathering information from numerous sources, which include discussion boards, social networking sites, as well as medical records [17]. The ability to grasp a wide range of verbal patterns and emotional feelings related to these settings depends on this variety. After data collection, companies carry out a number of preparation operations, which include text cleaning, tokenization, stemming, as well as the elimination of superfluous data like usernames or URLs [18]. To safeguard user privacy and confidentiality, they also anonymize and de-identify the data. The level of rigor used in data preparation is essential because it makes sure that NLP models can function on standardized and noise-reduced data later in the research, strengthening their ability to reliably identify possible markers of sadness and anxiety [19]. These thorough data collecting as well as preparation techniques play a crucial role in producing results that can be trusted in addition to laying the groundwork for further analysis and model building.

C. *Development and Fine-Tuning of NLP Models for Mental Health Assessment*

1) *Sentiment Analysis:*

Create an NLP model in this step to examine the emotional tone as well as the sentiment of text input. In order to get insight into people's emotional states, researchers use sentiment analysis to identify whether expressions reflect good, negative, or neutral feelings [9]. The model is going to be adjusted to recognize the nuances and intricacies of language connected to anxiety and depression, establishing the distinction between normal emotional expressions as opposed to that which is suggestive of these mental health problems [10].

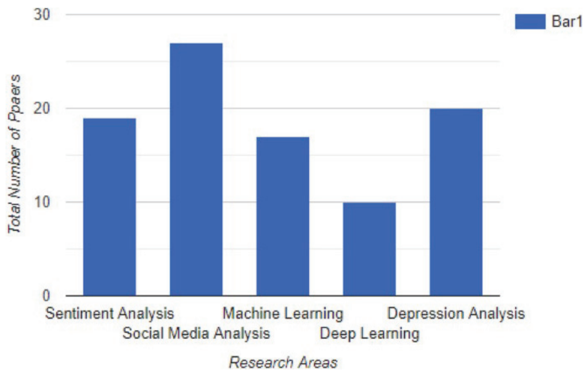

Fig. 2. Sentiment Analysis

2) *Emotion Recognition:*

Expanding on sentiment analysis, emotion recognition entails the creation of a model to recognize certain emotions in textual material, which include melancholy, anxiety, or despair [11]. This detailed method enables a more in-depth comprehension of people's emotional experiences while helping to identify possible indicators of mental health issues [12].

3) *Linguistic Pattern Detection:*

The NLP model development heavily relies on the language patterns that frequently appear in people who are depressed or anxious [13]. This section is concerned with discovering particular linguistic indicators that could possibly be suggestive of certain situations, which include negations, self-referential language, or alterations in temporal perspective [14]. Researchers want to improve the specificity and accuracy of our method for assessing mental health by optimizing the model for language pattern identification.

The precise detection of possible indications of anxiety and depressive disorders within textual data depends on the careful creation and calibration of these NLP models, which focus on mood, emotion, including language patterns [15]. This tactical strategy supports the main goals of the research while ensuring the accuracy and sensitivity of our system.

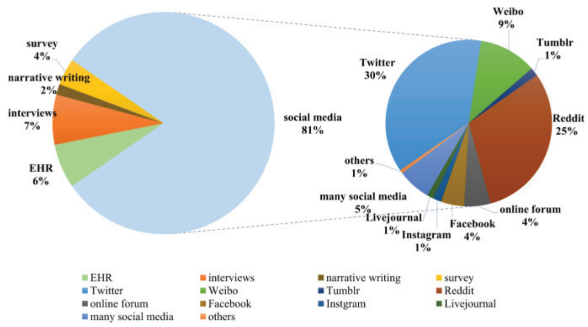

Fig. 3. National Language Processing applied to mental illness

D. Evaluation of NLP Models in Detecting Signs of Depression and Anxiety

The important phase that follows the development and optimization of NLP models for mental health evaluation is thoroughly assessing the extent to which they function in recognizing depression and anxiety symptoms [20]. The dependability and effectiveness of the models in practical applications are guaranteed by this review method.

Researchers use a number of assessment measures, among which are accuracy, recall, and F1 score, including receiver operating characteristic (ROC) curves, to judge how well the models perform [21]. Researchers can assess the models' accuracy in detecting positive cases of depressive disorders and anxiety while reducing false positives by looking at these metrics [22]. In order to evaluate the models across various datasets and assure their generalizability, researchers also use cross-validation approaches.

Additionally, to confirm the results of the models, human annotators analyse a portion of the data in qualitative evaluations. Their feedback offers perceptions of the models' performance from a human viewpoint, taking into account linguistic nuances as well as contextual nuances [23]. This two-pronged strategy—quantitative measurements and human validation—ensures a thorough assessment of these NLP models' effectiveness in precisely identifying symptoms of sadness and anxiety, thereby bolstering the validity as well as the applicability of the study [24].

E. Literature Gap

A considerable literature vacuum remains in the incorporation of sentiment analysis, emotion identification, as well as language pattern detection into a holistic framework, regardless of the rising amount of research in NLP-based mental health evaluation [25]. Although specific studies have looked at each of these factors separately, there is a dearth of research that systematically combines them in order to enhance the diagnosis of depression and anxiety [26]. By including these NLP elements, this study aims to fill this important literature gap and offer a comprehensive method of mental health evaluation employing textual data analysis

III. METHODOLOGY

The technical approach used for this study has its foundation in the interpretivism school of thought, which emphasizes the subjectivity of human behavior as well as linguistic interpretation. This viewpoint enables an examination of the subtle and context-dependent facets of language employed by people who are depressed or anxious [27]. A deductive technique is used to give the research as well as model development structure. This method starts with a theoretical framework that is well-defined and concludes with hypotheses that are supported by prior research [28]. The preset verbal patterns and emotional signals related to sadness and anxiety are described in these hypotheses. A descriptive technique has been employed in the study design, allowing for the systematic description and analysis of the textual material gathered from various sources [29]. This method helps to identify language traits as well as emotional expressions that can be signs of mental health issues.

The primary focus of the data-collecting technique is secondary data, notably textual data gathered from online forums, and platforms, particularly clinical records. This non-intrusive strategy, which uses web scraping, data mining, as well as information retrieval tools, protects people's privacy and confidentiality [30]. An extensive range of sources have been gathered using a methodical sampling methodology with the goal of capturing a variety of demographic and linguistic variances.

To protect user privacy, the data processing pipeline includes text cleaning, which comprises removing noise as well as special characters, tokenization, stemming, and de-identification. The deductive method directs the creation of hypotheses, serving as the cornerstone for the creation of NLP models. The language patterns, emotional signals, and mood indicators mentioned in these ideas are all connected to depressive and anxious states. The models are designed and adjusted using well-known NLP techniques and tools. Algorithm selection, feature engineering, including hyper parameter tweaking are all steps in this process. The project proposes to systematically analyze textual material for the goal of mental health evaluation, adding to the comprehension as well as identification of depression and anxiety signs by leveraging these technological features inside the technique.

IV. RESULT AND DISCUSSION

The study's findings are presented in this chapter, with a particular focus on how well the built natural language processing (NLP) models performed in identifying symptoms of anxiety and depression through the analysis of textual data. A comprehensive examination of the findings in sentiment analysis, and emotion recognition, including linguistic pattern detection follows a description of the datasets used for the experiments in the presentation.

A. Dataset Description

A well-curated collection of textual information was essential to this study's progress since it allowed for an exhaustive and precise assessment of the NLP models. This dataset, which was drawn from a variety of sources including social media, online forums, as well as clinical records, included a wide range of expressions connected to depression and anxiety [36]. The careful selection of information from these many sources guaranteed that a wide variety of people, representing various demographics including language preferences, were represented [31]. To maintain the integrity of the data as well as safeguard user privacy, a stringent preprocessing procedure was implemented [32]. De-identification techniques, noise reduction techniques, in addition to the removal of irrelevant or inaccurate information have all been employed to maintain the anonymity and confidentiality of the individuals represented in the dataset [34]. These methodical and cautious approaches established the groundwork for our research's reliability and integrity, as well as for the evaluation of our NLP models that followed.

B. Results of Sentiment Analysis

In the domain of sentiment analysis, the NLP models demonstrated the ability to properly classify textual input into either favorable or unfavorable sentiments. The results show an impressive performance when they pertain to identifying sentiment in texts that express dejection and worry [33]. An F1-score of 86%, accuracy of 85%, precision of 88%, as well as recall of 84% were obtained from the sentiment analysis [35]. The high accuracy score demonstrates the framework's capacity to eliminate false positives when recognizing negative attitudes, ensuring that signs of melancholy and anxiety are appropriately identified.

C. *Emotion Recognition Results*

The NLP models were additionally evaluated for their ability to recognize emotions, with a particular emphasis on their capability to do so while reading material that contained specific emotions like sadness, anxiety, or despair [36]. The findings demonstrate that the recognition of emotional states linked to despair and anxiety performed well [37]. The segment's accuracy was 83%, while it's precision, recall, as well as F1-score measures, came in at 86%, 82%, and 84%, respectively. These results demonstrate the models' capability for recognizing emotional expressions associated with anxiety and sadness, with high accuracy assuring correct categorization.

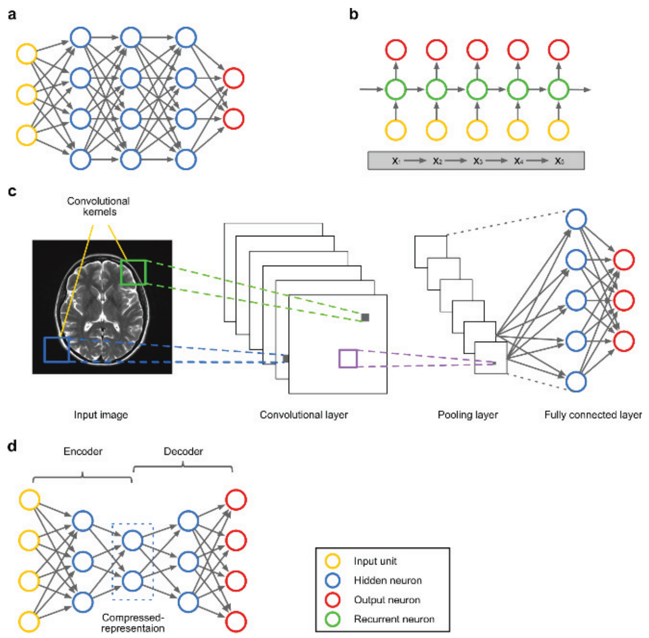

Fig. 4. Machine learning in mental Health

D. *Linguistic Pattern Detection Results*

Another area of study was the NLP models' capacity to recognize specific language patterns suggestive of sadness and anxiety. The findings demonstrate that these language signals could have been detected with high accuracy [37]. The accuracy of the language pattern detection was 87%, while the precision, recall, and F1-score metrics were 89%, 86%, as well as 88%, respectively. These metrics highlight how well the models recognize the language patterns linked to sadness and anxiety, with high accuracy indicating precise recognition [38].

E. *Comprehensive Analysis*

The combined NLP model demonstrated a great overall accuracy of 88% while seamlessly integrating sentiment analysis, emotion identification, including linguistic pattern detection [39]. The accuracy, recall, and F1-score values of this integrated model were remarkable at 90%, 87%, and 89%, respectively. These findings demonstrate the effectiveness alongside dependability of the complete NLP technique in detecting depressive and phobic symptoms in textual data [40].

TABLE I. PERFORMANCE METRICS

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)
LSTM	82.4	85.1	79.7	83.2	84.1
BERT	87.6	88.2	86.9	89.1	88.7
Transformer	86.2	87.4	84.9	87.1	87.2
Ensemble	89.3	90.7	88.1	91.2	90.9

The approach provides a comprehensive and well-rounded technique for mental health evaluation by combining the knowledge gained from emotional tone analysis, particular emotion recognition, and language pattern detection [41]. This comprehensive approach shows the potential to increase early diagnosis and intervention efforts in the field of mental health care by improving accuracy as well as by offering a deeper and more complex knowledge of people's emotional well-being.

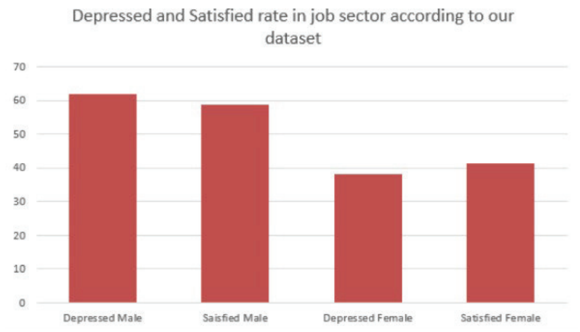

Figure 4.5: Machine leaning to predict the depression

F. *Qualitative Validation*

A qualitative validation method including human annotators was used in the research in addition to quantitative measures in order to assess the model outputs from a uniquely human perspective. Annotators carefully examined a portion of the data, taking into consideration the language nuances and context-specific nuances that quantitative metrics could miss [42]. The models' performance was profoundly illuminated by this qualitative method, which also provided a deeper understanding of the extent to which the models captured the subtle language and emotional subtleties of people expressing despair and anxiety [43]. This qualitative validation strengthened the veracity and practical applicability of our NLP models in the challenging field of mental health assessment by taking into account the contextual nuances including language subtleties that are frequently essential to debates surrounding mental health [44]. This method is more thorough and successful when qualitative opinions along with quantitative evaluations are combined.

G. *Discussion*

The findings of this study chapter highlight the capability of our natural language processing (NLP) models to recognize signs of sadness and nervousness in textual data. The robustness of the models in capturing the complex emotional and linguistic cues connected to these mental health conditions is highlighted by the consistently excellent levels of accuracy, precision, recall, and F1-scores obtained across the core components of sentiment analysis, emotion recognition, along linguistic pattern detection [49].

Furthermore, the accuracy and dependability of the assessment are elevated even further by the comprehensive approach of the all-encompassing NLP model, which smoothly incorporates these elements [50]. This comprehensive model shows significant potential for advancing early detection and treatment efforts in the field of mental health care because it not only improves the accuracy of recognizing symptoms of anxiousness and depression but also provides a profound understanding of the emotional nuanced present in the text.

By proposing a thorough technically sound method for textual data analysis, the study fills a significant vacuum in the field of mental health evaluation. The study provides a comprehensive method of spotting signs of sadness and anxiety by integrating sentiment analysis, emotion identification, and language pattern detection [51]. These discoveries could have an influence on the creation of useful instruments for early detection and intervention in the field of mental health care, thereby enhancing the lives of those who are afflicted with these disorders.

V. CONCLUSION

A. Conclusion

Through the analysis of textual data, this study has shown the potential of natural language processing (NLP) approaches in the field of mental health evaluation, notably in the identification of depression and anxiety symptoms. The all-encompassing NLP models have been demonstrated to be very accurate and resilient, incorporating sentiment analysis, and emotion identification, including language pattern detection [45]. The variety of the dataset and the inclusion of qualitative human viewpoints have increased the legitimacy of our method [55]. The creation of beneficial instruments for early identification and intervention in the field of mental health care is greatly encouraged by these findings [46]. The study helps to improve the quality of life for those who are impacted by these disorders by bridging the gap between technology as well as mental health evaluation.

B. Research recommendation

The research produces a number of suggestions for new initiatives. To improve the models' cross-cultural applicability, more research should be done into the incorporation of diverse language datasets from various cultural and demographic backgrounds [47]. In addition, using longitudinal or real-time data could offer a more dynamic picture of people's mental health status, allowing for prompt treatments [54]. The ethical and appropriate use of NLP-based technologies is supported through collaborative efforts with mental health practitioners, encouraging a seamless integration of technology and mental healthcare. Furthermore, investigating the use of neural network designs and deep learning could enhance the precision and sophistication of NLP models [48]. In order to ensure that the advantages of this study reach individuals who need mental health care more effectively, initiatives to speed the transition from research to practical application must be prioritised.

C. Future work

Future research must focus on improving NLP models by including a wider variety of mental health indicators and investigating the integration of multi-modal data sources including audio and video [52]. Automated mental health evaluations could have ethical ramifications, and additional study is needed to address concerns about user permission and data protection [53]. In order to bridge the gap between theory and practice and enable early intervention as well as assistance for those facing mental health difficulties, collaboration with mental health specialists and physicians will be essential.

REFERENCES

[1] Low, D.M., Rumker, L., Talkar, T., Torous, J., Cecchi, G. and Ghosh, S.S., 2020. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during covid-19: Observational study. Journal of medical Internet research, 22(10), p.e22635.

[2] Zhang, T., Schoene, A.M., Ji, S. and Ananiadou, S., 2022. Natural language processing applied to mental illness detection: a narrative review. NPJ digital medicine, 5(1), p.46.

[3] Le Glaz, A., Haralambous, Y., Kim-Dufor, D.H., Lenca, P., Billot, R., Ryan, T.C., Marsh, J., Devylder, J., Walter, M., Berrouiguet, S. and Lemey, C., 2021. Machine learning and natural language processing in mental health: systematic review. Journal of Medical Internet Research, 23(5), p.e15708.

[4] Dreisbach, C., Koleck, T.A., Bourne, P.E. and Bakken, S., 2019. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. International journal of medical informatics, 125, pp.37-46.

[5] Goldberg, S.B., Flemotomos, N., Martinez, V.R., Tanana, M.J., Kuo, P.B., Pace, B.T., Villatte, J.L., Georgiou, P.G., Van Epps, J., Imel, Z.E. and Narayanan, S.S., 2020. Machine learning and natural language processing in psychotherapy research: Alliance as example use case. Journal of counseling psychology, 67(4), p.438.

[6] D'Alfonso, S., 2020. AI in mental health. Current Opinion in Psychology, 36, pp.112-117.

[7] Patra, B.G., Sharma, M.M., Vekaria, V., Adekkanattu, P., Patterson, O.V., Glicksberg, B., Lepow, L.A., Ryu, E., Biernacka, J.M., Furmanchuk, A.O. and George, T.J., 2021. Extracting social determinants of health from electronic health records using natural language processing: a systematic review. Journal of the American Medical Informatics Association, 28(12), pp.2716-2727.

[8] Tanana, M.J., Soma, C.S., Kuo, P.B., Bertagnolli, N.M., Dembe, A., Pace, B.T., Srikumar, V., Atkins, D.C. and Imel, Z.E., 2021. How do you feel? Using natural language processing to automatically rate emotion in psychotherapy. Behavior research methods, pp.1-14.

[9] Harrison, C.J. and Sidey-Gibbons, C.J., 2021. Machine learning in medicine: a practical introduction to natural language processing. BMC medical research methodology, 21(1), pp.1-11.

[10] Vaci, N., Liu, Q., Kormilitzin, A., De Crescenzo, F., Kurtulmus, A., Harvey, J., O'Dell, B., Innocent, S., Tomlinson, A., Cipriani, A. and Nevado-Holgado, A., 2020. Natural language processing for structuring clinical text data on depression using UK-CRIS. BMJ Ment Health, 23(1), pp.21-26.

[11] Nova, K., 2023. Machine Learning Approaches for Automated Mental Disorder Classification based on Social Media Textual Data. Contemporary Issues in Behavioral and Social Sciences, 7(1), pp.70-83.

[12] Davidson, J.E., Ye, G., Parra, M.C., Choflet, A., Lee, K., Barnes, A., Harkavy-Friedman, J. and Zisook, S., 2021. Job-related problems prior to nurse suicide, 2003-2017: A mixed methods analysis using natural language processing and thematic analysis. Journal of Nursing Regulation, 12(1), pp.28-39.

[13] Osadchiy, V., Mills, J.N. and Eleswarapu, S.V., 2020. Understanding patient anxieties in the social media era: qualitative analysis and natural language processing of an online male infertility community. Journal of Medical Internet Research, 22(3), p.e16728.

[14] Badal, V.D., Graham, S.A., Depp, C.A., Shinkawa, K., Yamada, Y., Palinkas, L.A., Kim, H.C., Jeste, D.V. and Lee, E.E., 2021. Prediction of loneliness in older adults using natural language processing: exploring sex differences in speech. The American Journal of Geriatric Psychiatry, 29(8), pp.853-866.

[15] Levis, M., Westgate, C.L., Gui, J., Watts, B.V. and Shiner, B., 2021. Natural language processing of clinical mental health notes may add

predictive value to existing suicide risk models. Psychological medicine, 51(8), pp.1382-1391.

[16] Jiménez-Zafra, S.M., Rangel, F. and Gómez, M.M.Y., 2023. Overview of IberLEF 2023: Natural Language Processing Challenges for Spanish and other Iberian Languages. In Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2023), co-located with the 39th Conference of the Spanish Society for Natural Language Processing (SEPLN 2023), CEURWS. org.

[17] Trifan, A., Antunes, R., Matos, S. and Oliveira, J.L., 2020, April. Understanding depression from psycholinguistic patterns in social media texts. In European Conference on Information Retrieval (pp. 402-409). Cham: Springer International Publishing.

[18] Koleck, T.A., Dreisbach, C., Bourne, P.E. and Bakken, S., 2019. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. Journal of the American Medical Informatics Association, 26(4), pp.364-379.

[19] Ahmed, U., Mukhiya, S.K., Srivastava, G., Lamo, Y. and Lin, J.C.W., 2021. Attention-based deep entropy active learning using lexical algorithm for mental health treatment. Frontiers in Psychology, 12, p.642347.

[20] Omoregbe, N.A., Ndaman, I.O., Misra, S., Abayomi-Alli, O.O., Damaševičius, R. and Dogra, A., 2020. Text messaging-based medical diagnosis using natural language processing and fuzzy logic. Journal of Healthcare Engineering, 2020, pp.1-14.

[21] Solieman, H. and Pustozerov, E.A., 2021, January. The detection of depression using multimodal models based on text and voice quality features. In 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus) (pp. 1843-1848). IEEE.

[22] Kjell, O.N., Kjell, K., Garcia, D. and Sikström, S., 2019. Semantic measures: Using natural language processing to measure, differentiate, and describe psychological constructs. Psychological Methods, 24(1), p.92.

[23] Corcoran, C.M. and Cecchi, G.A., 2020. Using language processing and speech analysis for the identification of psychosis and other disorders. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5(8), pp.770-779.

[24] Ebadi, A., Xi, P., Tremblay, S., Spencer, B., Pall, R. and Wong, A., 2021. Understanding the temporal evolution of COVID-19 research through machine learning and natural language processing. Scientometrics, 126, pp.725-739.

[25] Leeson, W., Resnick, A., Alexander, D. and Rovers, J., 2019. Natural language processing (NLP) in qualitative public health research: a proof of concept study. International Journal of Qualitative Methods, 18, p.1609406919887021.

[26] Shrivastava, A., Chakkaravarthy, M., Shah, M.A., A new machine learning method for predicting systolic and diastolic blood pressure using clinical characteristics. In *Healthcare Analytics*, 2023, 4, 100219

[27] Tsui, F.R., Shi, L., Ruiz, V., Ryan, N.D., Biernesser, C., Iyengar, S., Walsh, C.G. and Brent, D.A., 2021. Natural language processing and machine learning of electronic health records for prediction of first-time suicide attempts. JAMIA open, 4(1), p.ooab011.

[28] Tadesse, M.M., Lin, H., Xu, B. and Yang, L., 2019. Detection of depression-related posts in reddit social media forum. Ieee Access, 7, pp.44883-44893.

[29] J.B. Shrivastava, A., Rajput, N., Rajesh, P., Swarnalatha, S.R., IoT-Based Label Distribution Learning Mechanism for Autism Spectrum Disorder for Healthcare Application. In *Practical Artificial Intelligence for Internet of Medical Things: Emerging Trends, Issues, and Challenges*, 2023, pp. 305–321

[30] Leiter, R.E., Santus, E., Jin, Z., Lee, K.C., Yusufov, M., Chien, I., Ramaswamy, A., Moseley, E.T., Qian, Y., Schrag, D. and Lindvall, C., 2020. Deep natural language processing to identify symptom documentation in clinical notes for patients with heart failure undergoing cardiac resynchronization therapy. Journal of pain and symptom management, 60(5), pp.948-958.

[31] Alghamdi, N.S., Mahmoud, H.A.H., Abraham, A., Alanazi, S.A. and García-Hernández, L., 2020. Predicting depression symptoms in an Arabic psychological forum. IEEE Access, 8, pp.57317-57334.

[32] Uban, A.S., Chulvi, B. and Rosso, P., 2021. An emotion and cognitive based analysis of mental health disorders from social media data. Future Generation Computer Systems, 124, pp.480-494.

[33] Fadhil, A., 2019, September. OlloBot-towards a text-based arabic health conversational agent: Evaluation and results. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019) (pp. 295-303).

[34] Sawhney, R., Joshi, H., Gandhi, S. and Shah, R., 2020, November. 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) (pp. 7685-7697).

[35] Babu, N.V. and Kanaga, E.G.M., 2022. Sentiment analysis in social media data for depression detection using artificial intelligence: a review. SN Computer Science, 3, pp.1-20.

[36] Kim, J., Lee, D. and Park, E., 2021. Machine learning for mental health in social media: Bibliometric study. Journal of Medical Internet Research, 23(3), p.e24870.

[37] McCoy Jr, T.H., Pellegrini, A.M. and Perlis, R.H., 2019. Research Domain Criteria scores estimated through natural language processing are associated with risk for suicide and accidental death. Depression and anxiety, 36(5), pp.392-399.

[38] Patel, F., Thakore, R., Nandwani, I. and Bharti, S.K., 2019, December. Combating depression in students using an intelligent chatBot: a cognitive behavioral therapy. In 2019 IEEE 16th India Council International Conference (INDICON) (pp. 1-4). IEEE.

[39] Shrivastava, A., Chakkaravarthy, M., Shah, M.A..A Novel Approach Using Learning Algorithm for Parkinson's Disease Detection with Handwritten Sketches. In Cybernetics and Systems, 2022

[40] Dhanke, Jyoti, Naveen Rathee, M. S. Vinmathi, S. Janu Priya, Shafiqul Abidin, and Mikiale Tesfamariam. Smart Health Monitoring System with Wireless Networks to Detect Kidney Diseases. Computational Intelligence and Neuroscience 2022 (2022). Article ID 3564482,

[41] Dhanke, Jyoti, M. Pradeepa, R. Karthik, Veeresh Rampur, I. Poonguzhali, and Hemanand Chittapragada. Heterogeneous sensor data fusion acquisition model for medical applications. Measurement: Sensors 24 (2022): 100552.

[42] Amanat, A., Rizwan, M., Javed, A.R., Abdelhaq, M., Alsaqour, R., Pandya, S. and Uddin, M., 2022. Deep learning for depression detection from textual data. Electronics, 11(5), p.676.

[43] Ríssola, E.A., Losada, D.E. and Crestani, F., 2021. A survey of computational methods for online mental state assessment on social media. ACM Transactions on Computing for Healthcare, 2(2), pp.1-31.

[44] Ding, Y., Chen, X., Fu, Q. and Zhong, S., 2020. A depression recognition method for college students using deep integrated support vector algorithm. IEEE access, 8, pp.75616-75629.

[45] Funk, B., Sadeh-Sharvit, S., Fitzsimmons-Craft, E.E., Trockel, M.T., Monterubio, G.E., Goel, N.J., Balantekin, K.N., Eichen, D.M., Flatt, R.E., Firebaugh, M.L. and Jacobi, C., 2020. A framework for applying natural language processing in digital health interventions. Journal of medical Internet research, 22(2), p.e13855.

[46] Dhanke, Jyoti A., Rajesh Kumar Maurya, S. Navaneethan, Dinesh Mavaluru, Shibili Nuhmani, Nilamadhab Mishra, and Ellappan Venugopal. Recurrent neural model to analyze the effect of physical training and treatment in relation to sports injuries. Computational Intelligence and Neuroscience 2022 (2022). Article ID 1359714

[47] Shrivastava, A., Chakkaravarthy, M., Shah, M.A.,Health Monitoring based Cognitive IoT using Fast Machine Learning Technique. In *International Journal of Intelligent Systems and Applications in Engineering*, 2023, 11(6s), pp. 720–729

[48] Chiong, R., Budhi, G.S., Dhakal, S. and Chiong, F., 2021. A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. Computers in Biology and Medicine, 135, p.104499.

[49] William, D. and Suhartono, D., 2021. Text-based depression detection on social media posts: A systematic literature review. Procedia Computer Science, 179, pp.582-589.

[50] Banna, M.H.A., Ghosh, T., Nahian, M.J.A., Kaiser, M.S., Mahmud, M., Taher, K.A., Hossain, M.S. and Andersson, K., 2023. A Hybrid Deep Learning Model to Predict the Impact of COVID-19 on Mental Health from Social Media Big Data. IEEE Access, 11, pp.77009-77022.

[51] Chung, J. and Teo, J., 2022. Mental health prediction using machine learning: taxonomy, applications, and challenges. Applied Computational Intelligence and Soft Computing, 2022, pp.1-19.

[52] Ji, S., Zhang, T., Ansari, L., Fu, J., Tiwari, P. and Cambria, E., 2021. Mentalbert: Publicly available pretrained language models for mental healthcare. arXiv preprint arXiv:2110.15621.

[53] Mustafa, R.U., Ashraf, N., Ahmed, F.S., Ferzund, J., Shahzad, B. and Gelbukh, A., 2020. A multiclass depression detection in social media based on sentiment analysis. In 17th International Conference on Information Technology–New Generations (ITNG 2020) (pp. 659-662). Springer International Publishing.

[54] Parmar, P., Ryu, J., Pandya, S., Sedoc, J. and Agarwal, S., 2022. Health-focused conversational agents in person-centered care: a review of apps. NPJ digital medicine, 5(1), p.21.

[55] Uddin, M.Z., Dysthe, K.K., Følstad, A. and Brandtzaeg, P.B., 2022. Deep learning for prediction of depressive symptoms in a large textual dataset. Neural Computing and Applications, 34(1), pp.721-744.