

Summary

Major depressive disorder constitutes the third leading cause of disability globally, with an estimated 5% of adults suffering from depression worldwide. However, different treatment trajectories and efficacy have been observed for different subgroups of patients. Traditionally, mental disorders have been considered as a collection of empirically related symptoms for unknown reasons. The heterogeneity of the disorder has been demonstrated mathematically; following the DSM-IV criteria for major depressive disorder 1497 different combinations have been calculated. These findings point to the need to achieve better diagnoses that allow applying more effective treatments according to the characteristics of the various types of patients.

The advent of new technologies such as artificial intelligence has provided a solution in other healthcare fields such as medicine. Natural language processing (NLP) is a branch within the field of artificial intelligence that employs computational techniques to learn, understand, and produce human language. This technology has proven useful in detecting constructs in unstructured clinical texts.

In view of the great potential that NLP has for analyzing linguistic data, the aim of this work is to conduct a systematic review using PRISMA methodology to find out how this technology is being used to diagnose and classify people suffering from depressive symptoms using clinical data. A search was conducted using a series of terms related to the research question in the PubMed database. After applying the PRISMA method, five articles were selected according to established criteria. The results of this systematic review suggest that the NLP technique has the potential to become a useful tool for health professionals in assessing and diagnosing depressive symptoms.

Key words: depression, diagnosis, natural language processing, clinical psychology, artificial intelligence.

Resumen

El trastorno por depresión mayor constituye la tercera causa de discapacidad a nivel global, se estima que el 5% de los adultos sufren depresión a nivel mundial. No obstante, se han observado diferentes trayectorias y eficacia de los tratamientos en función de diferentes subgrupos de pacientes. Tradicionalmente, los trastornos mentales se han considerado como una colección de síntomas empíricamente relacionados por razones desconocidas. La heterogeneidad del trastorno se ha demostrado matemáticamente; siguiendo los criterios del DSM-IV para el trastorno de depresión mayor se han calculado 1497 combinaciones distintas. Estos hallazgos apuntan a la necesidad de lograr mejores diagnósticos que permitan aplicar tratamientos más eficaces en función de las características de los diversos tipos de pacientes.

La llegada de nuevas tecnologías como la inteligencia artificial ha supuesto una solución en otros campos de la salud como la medicina. El procesamiento del lenguaje natural (NLP) es una rama dentro del campo de la inteligencia artificial que emplea técnicas computacionales para aprender, entender y producir lenguaje humano. Esta tecnología ha demostrado ser útil en la detección de constructos en textos clínicos desestructurados.

En vista del gran potencial que tiene el NLP para analizar los datos lingüísticos, se ha marcado como objetivo hacer una revisión sistemática utilizando la metodología PRISMA para averiguar de qué manera se está empleando esta tecnología para diagnosticar y clasificar a las personas que padecen síntomas depresivos utilizando datos clínicos. Se eligió la base de datos PubMed y se realizó una búsqueda con una serie de términos relacionados con la pregunta de investigación. Después de aplicar el método PRISMA, quedaron seleccionados cinco artículos de acuerdo a criterios establecidos. Los resultados de esta revisión sistemática apuntan a que la técnica de NLP tiene potencial para convertirse en una herramienta útil para los profesionales de la salud a la hora de evaluar y diagnosticar los síntomas depresivos.

Palabras clave: depresión, diagnóstico, procesamiento del lenguaje natural, psicología clínica, inteligencia artificial.



Natural language processing for the detection and classification of depressive patients: a systematic literature review

Valeria Marine Gonzalez Guerrero

Introduction

Major depressive disorder is the third leading cause of disability globally, with an estimated 5% of adults suffering from depression worldwide (WHO, 2021). Despite being extremely common, diagnostic remain unclear because of great heterogeneity of the disorder. (Ermers et al., 2020). This divergency has been demonstrated mathematically; following the DSM-IV criteria for major depressive disorder 1497 different combinations have been calculated (Østergaard et al., 2011). These findings point to the need to achieve better diagnoses that allow applying more effective treatments according to the characteristics of the various types of patients (Ermers et al., 2020). The advent of new technologies such as artificial intelligence has provided a solution in other healthcare fields such as medicine (Dwyer et al., 2018). Natural language processing (NLP) employs computational techniques to learn, understand, and produce human language (Hirschberg & Manning, 2015). This technology has proven useful in detecting constructs in unstructured clinical texts (Stewart & Velupillai, 2021). In view of the great potential NLP has for analyzing linguistic data, a systematic review using the PRISMA methodology was set as a goal to find out how this method is being used to diagnose and classify people suffering from depressive symptoms using clinical data.

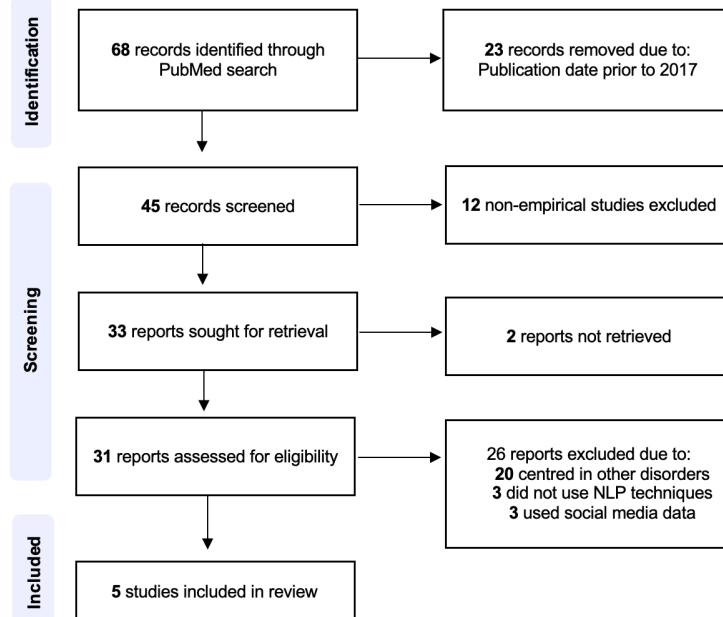
Method

Systematic literature review following PRISMA guidelines of empirical studies published between 2017 and 2022.

Search terms: *(natural language processing AND (depress*)) AND (diagnosis or prognosis)*

PubMed database

Figure 1: PRISMA flowchart



Results

Table 1: Characteristics of the reviewed studies

Authors	Procedure	Results
Adekanattu, Sholle, DeFerio et al., 2018	Clinical texts PHQ-9 Expert consensus	NLP technique identified that 53% of a sample of 1583 patients who had PHQ-9 score > 10 had not received a diagnosis of major depression.
Xu, Yang, Chakraborty et al., 2019	Semi-structured interview NSA-16 Audio recording and transcription	NLP technique found differences between the control group and the depression group with an accuracy around 72% when analyzing the recordings and conversational features.
Dai et al., 2021	Clinical texts Expert consensus	Novel pre-trained techniques had better performance than the others evaluated techniques for analyzing small and imbalanced data.
Kjell, Johnsson & Sikström, 2021	Answers to open-ended questions PHQ-9	The NLP-based "QCLA" method succeeds in capturing the diagnostic criteria for major depressive disorder. Valence may be contributing to explaining the model's prediction of the depression construct.
Burkhardt et al., 2021	Online therapy text messages BADs PHQ-9	The NLP technique was able to find differences according to the severity of the patients' depression. Measurement of "behavioral activation" through meaningful constructs with specific linguistic indicators. Association to changes in depressive symptomatology over time.

Discussion and conclusions

In general terms, there is a great heterogeneity in the reviewed studies. In common, it stands out the use of the PHQ-9 to measure depressive symptoms and the small size of the clinical samples. Future studies could focus on longitudinally studying the progression of patients through the process of remission and relapse, as well as the severity of the disorder. It is also important to replicate these studies with larger sample sizes.

The results of the studies tend to show that the NLP technique is able to classify and differentiate depressive symptoms with high accuracy. Therefore, in the future, these automation systems could represent a helpful tool for health professionals when assessing and diagnosing major depressive disorder.

References

- Adekanattu, P., Sholle, E. T., DeFerio, J., Pathak, J., Johnson, S. B., & Champion, T. R. (2018). Ascertaining Depression Severity by Extracting Patient Health Questionnaire-9 (PHQ-9) Scores from Clinical Notes. *AMIA Annual Symposium Proceedings, 2018*, 147.
- Burkhardt, H. A., Alexopoulos, G. S., Pullmann, M. D., Hull, T. D., Areán, P. A., & Cohen, T. (2021). Behavioral Activation and Depression Symptomatology: Longitudinal Assessment of Linguistic Indicators in Text-Based Therapy Sessions. *Journal of Medical Internet Research, 23*(7). <https://doi.org/10.2196/28244>
- Dai, H. J., Su, C. H., Lee, Y. Q., Zhang, Y. C., Wang, C. K., Kuo, C. J., & Wu, C. S. (2021). Deep Learning-Based Natural Language Processing for Screening Psychiatric Patients. *Frontiers in Psychiatry, 11*. <https://doi.org/10.3389/FPSYT.2020.533949>
- Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). *Machine Learning Approaches for Clinical Psychology and Psychiatry, 14*, 91–118. <https://doi.org/https://doi.org/10.1146/annurev-clinpsy-032816-045037>
- Ermers, N. J., Hagoort, K., & Scheepers, F. E. (2020). The Predictive Validity of Machine Learning Models in the Classification and Treatment of Major Depressive Disorder: State of the Art and Future Directions. *Frontiers in Psychiatry, 11*, 472. <https://doi.org/10.3389/FPSYT.2020.00472/BIBTEX>
- Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science, 349*(6245), 261–266. <https://doi.org/10.1126/science.aaa8685>
- Kjell, K., Johnsson, P., & Sikström, S. (2021). Freely Generated Word Responses Analyzed With Artificial Intelligence Predict Self-Reported Symptoms of Depression, Anxiety, and Worry. *Frontiers in Psychology, 12*. <https://doi.org/10.3389/FPSYG.2021.602581>
- Stewart, R., & Velupillai, S. (2021). Applied natural language processing in mental health big data. *Neuropsychopharmacology, 46*(1), 252. <https://doi.org/10.1038/S41386-020-00842-1>
- Xu, S., Yang, Z., Chakraborty, D., Victoria Chua, Y. H., Dauwels, J., Thalmann, D., Thalmann, N. M., Tan, B. L., & Chee Keong, J. L. (2019). Automated Verbal and Non-verbal Speech Analysis of Interviews of Individuals with Schizophrenia and Depression. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 225–228. <https://doi.org/10.1109/EMBC.2019.8857071>
- World Health Organization. (2021, September 13). *Depression*. WHO. Retrieved 30 May 2022, from <https://www.who.int/news-room/fact-sheets/detail/depression>
- Østergaard, S. D., Jensen, S. O. W., & Bech, P. (2011). The heterogeneity of the depressive syndrome: when numbers get serious. *Acta Psychiatrica Scandinavica, 124*(6), 495–496. <https://doi.org/10.1111/J.1600-0447.2011.01744.X>