2023 International Conference on Computational Science and Computational Intelligence (CSCI) | 979-8-3503-6151-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/CSCI62032.2023.00242

A Systematic Review on Machine Learning (ML) and Artificial Intelligence (AI) In UNDERSTANDING and ASSESSING women's health

Jones Yeboah
School Of Information Technology,
University of Cincinnati,
Cincinnati, Ohio, USA
yeboahjs@mail.uc.edu

Sophia Bampoh
Cleveland Clinic Florida
USA
BAMPOHS@ccf.org

Foster Addo Yeboah
School Of Information Technology,
University of Cincinnati,
Cincinnati, Ohio, USA
yeboahfr@mail.uc.edu

Isaac Kofi Nti
School Of Information Technology,
University of Cincinnati,
Cincinnati, Ohio, USA
0000-0001-9257-4295

Research track: CSCI-RTHI
Type of Submission: Full/ Regular Paper

***Abstract*—While depression remains a prominent health concern with its associated difficulties, recent years have witnessed the utilization of artificial intelligence (AI) and machine learning (ML) to analyze extensive datasets, thereby enhancing the detection and diagnosis of women's health risks. We conducted a systematic review of AI and ML applications in the context of women's health. This review was carried out using various databases, including PubMed, Scopus, ACM Digital Library, Web-of-Science, and IEEE Explore Digital. We employed specific search terms related to mental health and ML, supplementing our database searches with manual exploration. Our inclusion criteria encompassed articles in English published between 2010 and 2022, focusing on machine learning and artificial intelligence applications and algorithms in the domain of women health, particularly in relation to depression among women. Initially, we identified a total of 495 records based on abstract searches. After removing duplicates, 425 unique reports remained, of which 380 were subsequently excluded following abstract screening. We then assessed 45 full-text articles for eligibility, revealing a wide range of methodologies and outcomes. Notably, the results indicated a high level of accuracy in risk classification, exceeding 90%. Our findings provide valuable perspectives on the implementation and application of AI and ML within the domain of women's health. These insights underscore the potential of these technologies to propel advancements in interventions related to women's health. In summary, we present initial insights and propose future pathways for harnessing AI and ML technologies to tackle depression challenges to enhance women's health on a global scale.**

***Index Terms*—depression, Machine Learning, Artificial Intelligence, women health, mental health**

I. Introduction

In response to the worldwide impact of depression challenges on women health, public health initiatives have placed growing emphasis on the significance of prevention and treatment. To enhance the effectiveness of women health services, there has been a growing emphasis on the role of digital technology. This has led to the development of a diverse array of health technologies and applications aimed at improving access, engagement, and treatment outcomes.

Mobile applications and wearable devices have emerged as valuable tools for monitoring symptoms, conducting health risk assessments, delivering computerized treatments, and fostering peer support for individuals dealing with depression challenges. Additionally, the design of healthcare systems and people's interactions with everyday technology, along with the information collected in electronic health records (EHRs), have contributed to a substantial increase in the volume of data related to people's health and behavior. With the growth of data availability and improvements in computing power, machine learning (ML) research and applications have exploded. Machine Learning methods provides statistical and computational functionalities to build robust systems that can learn from data by itself. This approach is mostly used in gaming and will make great impact to understand large health data. This will improve the understanding of human behaviors, predicting patterns and outcomes. ML approaches is being adopted to revolutionize the issues related to women health. A growing area of research within the field of AI and ML is dedicated to developing technological solutions to address women's health concerns. While a substantial portion of this research concentrates on mobile health applications and tools, as technology keeps evolving the influx of AI and ML in our daily lives has received limited attention on women's health interventions. This paper presents a systematic review that investigates the existing body of research within the field of Human-Computer Interaction (HCI) and Machine Learning. Our focus is on examining the utilization of ML or AI in the domain of women's health, with a specific emphasis on the depression of women, including new mothers and pregnant women.

II. Background

Over the past few years, there has been research surveys and literature reviews in the fields of medical and clinical psychology that focus on the integration of ML with focus on the mental health. These studies have centered on assessing the precision, dependability, and efficiency of algorithms [1], while also exploring the prospects and obstacles linked to their real-world application [2]. Some research has also explored the performance of algorithms in relation to the outcomes of clinical interventions on mental well-being specific conditions [1], or the utilization of mobile sensing data to identify significant behavioral markers indicative of clinical mental well-being states [3].

A notable scoping review conducted by Shatte et al. [4] comprehensively mapped the fundamental concepts in this field by analyzing 300 literature records. Their work provided an overview of diverse applications of machine learning in mental health, highlighting four primary domains. Most of the research is centered around (i) the identification and diagnosis of mental health disorders, with some focusing on (ii) forecasting results, managing treatment, and offering assistance, (iii) investigating public health issues or (iv) addressing administrative and research-related aspects. The authors of this review concluded that machine learning has the potential to enrich clinical and research practices by offering fresh insights into mental well-being. Moreover, the widespread adoption and successful utilization of machine learning applications in real-world scenario may carry profound economic, societal, and personal implications. This expansion also raises ethical concerns related to responsibility and accountability in machine learning-driven decision-making [5]. Notably, machine learning outputs and biases may be fallible, and there is a potential for malicious misuse (as observed in related fields such as criminal justice, loan decisions, adversarial attacks on speech recognition [6], image processing [7], as well as exclusion from digital resources due to inadequate knowledge, access, or other technology-related obstacles [8].

III. Method

We adhered to the PRISMA literature review guidelines to organize the process of identifying and selecting pertinent articles for our review.

A. Search Strategy

We conducted a systematic literature search online, covering articles published between January 1st, 2010, and March 31st, 2022. Our search encompassed multiple databases, including PubMed, Scopus, ACM Digital Library, Web-of-Science, and IEEE Explore Digital. The search queries used were designed to capture articles related to depression, Artificial Intelligence, Machine Learning, prediction, classification, women, and menopause. In addition to database searches, we also employed manual search techniques to supplement our findings.

We utilized keywords in conjunction with designated filters as outlined in the PRISMA guidelines during our search process.

(A) IEEE Explore Digital: (”All Metadata”:”depression”) AND (”All Metadata”:”prediction” OR ”All Metadata”:”Classification”) AND (”All Metadata”:”pregnant” OR ”All Metadata”:”New Mother” OR ”All Metadata”:”Mother” OR ”All Metadata”: ”Women” OR ”All Metadata”:”Woman” OR ”All Metadata”: ”menopause”) AND (”All Metadata”:”Machine Learning” OR ”All Metadata”: ”Artificial Intelligence”)

(B) PubMed: (”depression”) AND (”prediction” OR ”Classification”) AND (”pregnant” OR ”New Mother” OR ”Mother” OR ”Women” OR ”Woman” OR ”menopause”) AND (”Machine Learning” OR ”Artificial Intelligence”)

(C) Web-of-Science: (”mental health”) AND (”prediction” OR ”Classification”) AND (”pregnant” OR ”New Mother” OR ”Mother” OR ”Women” OR ”Woman” OR ”menopause”) AND (”Machine Learning” OR ”Artificial Intelligence”)

(D) Scopus: TITLE-ABS-KEY ((”depression”) AND (”prediction” OR ”Classification”) AND (”pregnant” OR ”New Mother” OR ”Mother” OR ”Women” OR ”Woman” OR ”menopause”) AND (”Machine Learning” OR ”Artificial Intelligence”)) AND PUBYEAR ¿ 2009 AND PUBYEAR ¡ 2023

(E) ACM Digital Library:[All: ”depression”] AND [[All: ”prediction”] OR [All: ”classification”]] AND [[All: ”pregnant”] OR [All: ”new mother”] OR [All: ”mother”] OR [All: ”women”] OR [All: ”woman”] OR [All: ”menopause”]] AND [[All: ”machine learning”] OR [All: ”artificial intelligence”]] AND [Publication Date: (01/01/2010 TO 03/31/2022)]

B. Study selection

We conducted this review in accordance with the PRISMA guidelines, which provide an evidence-based framework for the selection and reporting of systematic reviews and meta-analyses [9]. Considering the widespread occurrence of depression concerns among women of diverse age groups and medical circumstances, diagnostic or age-related factors were not employed as exclusion criteria.

Initially, all articles identified through the database searches were considered based on their metadata, including titles and summaries. All works relevant to the subject matter were reviewed at this stage. However, due to the sheer volume of studies retrieved, a second selection was conducted using specific keywords. The utilization of keywords served to streamline the development of a classification scheme and ensured that the review encompassed the most current research [10]. In the third phase, we scanned the reference lists of related articles to identify any additional relevant studies. At the conclusion of this final phase, we identified 45 studies that met the eligibility criteria for inclusion in our review.

C. Data Collection Process

Two reviewers conducted independent assessments of both abstracts and full-text articles. In cases of disagreement, a third reviewer made the final decision. Source documents were evaluated based on the following inclusion criteria: (1) Articles published in the English language only, (2) Articles from the year 2010 through March 31st, 2022, (3) Relevance to depression, (4) Focus on women, (5) Involvement of Machine Learning and Artificial Intelligence applications and algorithms. We managed the studies identified through this search strategy using Microsoft Excel. Reports that did not meet the inclusion criteria were systematically excluded, and the reasons for their exclusion were documented. This encompassed studies unrelated to depression and classification specific to depression among women, clinical trials, and those that did not fulfill any of the inclusion criteria.

To visually represent the inclusion/exclusion of studies at each stage of the process (identification, screening, eligibility, and inclusion [11]), we created a PRISMA flowchart [9]. The selected reports were subsequently categorized based on their relevance to mental health, applications and algorithms in the field of AI and ML, along with their methodology. The latter category made a distinction between supervised learning, which entails predicting outcomes by training a classifier with a set of specified input values, and unsupervised learning, which operates without provided labels and aims to elucidate data patterns, frequently through clustering, based solely on input measures.

Our search covered international indexed journals and conferences pertaining to AI/ML in mental health affecting women's health. We searched the following electronic databases: ACM (n = 309), Scopus (n = 60), IEEE Explore Digital (n = 13), PubMed (n = 46), and Web-Of-Science (n = 67).

D. Threads to validity

Assessing the strengths and weaknesses of a systematic review requires careful consideration of factors that impact validity, as highlighted by previous research [12]. These factors predominantly relate to the selection of studies, data extraction, and the potential for researcher bias within the scope of this study.

To identify relevant studies, we conducted scans across the five aforementioned search engines. However, there is a possibility that other pertinent works may not have been captured in the search results. To mitigate this potential limitation, we manually searched the reference lists of the selected studies to uncover additional related research.

Data extraction represents a pivotal aspect of this work. To minimize the risk of inaccuracies in data extraction, we subjected the studies to two separate evaluations on different days, ensuring that the necessary data for addressing the research questions was accurately collected.

In the context of study selection and data extraction, it is worth mentioning the potential for researcher bias [13]. As a precautionary step, we implemented a systematic review approach wherein one researcher conducted the initial selection of studies, followed by independent verification by another researcher [14]. This approach was implemented with the involvement of three researchers to minimize the impact of any potential researcher bias.

E. Data Extraction

We utilized a data extraction form to gather pertinent information from the chosen studies in order to address our research question. The selected studies underwent three separate evaluations on different days, each conducted by different authors.

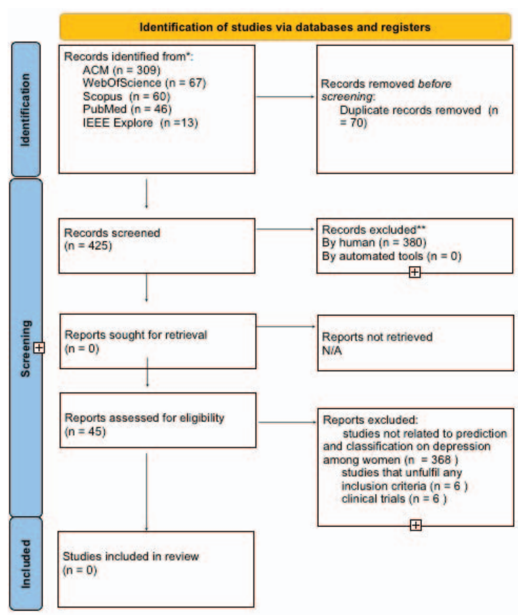

Fig. 1. Systemic Review Process

IV. FINDINGS

The final set of papers included in the review encompasses those published between 2010 and 2022. Notably, there has been a recent surge in publications, with almost two-thirds of all papers appearing in the past four years. Among the 45 studies analyzed, 28 were presented at conferences, comprising 9 abstracts, 6 short papers, and 13 full-length proceeding papers. Additionally, the review includes 11 journal articles and 6 symposium or workshop papers.

The primary focus of the majority of these papers is the adoption of Machine Learning approaches using specific datasets, as their principal research contribution (n = 25). Ten papers introduce various concepts, data methodologies, models, or systems, while three papers employ pre-existing ML algorithms to achieve a more profound understanding of the state of depression in women and enhance their interaction with healthcare providers. Furthermore, only a small subset of papers details the execution of practical studies involving comprehensive ML systems or the assessment of the accuracy of ML predictions. Lastly, two papers specifically address design considerations for user-centric, deployable ML systems.

V. Conclusion

Our review underscores a significant upswing in interest regarding the application of AI and ML in addressing depression through health interventions. This surge in interest is indicative of exciting research initiatives aimed at harnessing the potential of AI and ML to provide solutions for mitigating the challenges associated with depression. Our comprehensive review has provided valuable insights into this area of research, shedding light on the latest tools and applications in AI and ML that are predominantly utilized to enhance healthcare delivery for women. Furthermore, we have discussed the various approaches employed in Machine Learning systems that have shown promise in addressing depression issues to improve women's health.

However, challenges persist in the adoption of ML systems, particularly concerning the availability of large datasets. There is a pressing need for further research, especially in conducting empirical studies that encompass end-to-end ML systems and assessing the quality of ML predictions in the context of women health. Additionally, we have identified two papers that specifically delve into the design considerations for user-centric, deployable ML systems.

It is imperative to delve deeper into more innovative explorations within the design space, paving the way for fresh perspectives and guiding future research questions that will be crucial for ML to truly benefit this evolving domain with a multitude of novel tools.

References

[1] Yena L., Renee-Marie R., Rodrigo B. M., Justin J. B., Joshua D. R., Alisson T., Elisa B., et al. (2018). Applications of machine learning algorithms to predict therapeutic outcomes in mental health: A meta-analysis and systematic review. Journal of Affective Disorders 241, 519–532. DOI:https://doi.org/10.1016/j.jad. 2018.08.073

[2] Danilo B., Andreas M. (2018). Machine learning for precision psychiatry: Opportunities and challenges. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 3, 3 (2018), 223–230. DOI:https://doi.org/10.1016/j.bpsc.2017.11.007

[3] David C. M., Mi Z., Stephen M. S. (2017). Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annual Review of Clinical Psychology 13, 23–47. DOI:https://doi.org/10.1146/annurev- clinpsy- 032816- 044949

[4] Adrian B. R. S., Delyse M. H., Samantha J. T. (2019). Machine learning in mental health: A scoping review of methods and applications. Psychological Medicine 49, 9 (2019), 1426–1448. DOI:https://doi.org/10.1017/S0033291719000151

[5] Victoria B., Keith E. (2001). Intelligibility and accountability: Human considerations in context-aware systems. Human-Computer Interaction 16, 2–4 (2001), 193–212.DOI:https://doi.org/10.1207/S15327051HCI16234_05

[6] Chuan G., Mayank R., Moustapha C.,Laurens V. D. M. (2017). Countering adversarial images using input transformations. arXiv:1711.00117. DOI:https://arxiv.org/abs/1711.00117

[7] Moustapha M. C., Yossi A., Natalia N.,Joseph K. (2017). Houdini: Fooling deep structured visual and speech recognition mossdels with adversarial examples. In Proceedings of the s Neural Information Processing Systems (NIPS'17). 6977–6987.

[8] Ben G., Dan R., Sara S., Hannah C., Helena G., Til W. (2019). Digital exclusion among mental health service users: Qualitative investigation. Journal of Medical Internet Research 21, 1 (2019), e11696. DOI:https://doi.org/10.2196/11696

[9] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

[10] Petersen K, Feldt R, Mujtaba S, Mattsson M (2008) Systematic mapping studies in software engineering. In: 12th international conference on evaluation and assessment in software engineering (EASE) 12, pp 1–10

[11] Howick, J.; Chalmers, I.; Glasziou, P.; Greenhalgh, T.; Heneghan, C.; Liberati, A.; Moschetti, I.; Phillips, B.;Thornton, H. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Oxford Centre for Evidence-Based Medicine. Available online: https: //www.cebm.net/index.aspx?o=5653 (accessed on 31 July 2020).

[12] Roberto R, Lima JP, Teichrieb V (2016) Tracking for mobile devices: a systematic mapping study. Comput Graph 56:20–30

[13] Petersen K.,Gencel C.(2013)Worldviews,research methods,and their relationship to validity in empirical software engineering research. In: 2013 Joint Conference of the 23rd international workshop on software measurement and the 8th international con- ference on software process and product measurement. IEEE, pp 81–89

[14] Kitchenham B, Brereton OP, Budgen D, Turner M, Bailey J, Linkman S (2009) Systematic literature reviews in software engineering: a systematic literature review. Inf Softw Technol 51(1):7–15