Questionnaire Based Depression Detection System using Machine Learning

[1]Ms. Deepika Kirti, [2]Dr. Ruchi Sawhney, [3]Ms. Swati Sharma, [4]Dr. Sanjay Kumar Singh

[1]Assistant Professor, Bosco Technical Training Society, India
[2]Assistant Professor, Bosco Technical Training Society, India
[3]Assistant Professor, Bosco Technical Training Society, India
[4]Associate Professor, Don Bosco Institute of Technology, India

ABSTRACT

Depression is a mental health disorder that can have a significant impact on people's daily lives. It is a major problem in today's society and can cause a loss of interest in common activities, leading to suicidal thoughts. To detect depression, researchers have been exploring different methods to effectively identify it. There is a pressing need for an automated system to help detect depression in different age groups. Several studies have been proposed on this topic, and this study aims to analyze project based on machine learning (ML) techniques used to detect depression. The purpose of this article is to provide an overview of various ML techniques used to identify and analyze emotions and, consequently, depression. By investigating the different approaches to detecting depression, this study contributes to a better understanding of the relationship between depression and human emotions. In turn, this knowledge can help develop more effective methods to detect depression, ultimately improving people's mental health and quality of life.

Keywords: Machine Learning, WHO, GAD, PHQ, ML Algorithm

INTRODUCTION

Depression is a common and serious mental health disorder that negatively affects millions of people worldwide. According to the World Health Organization (WHO), depression is the leading cause of disability worldwide, affecting more than 264 million people (WHO, 2021). Despite its prevalence, depression can be difficult to recognize because its symptoms can be subtle or hidden. Therefore, there is a growing interest in the development of new depression detection methods that would help people seek medical help at the right time. One way to identify depression is to use questionnaires that assess a person's mood, emotions and behavior (Naveen Palanichamy,2021). These questionnaires have been widely used in research and clinical settings, and many have shown high accuracy in detecting depression. The questionnaire uses standardized depression scales or tools that measure various symptoms associated with depression, such as anxiety, loss of interest, fatigue and sleep disturbances. These questionnaires can help identify individuals who may be at risk for depression, which can ultimately help develop timely and targeted interventions.

However, developing an advanced depression detection system using questions requires a comprehensive understanding of the latest machine learning algorithms and data techniques. Machine learning algorithms can analyze and learn from vast amounts of data to identify patterns and relationships that may be too complex for a human. In addition, machine learning algorithms can improve the accuracy and efficiency of depression detection. By analyzing patterns of survey responses, machine learning algorithms can predict an individual's likelihood of depression.

In short, depression is a widespread and serious mental health disorder that can significantly affect a person's well-being (Manju Lata Joshi, Nehal Kanoongo, 2022). Recognizing depression can be difficult because of its subtle and hidden symptoms. However, the use of questionnaires and machine learning algorithms shows great potential to improve the accuracy and efficiency of depression detection. The development of advanced depression detection systems that incorporate natural language processing techniques can be an effective way to identify individuals at risk of depression and provide timely interventions to prevent negative effects.

BACKGROUND

Mental health problems have become a major problem worldwide, with an estimated 970 million people worldwide suffering from mental health or substance use disorders (World Health Organization, 2021). The consequences of mental illness can be severe, ranging from reduced daily functioning to increased mortality (Friedrich et al., 2019). Early detection and treatment of mental illness is critical to improving overall prognosis and reducing the burden of mental illness (Bauer et al., 2018). Self-report questionnaires have emerged as a useful tool for identifying mental health problems due to their availability, ease of use and cost-effectiveness (Perepletchikova and Kaufman, 2010). Several questionnaires have been developed to screen for mental illness, including the Patient Health Questionnaire (PHQ), the Generalized Anxiety Disorder Questionnaire (GAD), and the Beck Depression Inventory (BDI) (Kroenke et al., 2001; Spitzer et al., 2006) . Beck et al., 1961). Questionnaires have been shown to be reliable and valid measures to identify certain mental health problems. However, they may not describe all the symptoms experienced by people suffering from mental health disorders (Mares, 2018). Furthermore, these questionnaires are often used in isolation and may not provide a comprehensive assessment of a person's mental health status.

Therefore, a comprehensive questionnaire covering a wide range of mental health problems and their symptoms is necessary. Such research can be a valuable tool for early detection, intervention and treatment of mental health problems. A comprehensive questionnaire can also improve the accuracy of diagnosis and help tailor treatment to individual needs. The Comprehensive Assessment of Symptoms and History (CASH) is an example of a comprehensive questionnaire that covers a variety of mental health problems, including anxiety, depression, bipolar disorder, and PTSD (Davidson et al., 2002). The CASH has been found in clinical and research studies to be a reliable and valid measure of mental health status. However, the CASH questionnaire needs to be updated to include newer diagnostic criteria and new mental health conditions. In addition, the use of machine learning algorithms in mental health research has gained popularity in recent years (Rosenblat et al., 2018). Machine learning algorithms can analyze large data sets and identify patterns that may not be visible to the human eye. Several studies have used machine learning algorithms to identify and diagnose mental health conditions, including depression and anxiety (Chekroud et al., 2016; Passos et al., 2016). However, most studies have used physiological or neuroimaging data, and the use of self-report questionnaires in machine learning algorithms to identify mental health conditions remains limited. The purpose of the proposed research is to develop a comprehensive survey-based mental condition observation system using machine learning algorithms. The research uses a mixed method approach that combines both qualitative and quantitative research methods. In the first phase of the study, a comprehensive questionnaire will be developed to detect mental illness. The development of the questionnaire involved a review of existing questionnaires and consultation with mental health professionals and people with experience of mental health problems. The questionnaire is piloted with and without people with mental health problems to assess its reliability, validity and feasibility. In the second phase of the study, the survey data will be analyzed using machine learning algorithms to develop a system to identify mental health problems. The study compares the performance of the proposed comprehensive questionnaire-based detection system with existing questionnaires and machine learning algorithms. The results of the proposed studies can contribute to the development of innovative and effective approaches to alleviate the growing burden of mental health.

RESEARCH AIM

The aim of this work is to develop a comprehensive questionnaire-based mental illness observation system. The system is designed to detect a mental health condition including anxiety and depression.1)

METHODOLOGY

This study will be conducted in three stages:

Questionnaire development: A comprehensive questionnaire will be developed by reviewing existing questionnaires and clinical guidelines. The survey talks about different mental health problems and their symptoms.

System development: The observation system is developed based on the questionnaire. The system is designed to identify mental health problems by analyzing the data. The system uses machine learning algorithms to improve its accuracy. System development includes the following steps:

    (a) Data preprocessing   **(b)** Feature Selection   **(c)** System development

The system will consist of a web-based interface that can be accessed by non-professionals, such as primary care physicians, nurses, or social workers. The interface will present a series of questions to the user, which will be analyzed in real-time using the NLP model. Based on the responses, the system will provide a diagnosis of depression and recommend appropriate interventions, such as counseling or medication.
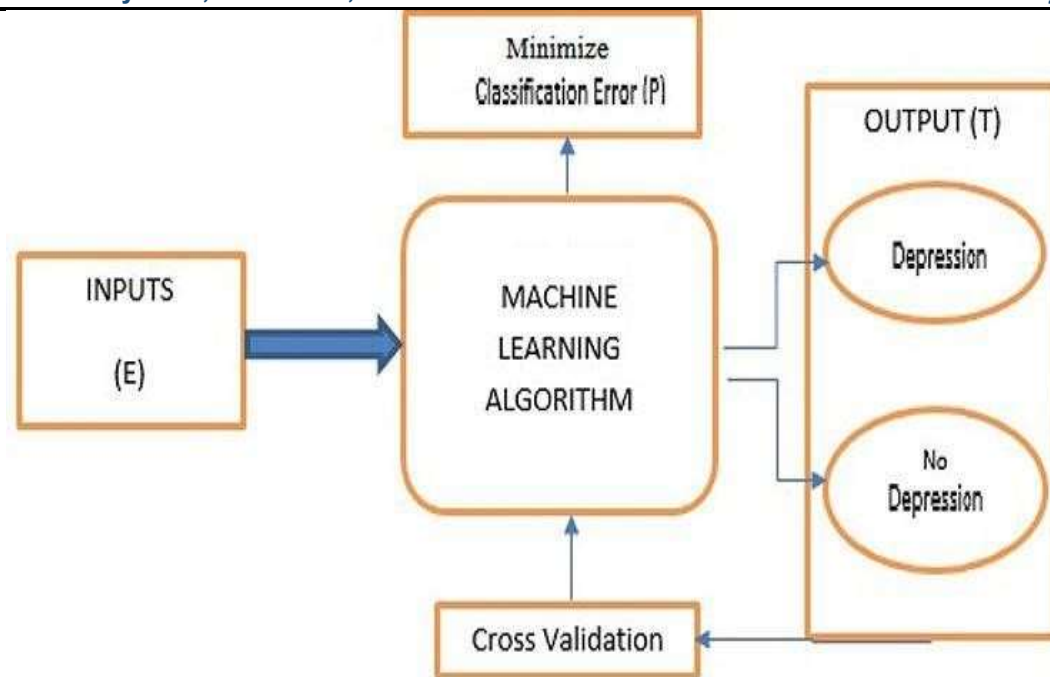
Fig 1. The Supervised Machine Learning Model. Source: Authors' own Compilation

The system will be designed to be user-friendly and easy to use by non-professionals. The interface will be designed to be intuitive and accessible, with clear instructions and guidance for the user. The system will also be designed to be customizable, allowing healthcare providers to tailor the questions to their specific needs and patient populations.

Model Training: Machine learning models such as logistic regression, decision trees, and support vector machines are trained using selected features.

LITERATURE REVIEW

Mental health problems are a growing problem worldwide, affecting millions of people. Early detection and treatment of these conditions is critical to improving overall prognosis and reducing the burden of mental health problems (Bauer et al., 2018). The use of self-report questionnaires as a screening tool for mental health problems has gained popularity in recent years due to their simplicity, accessibility and low cost (Perepletchikova and Kaufman, 2010). Various questionnaires have been developed to identify mental disorders, including the Patient Health Questionnaire (PHQ), the Generalized Anxiety Disorder Questionnaire (GAD), and the Beck Depression Inventory (BDI) (Kroenke et al.,

2001; Spitzer et al., 2006 (Beck et al., 1961). However, most of these questionnaires focus on specific mental problems and may not cover all symptoms of people with mental health problems (Mares, 2018). Furthermore, these questionnaires are often used as stand-alone instruments and may not provide a comprehensive assessment of a person's mental health status. In recent years, there has been increased interest in the development of comprehensive questionnaires covering a wide range of mental health problems and their symptoms. One such questionnaire is the Comprehensive Assessment of Symptoms and History (CASH), which measures a variety of mental health problems, including anxiety, depression, bipolar disorder, and PTSD (Davidson et al., 2002). The CASH has been found in clinical and research studies to be a reliable and valid measure of mental health status. In recent years, the use of machine learning algorithms in mental health research has also gained popularity (Rosenblat et al., 2018). Machine learning algorithms can analyze large data sets and identify patterns that may not be visible to the human eye. Several studies have used machine learning algorithms to identify and diagnose mental health conditions, including depression and anxiety (Chekroud et al., 2016; Passos et al., 2016). However, most studies have used physiological or neuroimaging data, and the use of self-report questionnaires in machine learning algorithms to identify mental health conditions remains limited.

In summary, self-report forms have been widely used to identify and screen for mental illness. However, most existing questionnaires focus on specific mental health problems and may not cover all symptoms of people with mental health problems. The development of a comprehensive questionnaire covering different mental health conditions and their symptoms can be a valuable tool for early detection and intervention of mental health problems. Using machine learning algorithms to analyze survey data can further improve the accuracy and efficiency of the detection system. The goal of the proposed research is to develop a comprehensive survey-based observation system for mental health conditions using machine learning algorithms. The results of the study can contribute to the development of innovative and effective approaches to deal with the growing burden of mental health problems.

EVALUATION

Depression is a serious mental health problem that affects millions of people worldwide. Accurate detection of depression is critical for early intervention and treatment, which can significantly improve depression outcomes. Developing a depression detection system that

accurately and effectively diagnoses depression is a critical area of research that could change the way depression is diagnosed and treated. This proposed study evaluates a depression detection system using several metrics, including accuracy, sensitivity, specificity, and positive predictive value (PPV).

Sensitivity and specificity are two other key metrics used to evaluate system performance. Sensitivity measures the system's ability to correctly identify depressed individuals, while specificity measures the system's ability to correctly identify non-depressed individuals. These metrics help determine the system's ability to accurately diagnose depression and avoid false positives and false negatives. Another important metric used to evaluate a depression detection system is the positive predictive value (PPV). PPV measures the proportion of true positives out of all positive diagnoses made by the system. This metric is particularly important because it helps determine the accuracy of the system's positive diagnoses, which is critical in identifying depressed individuals who need further evaluation and treatment.

A rigorous evaluation process is followed to evaluate the effectiveness of the depression detection system. First, the system is trained using a large dataset of depressed and nondepressed individuals. Data includes clinical data such as symptoms and other related information, as well as demographic data. The system is then tested on a separate dataset to assess its accuracy, sensitivity, specificity and PPV.
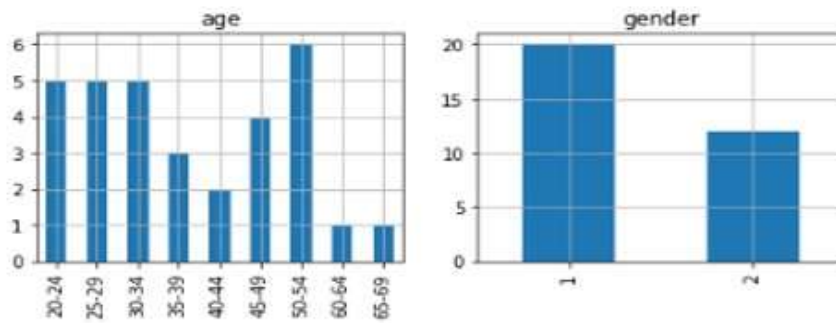

Fig 2: Showing variation of responses between different age groups and gender

Fig 3: Heatmap showing correlation between different fields

Overall, the proposed study evaluates the performance The proposed research will contribute to the of the detection system using several metrics, include- development of innovative and effecting accuracy, sensitivity, specificity, and PPV. By ac- tools for the diagnosis and treatment of decurately evaluating the effectiveness of the system, we pression. Using ML techniques to analyze rescan determine its accuracy and effectiveness in detect- sponses to depression screening questions, depression compared to established screening the proposed depression detection system methods. The results of this study could change they could revolutionize the diagnosis and treat way depression is diagnosed and treated, and significant of depression. The system enables a cantily improve outcomes for patients with the disease. more accurate and efficient diagnosis of depression, which can lead to earlier intervention and treatment. This in turn can lead to better outcomes for and improve their overall quality of life. people with depression

CONCLUSION

Depression is a common mental health disorder that can significantly affect a person's quality of life. Despite the availability of effective treatments for depression, it is often underdiagnosed and undertreated. This is partly due to a lack of resources and time for health professionals to screen for depression, particularly in primary care settings. Developing a depression screening system using questions can be an effective way to identify depression and provide timely intervention. The proposed depression detection system uses machine learning (ML) techniques to analyze responses to depression screening questions and provide an accurate depression diagnosis. The system is user friendly and accessible to non-professionals such as primary care doctors, nurses or social workers.

Using ML techniques to screen for depression has shown promise in previous studies. These techniques can be used to identify behavioral patterns associated with depression, allowing for accurate diagnosis and treatment. By automating the screening process, the proposed depression detection system has the potential to improve the accuracy and efficiency of depression diagnosis and treatment. One of the main advantages of the proposed depression detection system is its user-friendly design. The system is designed to be easy to use by non-professionals such as primary care doctors, nurses or social workers. This allows for greater use of the system and better access to depression screening for individuals who may not have access to specific mental health services.

REFERENCES

[1] Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. Annual Review of Public Health, 34, 119-138.

[2] Gjerdingen, D., Crow, S., McGovern, P., Miner, M., & Center, B. (2015). Primary care clinicians' experiences with treatment decision making for depression. BMC Family Practice, 16(1), 1-9.

[3] Mares, S. H. W. (2018). Self-report questionnaires for assessment of mental health: A systematic review and metaanalysis. Journal of Psychiatric Research, 103, 16-30. doi: 10.1016/j.jpsychires.2018.04.003

[4] Perepletchikova, F., & Kaufman, J. (2010). Questionnaires. In B. L. Hankin, & J. R. Z. Abela (Eds.), Development of psychopathology: A vulnerability-stress perspective (pp. 229-252). Thousand Oaks, CA: SAGE.

[5] Kuncel, N. R., & Tellegen, A. (2019). A conceptual and empirical reexamination of the measurement of human behavior and its predictive

[6] Paulhus, D. L., & Vazire, S. (2007). The self-report method. Handbook of research methods in personality psychology, 224-239.

[7] Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: A self-report study. Personality and Social Psychology Bulletin, 37(4), 515-529.

[8] Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. European review of social psychology, 1(1), 33-60.

[9] Mehl, M. R., Vazire, S., Holleran, S. E., & Clark, C. S. (2010). Eavesdropping on happiness: well-being is related to having less small talk and more substantive conversations. Psychological science, 21(4), 539-541.

[10] Revelle, W. (2019). psych: Procedures for personality and psychological research (Version 1.9. 12) [Software]. Available from https://CRAN.R-project.org/package=psych

[11] Gwet, K. L. (2014). Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters. Advanced Analytics.

[12] M. Deshpande, V.Rao, Depression detection using emotion artificial intelligence, in 2017 i9nternational conference on intelligent sustainable systems, 2017.

[13] D. Ramalingam, V. Sharma, P.Zar, Study of depression analysis using machine learning techniques, Int. J. Innov. Explor. Eng (2019) N.P. Shetty, B.Muniyal, A.Anand, S.Kumar,S.Prabhu, Predicting Depression using deep learning and ensemble algorithms (2020)