Vanshika Thakur, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Early Detection of Mental Distress in Gen Z

Vanshika Thakur[1], Sushmita Guha[2], Nafiya Kausar N.[3], Mrs.Jayahsree Kudari[4]

[1,2,3,] Student,School of Computer Science and Information Technology,Jain (Deemed to be University),Jayanagar,Bengaluru,Karnataka,India

[4]Associate Professor,School of Computer Science and Information Technology,Jain (Deemed to be University),Jayanagar,Bengaluru,Karnataka,India

Abstract- Mental distress among Generation Z has emerged as a significant public health concern due to increased exposure to digital environments, academic stress, and social pressures. This survey paper analyzes existing methodologies for early detection of mental distress using machine learning and artificial intelligence techniques. The study reviews various approaches including Natural Language Processing (NLP), sentiment analysis, and deep learning models applied to social media and behavioral datasets. It also explores commonly used datasets, particularly from Kaggle, and evaluates their effectiveness in predictive modeling. The survey identifies limitations in current systems such as data bias, privacy concerns, and lack of real-time adaptability. Furthermore, research gaps and future directions are discussed, emphasizing the need for multimodal data integration and ethical AI deployment.

Keywords: Mental Distress, Gen Z, Machine Learning, Deep Learning, Natural Language Processing (NLP), Sentiment Analysis, Early Detection, Social Media Analytics, Healthcare AI, Kaggle Datasets.

I. INTRODUCTION

Generation Z (individuals born approximately between 1997–2012) is increasingly experiencing mental health challenges such as anxiety, depression, and stress. Unlike previous generations, Gen Z is highly active on digital platforms, making their behavioral and emotional patterns more observable through online interactions.

The uploaded paper highlights that early detection of mental distress is critical to prevent severe psychological consequences such as self-harm or suicide. Traditional diagnostic approaches rely heavily on clinical interviews, which are time-consuming and not scalable. Therefore, integrating machine learning and AI techniques provides a proactive and scalable solution.

Recent advancements in AI enable the analysis of large-scale textual and behavioral data from platforms like Twitter, Reddit, and Instagram. These technologies can identify subtle linguistic cues, emotional patterns, and behavioral changes, making early detection feasible and efficient.

Objectives

The primary objectives of this survey paper are:

1. To analyze existing AI-based techniques for early detection of mental distress.
2. To review machine learning and deep learning models used in mental health prediction.
3. To examine datasets (especially Kaggle- based) used in research.
4. To identify limitations in current detection systems.
5. To highlight research gaps and suggest future improvements.

II. REVIEW OF LITERATURE

Mental health issues are a growing concern for Generation Z, and researchers are working hard to find new ways to detect them early on. One area that's getting a lot of attention is the use of artificial intelligence, psychology, and data science to identify patterns of emotional and psychological distress on social media platforms like Twitter, Reddit, and Instagram. Studies have shown that the way people write and behave online can be a strong indicator of their mental health, with things like linguistic patterns, emotional tone, and posting frequency all potentially signaling conditions like depression, anxiety, and loneliness. Natural Language Processing techniques are being used to analyze the text people post online, with tools like sentiment analysis, topic modeling, and emotion detection helping to identify

© 2026 Vanshika Thakur, This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

early warning signs. Deep learning approaches like Long Short-Term Memory and Convolutional Neural Networks are also being used, as they can pick up on contextual meaning and long-term patterns in language. Some researchers are focusing on datasets related to student mental health, which provide valuable information about the psychological conditions, academic pressures, and lifestyle factors that affect Gen Z.

However, despite the progress being made, there are still some big challenges to overcome. Many of the existing models are trained on limited datasets and aren't able to monitor mental health in real-time. There are also concerns around data privacy, and the fact that many datasets are not balanced or diverse. Additionally, most analyses are only looking at text data, and not taking into account other factors like images or videos. As a result, while AI-based approaches have improved our ability to detect mental health issues, they still need to be more robust and reliable. We need more diverse datasets and better models if we're going to make a real difference in people's lives. The good news is that this is an emerging area of research with a lot of potential, and by working together, we can create new tools and approaches that will help us detect mental health issues early on and provide better support to those who need it. It's worth noting that mental health is a complex issue, and there's no one-size-fits-all solution. But by using AI and machine learning to analyze online data, we may be able to identify patterns and warning signs that can help us intervene earlier and provide more effective support.

This could be especially important for young people, who are often more comfortable seeking help online than they are in person. By leveraging these technologies, we can create new pathways to support and care that are tailored to the needs of Gen Z. Overall, the use of AI and machine learning to detect mental health issues is a rapidly evolving field, with a lot of potential to make a positive impact. While there are still challenges to overcome, the benefits of early detection and intervention make it an area worth exploring further. By working together and sharing our knowledge and expertise, we can create new tools and approaches that will help us better support the mental health and wellbeing of Generation Z.

III. METHODOLOGY

The approach used in this study is a thorough and systematic one, combining a survey of existing research with a detailed analysis of the techniques used to detect mental distress in Gen Z. To start, the study involves a comprehensive review of research papers, journals, and conference publications that focus on using artificial intelligence and machine learning to identify mental health issues. This review helps to identify the most commonly used datasets for building predictive models, including those based on social media, student mental health, and surveys that capture emotional, behavioral, and psychological attributes. The next step is to examine the algorithms that are most often used in early detection systems. Traditional machine learning algorithms like Logistic Regression, Support Vector Machine, Decision Tree, and Random Forest are popular due to their simplicity and ease of interpretation. At the same time, deep learning models like LSTM, CNN, and neural network- based architectures are also studied for their ability to handle large amounts of textual data effectively.

A key part of the approach is comparing these algorithms based on how well they perform, using metrics like accuracy, precision, recall, and F1-score. Another important aspect is identifying the features that contribute most to detecting mental distress, such as the use of emotional words, how often negative sentiments are expressed, patterns of social media interaction, and behavioral characteristics of users. Overall, the goal of this approach is to provide a clear understanding of how different datasets and algorithms can be combined to create an effective early detection system for mental distress in Gen Z. By following this structured methodology, the study aims to contribute to the development of better tools for identifying and addressing mental health issues in this demographic.

The methodology is designed to be flexible and adaptable, allowing for the incorporation of new datasets and algorithms as they become available.

This will help to ensure that the detection system remains effective and up-to-date, and can continue to provide valuable insights into the mental health needs of Gen Z. In terms of the specific techniques used, the study will draw on a range of existing research and methodologies, including machine learning and artificial intelligence. The use of these techniques will be tailored to the specific needs and goals of the study, and will be designed to provide the most accurate and effective results possible.

By taking a systematic and comprehensive approach, the study aims to make a significant contribution to our understanding of mental distress in Gen Z, and to the development of effective tools for detecting and addressing this issue. The ultimate goal is to improve the mental health and wellbeing of young people, and to provide them with the support and resources they need to thrive. The study will also consider the potential limitations and challenges of the methodology, and will take steps to address these wherever possible. This will help to ensure that the results are reliable and valid, and can be used to inform future research and practice. Overall, the methodology used in this study is a rigorous and systematic one, designed to provide a detailed and comprehensive understanding of the issues surrounding mental distress in Gen Z. By following this approach, the study aims to make a valuable contribution to the field, and to help improve the mental health and wellbeing of young people.

DATASETS USED (Kaggle Datasets) Mental health is a complex issue that affects many people, and researchers are working hard to develop ways to detect mental distress using machine learning models. To do this, they need good datasets to train and test their models. Most researchers use datasets that are available to the public, like those on Kaggle, because they have structured and labeled data that is easy to use. One popular dataset is the depression detection dataset, which has social media posts labeled with emotions like depression, anxiety, and stress. This dataset is really useful for Natural Language Processing because it lets researchers analyze the tone and emotions in user posts. Another important dataset is the student mental health dataset, which focuses on Gen Z students and includes things like academic stress, sleep patterns, and emotional well-being. There are also survey-based datasets like the mental health in tech dataset, which provides information about psychological conditions and behaviors.

Social media datasets like Twitter sentiment and Reddit mental health datasets are also widely used because they give real-time emotional data from users. These datasets help researchers build models that can identify early symptoms of mental distress. However, the paper also points out that the datasets need to be improved in terms of quality and diversity. This is important because it will help researchers build more accurate and generalizable models. By using these datasets, researchers can develop models that can detect mental distress early on, which can be really helpful in preventing more serious problems. Overall, datasets play a crucial role in developing accurate machine learning models for mental distress detection, and researchers are working hard to improve them.

For example, imagine you're a researcher trying to build a model that can detect depression in social media posts. You would use a dataset like the depression detection dataset to train your model, and then test it on a separate dataset to see how well it works. If your model is accurate, it could be used to help people who are struggling with depression by identifying early symptoms and providing support. This is just one example of how datasets can be used to improve mental health outcomes.

In addition, datasets can also be used to identify patterns and trends in mental health. For instance, a dataset might show that certain age groups or demographics are more likely to experience mental distress. This information can be used to develop targeted interventions and support services. By analyzing datasets and identifying patterns, researchers can gain a better understanding of mental health issues and develop more effective solutions. Overall, datasets are a critical component of mental health research, and they have the potential to make a big impact on our understanding and treatment of mental health issues. By improving the quality and diversity of datasets, researchers can

develop more accurate and generalizable models that can help people who are struggling with mental distress.

III. LIMITATIONS

Despite the growing use of machine learning and artificial intelligence in mental health detection, the paper identifies several limitations in existing approaches. One of the major limitations is the issue of data privacy, as mental health data is highly sensitive and cannot always be collected or shared openly. Another limitation is the presence of biased datasets, where most of the available data is collected from specific regions, age groups, or social media platforms, which reduces the generalization capability of machine learning models. The paper also highlights that many existing models focus only on textual data and ignore other important factors such as facial expressions, voice patterns, and behavioral changes. This limitation reduces the accuracy of early detection systems because mental distress cannot always be identified through text alone. In addition, the paper mentions that most research studies focus only on offline datasets and do not provide real-time detection systems that can continuously monitor the mental health condition of individuals. Another limitation is the occurrence of false positive and false negative predictions, which can create serious problems in real-world applications. Therefore, although current technologies show promising results, they still require significant improvements to become reliable tools for mental health detection.

IV. RESEARCH GAP

Mental health issues are a big problem for Gen Z, and we need to find better ways to detect them early. When we look at what's already been studied, we can see some big gaps in what we know. One of the main issues is that most studies only look at what people write on social media, which isn't enough to really understand what's going on. We need to use lots of different types of data, like what people say, how they look, and how they behave, to get a more complete picture. Another problem is that we don't have systems that can watch what people are doing in real- time and catch the early signs of mental health issues. We also need to make sure that our models can explain why they think someone is struggling, which is important for using these systems in healthcare. Right now, a lot of studies just focus on getting the right answer, but they don't tell us why they got that answer. We also need to study Gen Z more, since they're the ones who are most active online and most at risk for mental health issues. And we need to look at how mental health issues affect different groups of people, because what works for one group might not work for another.

All of these gaps in what we know mean that we need to do more research to find better ways to detect mental health issues early. For example, imagine if we had a system that could look at all the different things someone does online, like what they post, how they interact with others, and how they behave, and use that to detect if they're struggling with their mental health. That would be a much more powerful tool than just looking at what they write. And if we could make sure that our models are explainable, that would make it easier for healthcare professionals to use them and trust the results. By focusing on these areas, we can make a big difference in how we detect and treat mental health issues in Gen Z. We just need to keep doing research and working to fill in the gaps in what we know. This will help us create better systems for detecting mental health issues, and ultimately, help more people get the support they need. It's time for us to take mental health seriously and do more to support the people who are struggling. We owe it to ourselves, our friends, and our family members to make sure we're doing everything we can to detect mental health issues early and provide the right support. By working together, we can make a real difference and create a healthier, happier world for everyone.

FUTURE SCOPE

The potential for early detection of mental health issues in Gen Z is huge and looks promising, thanks to how fast artificial intelligence and data science are advancing. To take this further, researchers should work on creating systems that can look at lots of

different types of data - like what people write, their facial expressions, how they speak, and information from wearable devices - to make predictions more accurate. Another key area to focus on is making systems that can monitor mental health in real-time, constantly checking user behavior and sending out warnings when it notices signs of trouble. It's also important that these systems can explain why they're making certain predictions, so healthcare professionals can trust them. We should also be working on systems that are tailored to each individual, rather than using the same approach for everyone.

And, we need to find ways to integrate these systems with healthcare platforms and mobile apps. Using advanced language models and deep learning techniques can make a big difference in how well these systems work. So, the future of this area of research is really strong, and if we can improve the data, algorithms, and ethics involved, AI-based systems could play a major role in improving mental health outcomes for Gen Z. For instance, imagine having a system that can pick up on subtle changes in someone's behavior or mood, and then provide them with personalized support and resources to help them cope. This could be especially important for young people who are struggling with mental health issues, but might not know where to turn or how to ask for help. By using AI and data science to detect these issues early, we might be able to prevent more serious problems from developing down the line. And, by making these systems more explainable and trustworthy, we can help build confidence among healthcare professionals and the people using these systems. Overall, the possibilities here are exciting, and it's an area that's definitely worth exploring further. With the right approach, we could make a real difference in the lives of young people and help them build healthier, happier relationships with their mental health.

REFERENCES

1. R_D0.1: Mansoor, M. A., & Ansari, K. H. (2024). Early Detection of Mental Health Crises through Artificial- Intelligence-Powered Social Media Analysis: A Prospective Observational Study. Journal of Personalized Medicine, 14(9), 958.
2. R_D0.2: Wang, L., Wang, C., Li, C., Murai, T. Bai, Y. Song, Z. Zhang, S. Zhang, Q. Huang, Y., Bi, X., & Jiang, J. (2025). AI-assisted multi-modal information for the screening of depression: a systematic review and meta-analysis. npj Digital Medicine, 8, 523.
3. R_D0.3: (Not Specified). (2024). Screening for Depression Using Natural Language Processing. Journal of Medical Internet Research, 13, e55067.
4. R_D0.4: (Not Specified). (2022). Machine Learning on Early Diagnosis of Depression. Psychiatry Investigation, 19(8), 597–605.
5. R_D0.5: (Not Specified). (2025). Predicting depression in healthy young adults: A machine learning approach. NeuroImage.
6. R_D0.6: (Not Specified). (N/A). How to Present Study Limitations and Alternatives. Wordvice Blog.
7. R_D0.7: (Not Specified). (2024). Screening for Depression Using Natural Language Processing. JMIR, 13, e55067.
8. R_D0.8: (Not Specified). (N/A). Writing a Research Paper Conclusion. Scribbr.
9. R_D0.9: (Not Specified). (2022). A deep learning-based model for detecting depression
10. R_D0.3: (Not Specified). (2024). Screening for Depression Using Natural Language Processing. Journal of Medical Internet Research, 13, e55067
11. R_D0.4: (Not Specified). (2022). Machine Learning on Early Diagnosis of Depression. Psychiatry Investigation, 19(8), 597–605.
12. R_D0.5: (Not Specified). (2025). Predicting depression in healthy young adults: A machine learning approach. NeuroImage.
13. R_D0.6: (Not Specified). (N/A). How to Present Study Limitations and Alternatives. Wordvice Blog.
14. R_D0.7: (Not Specified). (2024). Screening for Depression Using Natural Language Processing. JMIR, 13, e55067.
15. R_D0.8: (Not Specified). (N/A). Writing a Research Paper Conclusion. Scribbr.
16. R_D0.9: (Not Specified). (2022). A deep learning-based model for detecting depression

17. R_D0.2: Wang, L., Wang, C., Li, C., Murai, T. Bai, Y. Song, Z. Zhang, S. Zhang, Q. Huang, Y., Bi, X., & Jiang, J. (2025). AI-assisted multi-modal information for the screening of depression: a systematic review and meta-analysis. npj Digital Medicine, 8, 523.
18. R_D0.2: Wang, L., Wang, C., Li, C., Murai, T., Bai, Y., Song, Z., Zhang, S., Zhang, Q., Huang, Y., Bi, X., & Jiang, J. (2025). AI-assisted multi-modal information for the screening of depression: a systematic review and meta-analysis. npj Digital Medicine, 8, 523.
19. R_D0.3: (Not Specified). (2024). Screening for Depression Using Natural Language Processing. Journal of Medical Internet Research, 13, e55067.
20. R_D0.4: (Not Specified). (2022). Machine Learning on Early Diagnosis of Depression. Psychiatry Investigation, 19(8), 597–605.
21. R_D0.5: (Not Specified). (2025). Predicting depression in healthy young adults: A machine learning approach. NeuroImage.
22. R_D0.6: (Not Specified). (N/A). How to Present Study Limitations and Alternatives. Wordvice Blog.
23. R_D0.7: (Not Specified). (2024). Screening for Depression Using Natural Language Processing. JMIR, 13, e55067
24. R_D0.8: (Not Specified). (N/A). Writing a Research Paper Conclusion. Scribbr.
25. R_D0.9: (Not Specified). (2022). A deep learning-based model for detecting depression