A Deep Learning Approach for Detecting and Mitigating Mental Health Conditions Through a User-centred Interactive Interface

Raveen Rajapaksha
Department of Computer Science and Informatics
Uva Wellassa University
Badulla, Sri Lanka
cst20080@std.uwu.ac.lk

Kavindu Buddhika
Department of Computer Science and Informatics
Uva Wellassa University
Badulla, Sri Lanka
cst20064@std.uwu.ac.lk

Kavindu Manahara
Department of Computer Science and Informatics
Uva Wellassa University
Badulla, Sri Lanka
cst20003@std.uwu.ac.lk

Nisal Thiwanka
Department of Computer Science and Informatics
Uva Wellassa University
Badulla, Sri Lanka
nisal@uwu.ac.lk

***Keywords*—AI in healthcare, deep learning, knowledge-based recommendation system, mental health detection, user-centred interface**

I. Introduction

Mental health disorders, such as depression, anxiety, attention deficit hyperactivity disorder (ADHD), bipolar disorder, personality disorder (PD), and post-traumatic stress disorder (PTSD), are growing global concerns that affect millions of individuals across all age groups and backgrounds. These conditions not only impair emotional and psychological well-being, but also hinder social relationships, academic performance, and professional productivity [1]. Deep neural networks have achieved high accuracy in detecting conditions from social media text and other user-generated content [2].

This research focuses on developing an AI-driven mental health framework comprising a user-centred interactive interface for intuitive engagement, a deep learning (DL) model for mental health status detection, and a knowledge-based recommendation system (KBRS) for adaptive, personalised interventions. With the DL model successfully developed and validated, the next phase focuses on implementing the user-centred interface and enhancing the KBRS.

A. Current Discussions and Research Gaps

DL has greatly advanced mental health research, with models that achieve high precision in detecting conditions such as depression and anxiety [3]. Hybrid models that combine textual and audio features further enhance the accuracy of the classification [4]. However, most DL-based systems focus on single-condition classification, making them less effective in detecting co-occurring disorders such as bipolar disorder [1] and personality disorder [5].

Another key gap is the lack of adaptive intervention mechanisms. Although KBRS has been applied in mental health, many remain static, unable to dynamically adapt to user inputs and changing conditions [6]. Studies show that integrating sentiment analysis with KBRS improves personalised support, but adaptive intervention mechanisms in real time remain underexplored [7]. In addition, many applications lack accessible and engaging interfaces that are tailored for people in distress. Research in Human-Computer Interaction (HCI) suggests that HCI features significantly improve engagement but remain underutilised in AI-based tools [8].

This research tackles these gaps by integrating DL-based diagnostics with an adaptive KBRS, supported by a user-centred interface. Unlike existing systems that focus solely on detection or intervention, this framework ensures a seamless transition from assessment to personalised care, bridging the gap between diagnosis and customised mental health support.

II. Materials and Methods

This research integrates principles from HCI, DL, and KBRS to develop an interactive system for the detection and intervention of mental health.

A. Research Design

This study employs a qualitative approach, using text-based data to refine diagnostic models and KBRS functionalities, while proportional stratified sampling ensures a

https://doi.org/10.31705/ADScAI.2025.09

diverse representation of mental health conditions. To ensure an effective deep learning model for detecting mental health conditions, multiple transformer-based models were trained and evaluated. The models explored in this study included BERT-base-uncased, DistilBERT-base-uncased, ELECTRA-base-discriminator, ELECTRA-base-generator, and MiniLM-L12-H384-uncased. Each model was trained using the same dataset and evaluated based on standardised performance metrics, including accuracy, precision, recall, and F1 score.

By comparing the predictive strengths and weaknesses of each model, a final selection was made based on the balance of efficiency, accuracy, and reliability. Complementing the model, a user-centred interface design with progressive disclosure, accessibility-focused features, and soothing colour schemes to accommodate individuals with mental health challenges. The KBRS component integrates decision-making theories and is enhanced with time-based rules and feedback loops to support personalised, adaptive, and context-sensitive interventions based on user progress and situational triggers.

B. Data Collection Methods

This study relies on publicly available secondary datasets to develop and evaluate the system.

- Sentiment Analysis for Mental Health Dataset[1]: Provides textual data with sentiment annotations, helping to identify emotional patterns and mental health indicators.
- Reddit-Mental-Illness-82 dataset[2]: Contains user-generated mental health related posts, categorised by conditions such as depression, anxiety, and PTSD.

The datasets were combined, modified, and preprocessed under the guidance of a psychotherapist to create a comprehensive dataset for training and validation, ensuring diverse linguistic and contextual coverage.

III. Results and Discussion

BERT-base-uncased demonstrated a strong classification capabilities, achieving an overall accuracy of 82.03%, but minor misclassifications in bipolar & personality disorders. DistilBERT-base-uncased, showed a slight improvement over its predecessor with an accuracy of 82.95% also efficient, and lower computational cost. ELECTRA-base-discriminator was also assessed and achieved an accuracy of 81.92%, slightly lower than both BERT and DistilBERT, and it also struggled with personality disorder classification. ELECTRA-base-generator exhibited a noticeable decrease in performance, achieving an accuracy of 79.01%, and misclassified personality disorder. MiniLM-L12-H384-uncased recorded the lowest overall performance with an accuracy of 75.64%.

The best-performing model, DistilBERT-base-uncased was selected based on standardised performance metrics. Following the fine-tuning process, as shown in TABLE I, the model demonstrated a high classification accuracy of 97.15%, particularly in detecting our targeted mental health disorders.

[1]https://kaggle.com/datasets/suchintikasarkar/sentiment-analysis-for-mental-health/

[2]https://huggingface.co/datasets/mavinsao/reddit-mental-illness-82/

TABLE I
Classification Report

Label	Precision	Recall	F1 score	Support
ADHD	0.98	0.97	0.97	1137
Anxiety	0.96	0.96	0.96	1241
Bipolar	0.96	0.97	0.96	1163
Depression	0.98	0.98	0.98	2120
Normal	0.99	0.98	0.99	1866
PTSD	0.95	0.96	0.96	1188
PD	0.96	0.97	0.97	1724

IV. Conclusion and Future Work

This research introduces an integrated framework that combines a user-centred interface, a DL model, and a KBRS to detect and Mitigate Mental Health Conditions. The DL model for mental health status detection has been successfully developed, and future work will focus on implementing the interactive interface and improving KBRS, ensuring a holistic AI-driven mental health support system.

References

[1] O. F. Kernberg and F. E. Yeomans, "Borderline personality disorder, bipolar disorder, depression, attention deficit/hyperactivity disorder, and narcissistic personality disorder: Practical differential diagnosis," *Bulletin of the Menninger Clinic*, vol. 77, no. 1, p. 1–22, Mar. 2013. [Online]. Available: http://dx.doi.org/10.1521/bumc.2013.77.1.1

[2] Vandana, N. Marriwala, and D. Chaudhary, "A hybrid model for depression detection using deep learning," *Measurement: Sensors*, vol. 25, p. 100587, Feb. 2023. [Online]. Available: http://dx.doi.org/10.1016/j.measen.2022.100587

[3] Z. Zhang, W. Lin, M. Liu, and M. Mahmoud, "Multimodal deep learning framework for mental disorder recognition," in *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, Nov. 2020, p. 344–350. [Online]. Available: http://dx.doi.org/10.1109/FG47880.2020.00033

[4] D.-A. Phan, Y. Matsumoto, and H. Shindo, "Autoencoder for semisupervised multiple emotion detection of conversation transcripts," *IEEE Transactions on Affective Computing*, vol. 12, no. 3, p. 682–691, Jul. 2021. [Online]. Available: http://dx.doi.org/10.1109/TAFFC.2018.2885304

[5] P. Asherson, A. H. Young, D. Eich-Höchli, P. Moran, V. Porsdal, and W. Deberdt, "Differential diagnosis, comorbidity, and treatment of attention-deficit/hyperactivity disorder in relation to bipolar disorder or borderline personality disorder in adults," *Current Medical Research and Opinion*, vol. 30, no. 8, p. 1657–1672, May 2014. [Online]. Available: http://dx.doi.org/10.1185/03007995.2014.915800

[6] R. L. Rosa, G. M. Schwartz, W. V. Ruggiero, and D. Z. Rodriguez, "A knowledge-based recommendation system that includes sentiment analysis and deep learning," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, p. 2124–2135, Apr. 2019. [Online]. Available: http://dx.doi.org/10.1109/TII.2018.2867174

[7] A. Hasan, R. Singh, S. Sanjana, S. Mudi, and S. Singh, "Knowledge based recommendation system including sentimental analysis and deep learning," *International Journal of HRM and Organizational Behavior*, vol. 12, no. 2, p. 50–56, Apr. 2024. [Online]. Available: https://ijhrmob.org/index.php/ijhrmob/article/view/179

[8] M. M. Dedhia, M. A. Hegishte, and M. A. Nair, "Recommendation system for e-learning based on personality type and learning style," *system*, vol. 9, no. 04, 2022.