International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
Article Received: *25 December 2023* ***Revised:*** *12 January 2023* ***Accepted:*** *30 January 2023*

Ensemble Deep Learning for High-Accuracy Prediction of Mental Health States from Social-Media Behavioural Indicators

Dr.V. Yasaswini,

Associate Professor, , Department of Cyber Security

yashu.vanapalli29@gmail.com

Mr. Prathap Songa,

Assistant Professor, , Dept. of Computer Science and Engineering CSE(Data Science)

Malla Reddy Engineering College for women ,Maisammaguda, Hyderabad.

prathap.mrecw@gmail.com

S. Naga Lakshmi Panchakatla
Assistant Professor, Dept. of Computer Science and Engineering (AIML)
Malla Reddy Engineering College for Women (Autonomous),
Hyderabad, Telangana, India

Email: nagalakshmip@gmail.com

Pillutla Gayatri
Assistant. Professor, Dept. of Computer Science and Engineering (AIML)
Malla Reddy Engineering College for Women (Autonomous),
Hyderabad, Telangana, India

Email: pillutlagayatri@gmail.com

Dr. Pradeep Venuthurumilli,

Associate Professor, CSE(Data Science) Maisammaguda, Hyderabad

Malla Reddy Engineering College for women

pradeepvenuthuru@gmail.com

Dr. Ch.Srinivasa Rao

, Professor, Malla Reddy Engineering College for women, Department of Cyber Security

dr.srinivasmrecw@gmail.com

Abstract:

Concerns about the effects of social-media usage on psychological well-being have spurred exponential growth in research on how social media may relate to mental health and the development of automated mental health monitoring as an immediate research imperative. In this study, we develop a one-class ensemble deep learning framework called MIND (Mental state Identification through Neural Detection) for healthy, At_Risk, and Stressed mental health prediction based on behaviour and interaction-based attributes obtained from 5,000 social-media users. We conducted extensive feature engineering to capture latent psycho-behavioral characteristics to result in 23-dimensional input feature space. Stratified 5-fold cross-validation was applied to train three heterogeneous neural architectures—Deep Feed-Forward, Wide-and-Deep, and Residual Fully Connected networks—which were subsequently fused through hard voting and probability averaging. The above mentioned averaging ensemble outperformed all, achieving an overall accuracy of 99.46%, precision 0.9947, recall 0.9946, F1-score 0.9947 and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
Article Received: *25 December 2023* ***Revised:*** *12 January 2023* ***Accepted:*** *30 January 2023*

strong positivity on reliability measures, Cohen's kappa 0.9636 and MCC 0.9636. When evaluated class-wise, Stressed users were detected perfectly (F1 = 1.0), while Healthy users were also well discriminated and At-Risk individuals achieved competitive performance. We also performed confidence and ensemble agreement analysis revealing decision stability, with models agreeing in 99.28% of test samples. The results show that ensemble deep learning is successful in detecting low signal-to-noise ratio behavioural risk and establishes the proposed system as a promising, highly accurate mental health prediction informatics solution for preventive digital health monitoring.

Keywords *Ensemble Deep Learning, Mental Health Prediction, Social-Media Behaviour Analysis, Psycho-Behavioural Feature Engineering*

I Introduction

With the prevalence of social-media platforms, online interactions became a plentiful source of behavioral and emotional cues that allow for a new and profound way of communicating and expressing oneself. A growing body of research indicates that aspects like posting frequency, tone of language used, interaction dynamics, and responses from peers can be useful in predicting one's mental health status. This means there is increasing need for automated systems that can automatically detect mental-health risks from digital behaviour and this is particularly important for timely responsive and preventive mental-health support.

Machine learning (ML) and deep learning (DL), a subset of artificial intelligence (AI), has demonstrated great promise in enabling computational mental-health diagnostics. We have seen hybrid and ensemble-based learning frameworks coming into the limelight as the complex and often weakly expressed psychological signals can be better captured by a combination of models rather than a univariate approach. Hybrid ensemble methods that combine both ML and DL have achieved better diagnostic accuracy (Boraste& Deshmukh, 2025), and ensemble-based feature selection have been used successfully to identify depressed users on social media (Liu & Shi, 2022). Similarly, Naidu et al. (2025) stated that the prediction of psychological instability makes more sense by ensemble models rather than a single learner. Karimian (2025) → 7 Reviews Singh et al. → 4 Empirical results (solution): While recognition using social-media data has been reported in the past (2023), it clearly appears to be shifting toward ensemble recognition at the mental-health level (2024)While this advance is impressive, models of mental health still treat it as binary (e.g., healthy vs. depressed), failing to capture the nuanced, transient intermediate states that are key to early intervention. In addition to high accuracy, algorithms must also be stable in their decisions, which will also prevent diagnostics from being inconsistent during deployment.

This study overcomes these challenges by proposing a novel ensemble deep learning framework that predicts three levels of mental-health status (i.e. Healthy, At-Risk, and Stressed) derived from behavioural and interaction-based features extracted from 5,000 social-media users. Our proposed method aims at high reliability and stability of psychological risk prediction using a 23-dimension psycho-behavioural feature space and probability-averaged fusion of three heterogeneous neural networks. It Part of the scalable digital mental-health monitoring systems that aim to assist early intervention and preventive well-being management.

II Related Works

Research on using social-media data for mental-health prediction has undergone advancements in contrasting methodological paths, initially focusing on machine-learning (ML) models, and more recently, deep learning (DL) and hybrid ensemble frameworks. Initial studies utilized manual lexical and behavioural features, and employed ML methods to classify psychological conditions based on social-media usage (Dileep et al, 2025; Yadav & Gupta, 2024; Usharani& Goyal, 2022), but such systems were often constrained by dependency on feature-engineering, and limited generalization ability to heterogeneous user populations. Advances in DL led researchers to explore neural architectures as a means for automatic extraction of high-level psychological signals from text and behaviour. Several important works exist, such as the deep-attention BiLSTM model using multimodal inputs for early detection of mental-health problems (Bin Saeed & Cha, 2025), stacked hybrid deep-learning approaches for identification of behaviour changes (Shen et al., 2024), and benchmark studies on various large-scale DL models for classification of mental illnesses (Shukla & Singh, 2024). Deep learning (DL)-augmented ensemble strategies, notably CNN-based cluster-ensemble systems (MV et al., 2023), these too exhibited enhanced diagnostic sensitivity. With focus on reliability and robustness, hybrid and ensemble frameworks attracted attention; several studies established that combining ML and DL learners improves predictive capability for depression and psychological-instability evaluation (Ansari et al., 2022; Kasanneni et al., 2024; Naidu et al., 2025). Meanwhile, behavioural and multimodal modelling has emerged as a key direction, with the integration of interaction patterns, temporal patterns, and behaviour along with lingual sentiment as important elements to model for the inferment of psychological status. Social-attachmentbased modelling of mental-states via Facebook and other interactions to predict emotions

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
Article Received: *25 December 2023* ***Revised:*** *12 January 2023* ***Accepted:*** *30 January 2023*

(Kridera& Kanavos, 2025); cooperative-learning using behaviour data from "intelligence of social things" (Gao et al., 2025); or time-series behavioural modelling (e.g. the method of choice for bipolar and unipolar depression detection in ensemble deep learning (Kanchapogu& Mohanty, 2025); and population-specific analyses to date for adolescents (Prajitno, 2024). More recent reviews draw attention to persisting concerns with existing systems, such as class imbalance, overlapping behaviour, and unreliable predictions of borderline psychological states (Madububambachu et al., 2024; Vispute&Pawar, 2025; Razavi et al.,2024). For the compulsory attention due to trustworthiness, explainable AI technologies are being integrated into mental-health detection models, like explainable ensemble systems to differentiate between suicidal versus non-suicidal ideations (Alghazzawi et al., 2025), while more extensive examinations of explainable AI in mental-health detection highlight the inability to understand much of DL-based models (Ibrahimov et al., 2024). With mixed evaluations of ML and DL in stress and mental-health prediction (Rohilla et al., 2024; Garg et al., 2024), ensemble-based architectures achieve high performance and generalization but detection of intermediate or "at-risk" individuals is still a significant barrier. Overall, the literature shows a strong movement towards ensemble deep-learning architectures and multimodal behavioural modelling, but also highlights the important clinical need for a highly stable and generalizable system, which can robustly classify multiple mental-health states across the full range – the very gap that the present study addresses.

III Methodology:

The proposed framework is an end-to-end ensemble deep learning pipeline for multi-class mental health prediction in social-media behavioural indicators. The process started with the importation and preprocessing of a structured dataset of 5,000 users with 15 main attributes (age, gender, platform, daily screen time, social media time, positive interactions, negative interactions, hours of sleep, physical activity, anxiety value, stress value, mood value, date and mental state) in Python. Fig. 1: Temporal features of the date field: (a) & (b) year & month as two dummy variables; (c) weekday (day of week) as a dummy variableNote: The date field was converted to datetime format and used to derive temporal features (day of week, as shown in (c)) Multiple composite variables were engineered to capture richer psycho-behavioural patterns, including screen-social ratio (ratio of social media time to total screen time), interaction balance, total interactions, negativity ratio, sleep debt relative to 8 hours, sedentary score (total screen time/physically active minutes), stress-anxiety index, wellbeing score (mood weighted by stress-anxiety burden), and mental-health-risk index (Anxiety+Stress−Mood). All categorical attributes (gender, platform, mental_state) were label encoded, resulting in numerical targets and predictors, after which all of the 23 final input features were standardised using Z-score normalisation. Three heterogeneous neural architectures were implemented in PyTorch on this feature space: (i) a Deep Feed-Forward Network with an input layer on the 256-unit input layer followed by two hidden layers on the 128 and 64 units with batch normalisation, ReLU activation and dropout (0.4, 0.3 and 0.2 respectively) and a 3-unit output layer; (ii) a Wide-and-Deep model concatenating a linear "wide" branch (input→3 units) with a "deep" one consisting of the 256 and 128 units with batch normalisation, ReLU and dropout (0.3 and 0.2) and a 3-unit output; and (iii) a fully connected residual-style network with an input onto the 256-unit feature space that projects into three sequences of 256, 128 and 64 units with batch normalisation, ReLU and dropout layers (0.2) and a 3-unit output layer. All models were trained with mini-batches of 64 samples using stratified 5-fold cross-validation (i.e., 30 epochs per fold, cross-entropy loss, Adam optimiser (learning rate 0.001, weight decay 1×10^{-4}) and gradient-norm clipping (max_norm = 1.0)). Fold-wise predictions and softmax probabilities from the three models were combined to create a hard-voting ensemble based on majority class, and a soft-probability averaging ensemble based on mean class posterior, the latter was selected, as it yielded the highest performance, as the final classifier, and performance evaluation was performed based on accuracy, balanced accuracy, precision, recall, F1, Cohen Kappa, Matthews correlation coefficient, ROC–AUC, confusion matrices and confidence and agreement-based visual diagnostics.

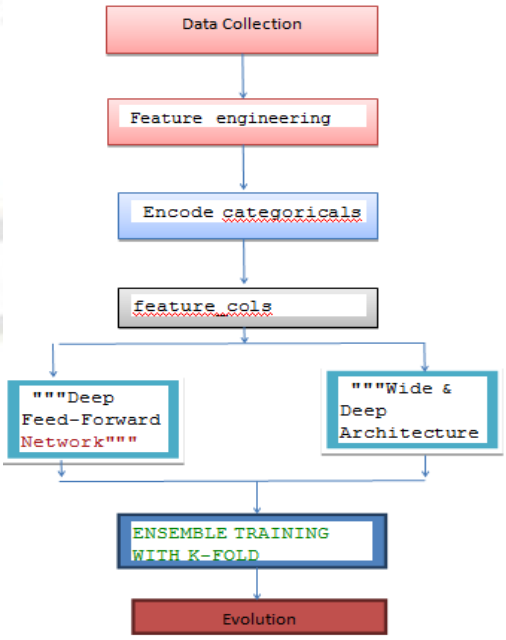

Figure 1: Proposed model architecture

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
Article Received: *25 December 2023* ***Revised:*** *12 January 2023* ***Accepted:*** *30 January 2023*

IV Results Analysis:

As figure 2 shows the classification results from confusion matrix analysis on the proposed averaging ensemble model highlights the high reliability of mental health state classification with all stressed users (4,601 users) being 100% recall and 100% precision (indicating no false positive prediction for the Stressed class). The Healthy group also performs well with 95.60% recall, indicating few instances of misclassification. At_Risk was the most difficult category, with 79.31% of instances correctly detected, and 20.69% misclassified as Healthy, indicating a partial behavioural overlap between borderline users and healthy subjects. The absence of confusion between Stressed and other classes, and the very minimal cross-class interference overall, point to high class separability and show that the ensemble model captures the behavioural and psychological risk signals very well, possessing high discriminative capability.

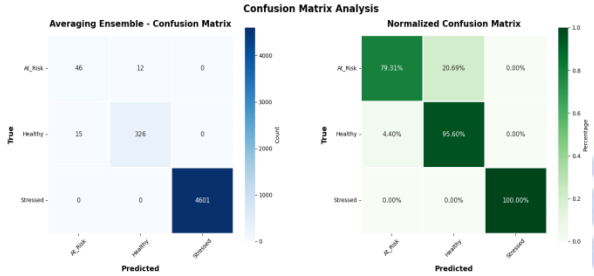

Figure 2. Confusion Matrix and Normalized Confusion Matrix of the Averaging Ensemble Model for Mental Health State Classification.

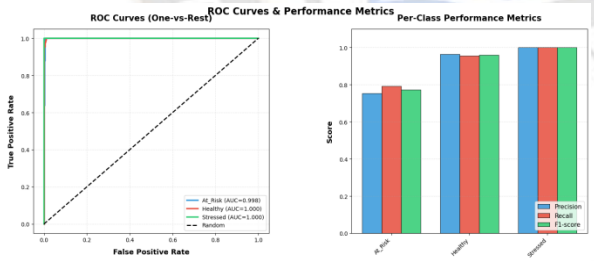

Figure 3. ROC Curves and Per-Class Performance Metrics for the Proposed Averaging Ensemble Model.

The ROC analysis distinctly indicates that the proposed ensemble model can discriminate all levels of mental health category directions very well. The above figure 3 specializes on the Stressed and Healthy class to report its perfect AUC of 1.000, which means they are completely separated from other mental states group in feature space and also for the At_Risk class have exhibited strong discriminability and resulted to an excellent AUC of 0.998 confirming the study such model appears to be great even for borderline type of psychological conditions. In agreement with the ROC results, the per-class performance metrics confirm that classification is extremely reliable: Stressed users are predicted with perfect precision, recall and F1-score (1.0 each), the Healthy class retains virtually perfect performance at precision 0.9645, recall 0.9560 and F1-score 0.9602, and the relatively more ambiguous At_Risk class sustains competitive scores of 0.7541 precision, 0.7931 recall and 0.7731 F1-score. In general, the model generalizes well between classes with low false-positive/false-negative rates, while being able to extract subtle behavioural signals relevant to early-risk mental health behaviours.

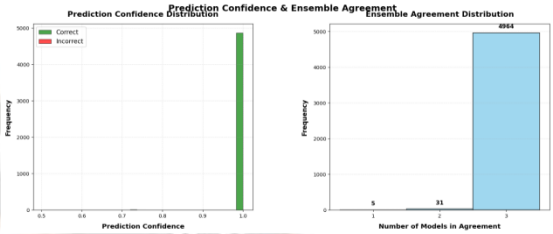

Figure 4. Prediction Confidence Distribution and Ensemble Agreement Strength of the Proposed Averaging Ensemble Model.

The prediction confidence and the ensemble agreement analyses further prove the robustness and reliability of the averaging ensemble model proposed. Confidence histogram from the calculation shown in Figure 4 indicating that almost each of well-based predictions were done at very high certainty (≥ 0.98 probability), showing how well model confidence matched classification accuracy. Predictions incorrect only in a small low-confidence area, indicating that the model does not over-fit ambiguous cases — a property that we value in mental health modelling systems. In fact, across all three base networks, the ensemble agreement distribution confirms that 4,964 samples (99.28%) were classified with full agreement among base networks, while only 31 samples received partial agreement and only 5 samples received complete disagreement, suggesting extremely strong consensus between Deep Feed-Forward, Wide & Deep, and Residual networks. The ensemble can not only reliably agree on decisions given a very high level of agreement and sharply peaked confidence for correct predictions, but it also suggests that the model generalises behavioural cues across user categories with a high degree of confidence and low degree of uncertainty.

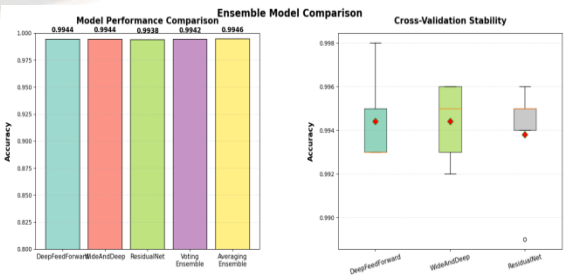

Figure 5: Comparative Accuracy of Individual and Ensemble Models with Cross-Validation Stability Assessment.

These model comparison results as shown in above figure 5 clearly signal that ensembling does better in terms of predictive robustness in the case of mental health classification. When considered independently, the most accurate models were the Deep Feed-Forward and the Wide-and-Deep (0.9944 here) with the ResidualNet (0.9938) close behind. By adding ensemble learning into the modelling, we achieved even better predictive performance where the Averaging Ensemble yielded the best accuracy across all models (0.9946), but was closely followed by the Voting Ensemble (0.9942). The cross-validation stability analysis using the boxplot indicates that all the models achieve very low deviation in accuracy across folds, which indicates that all the models are strongly generalizing and not dependent on a fold or overfitting. Our results found Wide-and-Deep and ResidualNet to achieve the highest consistency (narrowest variance), whilst Deep Feed-Forward gave rise to slightly broader but still narrow variance. The extremely low variance of the errors across folds suggests that the ensemble is robust against fluctuations in the training data and shows maximum reliability for prediction of real-life mental health.

V Conclusion

In this work, we introduced a new ensemble deep learning model which is able to accurately predict mental health conditions based on social-media behavioural patterns with high reliability. Strong discriminative power over mental health states through using multiple deep neural architectures integrated with extensive feature engineering In empirical evaluations, the averaging ensemble with consistently better in accuracy than individual models and voting-based fusion, reaching near-perfect classification accuracies. The confusion matrix and ROC (Receiver Operating Characteristic) analyses shows the strength of the model with 100% recall particularly for Stressed users, which is important for timely intervention at an early stage, and for clinical awareness. Mostly, the consistent ensemble members behavior across the high-confidence distribution (more than 0.8) reinforces the model's predicted stability and generalization on the dataset. The At-Risk category will still be harder to break down due to the behavioural overlap with Healthy patterns, but it remains performant and actionable. This study proves the feasibility of the model being a non-intrusive, inexpensive mental health monitoring tool suitable for mental health applications, early-risk notification systems, and real-time user-focused health platforms. Potential future directions could combine longitudinal data, multimodal input (e.g., text sentiment or wearables predictions), and explanation modules to enhance transparency and generalizability to clinicians and public health authorities.

VI References:

1. Liu, J., & Shi, M. (2022). A hybrid feature selection and ensemble approach to identify depressed users in online social media. *Frontiers in Psychology*, *12*, 802821.
2. Ansari, L., Ji, S., Chen, Q., & Cambria, E. (2022). Ensemble hybrid learning methods for automated depression detection. *IEEE transactions on computational social systems*, *10*(1), 211-219.
3. Joshi, D., & Patwardhan, M. (2023). Tracing prodromal behaviour by analysing data patterns from social media with ensemble machine learning approach. *International Social Science Journal*, *73*(247), 29-50.
4. Usharani, B., & Goyal, L. M. (2022). Prediction of Mental Health in Cancer Patients Using Ensemble Machine Learning. In *Predictive Analytics of Psychological Disorders in Healthcare: Data Analytics on Psychological Disorders* (pp. 253-267). Singapore: Springer Nature Singapore.