Int. J. Appl. Math. Comput. Sci., 2024, Vol. 34, No. 4, 565–577
DOI: 10.61822/amcs-2024-0038

A DEEP LEARNING BASED HYBRID MODEL FOR MATERNAL HEALTH RISK DETECTION AND MULTIFACETED EMOTION ANALYSIS IN SOCIAL NETWORKS

R. GEETHANJALI [a,*], A. VALARMATHI [b]

[a]Faculty of Information and Communication Engineering
Anna University
UCE-BIT Campus, Mandaiyur, Tiruchirappalli 620 024, Chennai, India
e-mail: geethanjaliyokesh18@gmail.com

[b]Department of Computer Applications
Anna University
UCE-BIT Campus, Mandaiyur, Tiruchirappalli 620 024, Chennai, India
e-mail: valar1030@yahoo.com

In the field of public health, accurately identifying maternal health risks through social network data is both vital and challenging due to the complexities of multimodal sentiment analysis. Our study addresses this challenge by introducing the maternal health risk factor detection using deep learning approach (MHRFD-DLA), a novel framework that integrates convolutional neural networks, long short-term memory networks, and attention mechanisms. This approach enhances sentiment analysis and risk detection in maternal health, with the focus on critical areas such as prenatal care, mental health, and nutrition. MHRFD-DLA utilizes multimodal data, including text and electrocardiogram (ECG) signals, offering a comprehensive assessment of maternal health risks. Our model outperforms existing multimodal sentiment analysis models, achieving an accuracy of 98.4%, a precision of 97.6%, a recall of 95.6%, and an F1 score of 98.4%. Through performance evaluations, visualizations such as the confusion matrix and class distributions further validate its robustness. The MHRFD-DLA model not only bridges significant gaps in current methodologies, but it also sets a new benchmark for maternal health surveillance and intervention, demonstrating its practicality and effectiveness in real-world applications.

Keywords: multifaceted emotion analysis, social networks, maternal health risk factor detection, deep learning, hybrid approach.

1. Introduction

Understanding the varied material published on social networks, which includes both text and auditory information, is greatly assisted by multimodal sentiment analysis (Gupta *et al.*, 2021). By considering a variety of inputs, rather than just text, this analytical method provides a more complete picture of user sentiment than is possible with conventional sentiment assessment (Afyouni *et al.*, 2022). Understanding the complexities of user-generated material is crucial in the ever-changing world of social media, as it provides insight into people's multifaceted emotions (Nazir *et al.* 2022). Regarding online conversations around mother fitness, multimodal sentiment analysis is a lifesaver (Ghosh *et al.*, 2023). Addressing issues and increasing monthery well-being requires understanding the views expressed in these talks (Geethanjali and Valarmathi, 2022). A sophisticated comprehension of the feelings expressed in these conversations is necessary to identify possible risk factors for maternal health (Zhang *et al.*, 2020).

Understanding user attitudes allows for the prompt implementation of treatments to tackle new issues and reduce risks, ultimately resulting in better maternal outcomes (Acharya *et al.*, 2023). Communities, healthcare providers, and lawmakers may all benefit from sentiment analysis since it goes beyond simple observation and gives them the tools they need to

*Corresponding author

customize solutions (Oueslati *et al.*, 2020; García-Díaz *et al.*, 2020).

Understanding the sentiment expressed in social network data is crucial for addressing maternal health challenges effectively. Multimodal sentiment analysis, which considers various inputs like text and auditory information, provides a comprehensive understanding of user sentiment, essential in the dynamic landscape of social media. Analyzing user-generated content helps in grasping multifaceted emotions, especially pertinent in discussions concerning maternal fitness. By comprehending the sentiments conveyed in these conversations, potential risk factors for maternal health can be identified promptly, enabling timely interventions and improved outcomes. Such insights empower communities, healthcare providers, and policymakers to tailor solutions effectively, transcending mere observation and fostering proactive approaches towards maternal well-being.

Furthermore, the nuanced understanding gleaned from sentiment analysis aids in the customization of interventions to address specific needs within maternal health. By delving deeper into the sentiments expressed across social networks, stakeholders gain valuable insights into the concerns, challenges, and aspirations of mothers. This enables the development of targeted strategies aimed at improving maternal outcomes and promoting overall well-being. Moreover, the timely identification of emerging issues through sentiment analysis facilitates the rapid implementation of responsive measures, fostering a proactive approach to maternal healthcare. Ultimately, leveraging sentiment analysis in maternal health not only enhances our understanding of maternal experiences but also empowers us to enact meaningful change for the betterment of maternal and child health outcomes.

1.1. Problem statement. This study delves into the complex landscape of social networks, aiming to decipher maternal health risks and sentiments effectively. It introduces a cutting-edge hybrid deep learning model, amalgamating convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and attention mechanisms. This model provides a holistic approach to sentiment analysis, considering both textual and auditory data, thereby offering a more nuanced understanding of maternal well-being within online communities. By outperforming existing methods, it not only enhances the detection of maternal health risks but also facilitates timely interventions, ultimately contributing to improved maternal outcomes.

1.2. Motivation of the work. The motivation behind this work stems from the pressing need to address maternal health risks effectively in the digital age. With the proliferation of social networks, there is a wealth of data available, presenting an opportunity to gain insights into maternal well-being and potential risks. However, existing methods often struggle to accurately analyze the multifaceted emotions and sentiments expressed in these networks, hindering timely interventions. Therefore, the study seeks to pioneer a comprehensive approach that leverages advanced deep learning techniques to bridge this gap. By integrating CNNs, LSTMs, and attention mechanisms, the aim is to develop a model capable of not only detecting maternal health risks, but also understanding the nuanced sentiments surrounding them. Ultimately, the goal is to empower healthcare providers, communities, and policymakers with actionable insights to improve maternal outcomes and support networks.

1.3. Technical gap. The technical gap addressed by this study lies in the inadequacy of existing methods to effectively analyze multifaceted emotions and sentiments in social network data related to maternal health. Traditional approaches often struggle to capture the nuanced expressions of maternal well-being, hindering accurate risk detection and timely interventions. By pioneering a comprehensive deep learning approach integrating CNNs, LSTMs, and attention mechanisms, this study aims to fill this gap by developing a more sophisticated model capable of understanding and analyzing complex sentiments surrounding maternal health risks in digital platforms.

1.4. Contribution to the work. A hybrid CNN LSTM attention model has been implemented to tackle the intricate issue of multimodal sentiment assessment within the framework of maternal fitness threat detection (Rahman and Alam, 2020). More study and development are required for these methods to reach their maximum potential and improve motherly well-being in online communities. Here are specific objectives to be attained:

- With a focus on identifying risks to maternal health, create an effective method for multimodal sentiment analysis in social networks. A more sophisticated comprehension of the ever-changing nature of user-generated content on social media platforms can be achieved by implementing a CNN + LSTM + attention hybrid model, which expertly combines visual and textual data.

- To monitor and detect maternal health risk indicators across several social media platforms in real-time, implement the developed MHRFD-DLA model. Facilitate the prompt detection of feelings about variables impacting a mother's health, such as prenatal care, diet, and mental health.

- To ensure the suggested MHRFD-DLA model is accurate, run comprehensive simulation studies with various social network datasets. Measure its efficacy compared to preexisting methods by calculating its F1 score, accuracy, and recall. Exhibit the model's strength and effectiveness in detecting intricate sentiment patterns linked to factors that influence maternal health.

1.5. Structure of the article. The literature review is presented in Section 2. Section 3 contains the mathematical study of the maternal health risk factor detection using a deep learning approach (MHRFD-DLA). Results and discussion are presented in Section 4, Future research directions presented in Section 5. A brief overview and some concluding recommendations are included in Section 6.

2. Literature review

Gopalakrishnan *et al.* (2023) developed the attribute selection hybrid network (ASHN) to identify postpartum depression (PPD) from social media posts. ASHN analyzes characteristics of Facebook postings, such as tone and reflective thinking, and effectively extracts psycho-linguistic signals. It outperforms baseline methods in detecting signs of depression in online posts using metrics like F1 score, recall, and precision.

Nadeem *et al.* (2022) introduced the sequence, semantic, context learning (SSCL) framework for depression diagnosis using a Twitter dataset. This deep learning method combines LSTM, CNN, GRUs, and self-attention, achieving 82.9% accuracy for ternary labels and 97.4% for binary labels. Its 94.4% F1-score on unseen data and superior performance in sarcasm detection highlight its robustness.

Shoumy *et al.* (2020) enhanced affective computing by integrating physiological modalities with text, audio, and visual data. They address limitations in anti-spoofing and environmental resilience. Cimtay *et al.* (2020) developed a multimodal emotion identification system combining EEG, facial expressions, and galvanic skin response (GSR), achieving up to 91.5% accuracy on the DEAP dataset and 81.2% on a custom dataset.

Togunwa *et al.* (2023) proposed a hybrid deep learning model combining artificial neural networks and random forest algorithms for maternal health risk classification. The model shows promising results in improving health outcomes for pregnant women and infants. Lilhore *et al.* (2024) introduced a hybrid framework for detecting postpartum depression using text and audio data, demonstrating better performance than existing models.

Nti and Owusu-Boadu (2022) utilized machine learning to predict maternal health risks, with gradient boosting showing the highest accuracy. Explainable AI methods like LIME and SHAP provided insights into prediction rationales. Acharya *et al.* (2023) developed a hybrid boosting ensemble model, integrating XGBoost and CatBoost, to predict maternal mortality risk, offering valuable insights for maternal healthcare, especially in developing countries.

Gupta and Sharma (2023) emphasized early detection of postpartum depression (PPD) due to its impact on maternal and infant well-being. They review current detection techniques, noting the limitations of questionnaire-based tools and highlighting the effectiveness of support vector machine classifiers and EEG-based methodologies.

Byeon (2023) introduced SSDD for depression detection on social media using semantic representation and semi-supervised deep learning. SSDD combines hybrid feature analysis and deep autoencoders with Bi-LSTM for detecting depressive content on platforms like Twitter, outperforming existing models.

Teisseyre (2022) explored the role of explainable AI (X-AI) in mental health assessments for depression. The study reviewed methods using SHAP and LIME, emphasizing the need for transparent AI predictions to improve depression prediction and treatment. Feature selection methods using weighted penalized empirical risk minimization, demonstrating effectiveness in handling unobserved variables and improving predictive accuracy.

3. Proposed methodology: Multifaceted emotion analysis for maternal health risk detection

3.1. Operating environment. Utilizing Google Colab as the platform, our proposed methodology focuses on multifaceted emotion analysis for maternal health risk detection. Leveraging the collaborative and cloud-based features of Colab, we integrate deep learning techniques to analyze diverse emotional expressions in maternal health discussions on social networks. This approach enables real-time processing of textual and audio data, facilitating the detection of nuanced sentiments and potential health risks among pregnant women.

3.2. Dataset descriptions. In this research, two datasets were utilized to conduct sentiment analysis on different modalities: the `maternal health risk data` dataset was employed for text sentiment analysis, while the `heartbeat data` was utilized for audio sentiment analysis. The integration of these datasets allowed for a comprehensive analysis of sentiments related to maternal health, encompassing both textual and auditory aspects, thereby providing a more holistic understanding of maternal well-being in the context of healthcare.

3.2.1. Text SA dataset. The `maternal health risk data` dataset contains essential health attributes of pregnant patients, such as age, blood pressure, blood glucose levels, and heart rate, aiming to predict maternal risk intensity during pregnancy. It enables researchers to develop predictive models for maternal health risks and improve prenatal care (Geethanjali, 2024).

3.2.2. Audio SA dataset. `Heartbeat data` includes segmented and preprocessed electrocardiogram (ECG) signals from multiple sources. It features normal and abnormal heartbeat patterns, facilitating research in heartbeat classification using deep learning and transfer learning techniques for cardiovascular health assessment and diagnostic advancements (Fazeli, 2024).

3.3. Data preprocessing. In the preprocessing phase of text sentiment analysis using the `maternal health risk data` dataset, tasks included text cleaning (removing special characters and punctuation), tokenization (splitting text into words), stopword removal (eliminating common but non-informative words), and stemming or lemmatization (normalizing words to their root forms). The processed text was then encoded into numerical formats, such as one-hot encoding or word embeddings, for deep learning models.

For audio sentiment analysis with the `ECG heartbeat categorization dataset` pre-processing involved signal processing to remove noise, segmentation to divide ECG signals into heartbeat segments, and feature extraction to capture characteristics like amplitude and frequency. The extracted features were normalized, and the data were converted into numerical formats, such as feature vectors or spectrograms, for input into deep learning models.

3.4. Feature extraction. For text sentiment analysis using the `maternal health risk data` dataset, feature extraction involves converting textual data into numerical forms suitable for deep learning. Techniques include bag-of-words (BoW), which represents documents as vectors based on word frequencies; TF-IDF (term frequency-inverse document frequency), which weights terms according to their importance in individual documents and across the corpus; and word embeddings, which encode contextual relationships between words in dense vector spaces, capturing semantic information.

In audio sentiment analysis with the `heartbeat data` feature extraction focuses on capturing key characteristics of ECG signals. Methods such as mel-frequency cepstral coefficients (MFCCs) are used to extract the spectral envelope, providing insights into acoustic features like pitch. Spectrograms are also utilized to represent the frequency content of the ECG signals over time, allowing visualization of changes in frequency components. Combining MFCCs and spectrograms offers a comprehensive understanding of the audio data's acoustic properties, enhancing the deep learning analysis.

3.5. Proposed MHRFD-DLA model. Our maternal health risk factor detection using a deep learning approach (MHRFD-DLA) employs advanced sentiment analysis on social media to identify maternal health risks. By integrating CNNs and LSTMs with an attention mechanism, MHRFD-DLA forms a hybrid framework for analyzing multimodal data, including text and audio. This approach enhances sentiment assessment for early detection of risk factors in maternal health by processing and analyzing social media content. The hybrid CNN + LSTM + attention model allows for nuanced understanding of evolving user-generated content, improving real-time tracking and detection of maternal health risk indicators.

The model leverages multimodal fusion to integrate textual and audio data, enhancing its ability to detect complex emotional correlations. This hybrid approach provides a more comprehensive view than unimodal methods by combining sentiments from both modalities to identify maternal health risk factors.

Figure 1 illustrates the detailed multimodal sentiment analysis process used for early identification and intervention in maternal health issues, aiming to deliver deeper insights through the fusion of text and audio data. The accuracy of a CNN for detecting maternal health risks is assessed using

$$\mathrm{Acc} = \frac{1}{O}\sum_{j=1}^{O}\Big(e^{-\vartheta_{\mathrm{DF}}\left(P_{\mathrm{CNN}}^{(j)},Z^{(j)}\right)\times\alpha\left(P_{\mathrm{CNN}}^{(j)},Z^{(j)}\right)} - \lambda S\Big). \tag{1}$$

This equation evaluates performance by comparing the model's predictions $P_{\mathrm{CNN}}^{(j)}$ with the true labels $Z^{(j)}$ using a cross-entropy loss. An exponential function scales the loss to penalize larger discrepancies, while a regularization term, controlled by λ, prevents overfitting by penalizing complex models. The Kronecker delta function α contributes to the accuracy calculation by assigning a value of 1 if the prediction matches the label and 0 otherwise. This comprehensive approach integrates loss, regularization, and prediction accuracy to measure the CNN's effectiveness in identifying maternal health risk factors.

Equation (2) for precision evaluates how well the model identifies maternal health risk factors. Precision Q is calculated using correct positive forecasts (UQ), true negative predictions (GO), and true positives (GQ). It

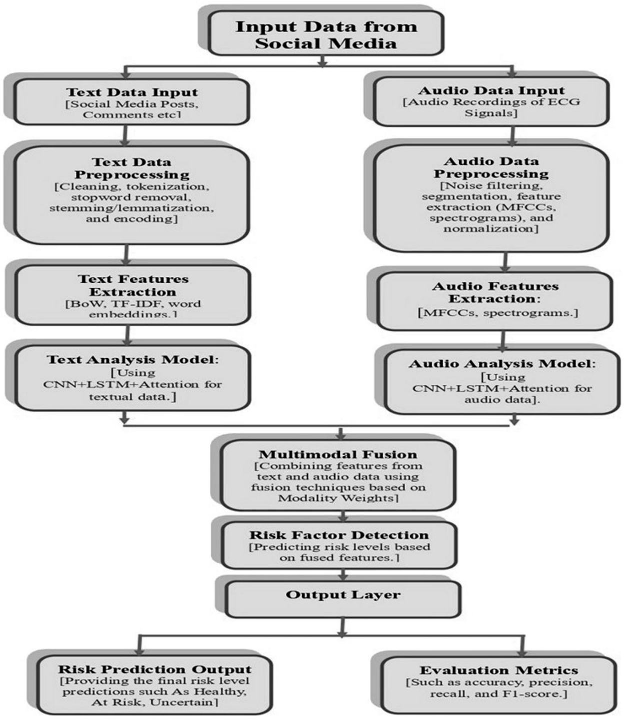

Fig. 1. Proposed MHRFD-DLA architecture for multimodal sentiment analysis to identify risk factors in maternal health.

incorporates a weighting factor γ to balance recall and precision, while δ and α account for additional feature contributions and impacts. Regularization parameters β and λ manage model complexity, influencing true positives and false negatives through nonlinear functions and offset terms (σ and ε), and false positives through $\varnothing$. A small constant ϖ prevents division by zero. This detailed precision equation integrates multiple variables and their interactions, highlighting the value of combining audio and text data in social media sentiment analysis for maternal health assessment, as shown in Fig. 1. Equation

$$\begin{aligned} d_u &= f_u \times d_{u-1} \\ &+ j_u \times \frac{\tanh(X_{jd} \times y_u + c_{jd} + V_{jd} \times i_{u-1})}{\sigma(W_{jd} \times y_u + d_{jd} + U_{id} \times i_{u-1})} \end{aligned} \tag{3}$$

updates the state of an LSTM cell using several key variables. The current cell state d_u and the previous state d_{u-1} are updated by the forget gate output f_u and influenced by the input gate j_u. The input gate involves weight matrices X_{jd} and V_{jd}, while y_u represents the time step and c_{jd} is the bias term. The nonlinear transformation is given by $\tanh(X_{jd} \times y_u + c_{jd} + V_{jd} \times i_{u-1})$, where W_{jd} and d_{jd} are input weight matrices and biases, and U_{id} is the hidden state weight matrix. The sigmoid function σ scales this transformation to adjust the cell state. Each variable contributes to the LSTM cell's ability to manage temporal patterns in input data. In

$$T = \sum_{j=1}^{o} \frac{e^{(X_2 \times \tanh x_1[i_u;\tilde{i}_j]+c_1)+c_2}}{\sum\limits_{k=1}^{o} e^{X_2 \times \tanh(X_1 \cdot [i_u;\tilde{i}_k])+c_2}} \times \tilde{i}_j \tag{4}$$

attention weights manage the relevance of each detail, where T represents the weighted sum of transformed input elements $\tilde{i}_j$ in a sequence. A multi-layer perceptron (MLP) with learnable parameters X_1, X_2, c_1, and c_2 computes these weights. At each time step, the hidden state i_u is combined with the transformed input element $\tilde{i}_j$ to form $[i_u;\tilde{i}_j]$. The hyperbolic tangent function $\tanh$ adds nonlinearity, and the attention weights are normalized using softmax to sum to one, reflecting each element's relative importance in the sequence.

3.6. Attention model. The attention model in the proposed MHRFD-DLA algorithm dynamically weighs the relevance of different modalities and segments based on learned attention scores. By assigning attention scores to each element in the sequence, it allows the model to focus on pertinent information during sentiment analysis for maternal health risk factor detection. Through nonlinear activation functions and bias terms, the attention mechanism enhances the model's ability to capture subtle sentiment patterns across multiple modalities in social network data related to maternal health. In

$$S = \frac{\sum\limits_{j=1}^{O} \left(\frac{UQ_j \times \text{Attention}_j}{\sqrt{2 \times \text{Attention}_j}} \right)}{\sum\limits_{j=1}^{O} \left(UQ_j + GO_j + \exp\left(\frac{\text{Attention}_j^2}{\max(\text{Attention})} \right) \right)} \tag{5}$$

the recall metric S assesses the MHRFD-DLA model's ability to detect maternal health risk factors. The total dataset occurrences are denoted by O, with UQ_j representing true positives for the j-th sample and GO_j indicating false negatives. The attention mechanism adjusts modality weights dynamically, with Attention_j^2 showing the weight for the j-th instance.

The risk factor detection output, the final block, reflects the model's prediction on maternal health risks after analyzing social media data. This output is crucial for continuous monitoring and early risk identification, leveraging multi-source and multi-modal information. Algorithm 1 shows the algorithmic steps for the MHRFD-DLA procedure in multimodal sentiment analysis for maternal health risk detection.

Equation (6) describes the attention mechanism for determining modality weights in sentiment analysis. At the j-th time step, the weight β_{jk} reflects the importance of the k-th modality. The integration of textual features and modalities is facilitated by weight matrices W_1 and W_2, with a nonlinear activation function ϕ enhancing

$$Q = \frac{(1+\gamma^2)\cdot\sigma(UQ+\rho)+\delta\cdot\alpha}{(1+\gamma^2)\cdot\sigma(UQ+\rho)+\gamma^2\cdot\varnothing(GO+\eta)+\varepsilon\cdot GQ+\varpi+\beta\cdot\lambda}. \tag{2}$$

$$\beta_{jk} = \frac{\exp\left(\varnothing\left(X_1 i_{j-1}+X_2 i_k+c_1\right)\phi\left(W_1 f_{j-1}+W_2 f_k+C_2\right)\right)}{\sum_{l=1}^{U}\exp\left(\varnothing\left(X_2 i_{j-1}+X_2 i_l+c_1\right)\phi\left(W_2 f_{j-1}+W_2 f_l+C_2\right)\right)}. \tag{6}$$

Algorithm 1. Proposed MHRFD-DLA algorithm for multimodal sentiment analysis.

Input:
- **text_model** : The CNN+LSTM+Attention model for text sentiment analysis.
- **ecg_model** : The CNN+LSTM+Attention model for ECG sentiment analysis.
- **num_classes**: The number of classes for final prediction.

Output:
- **fusion_model**: The combined CNN+LSTM+Attention model for joint sentiment analysis.

Begin

Create Fusion Model
Step 1: Retrieve the output of the CNN+LSTM+Attention text model.
Step 2: Retrieve the output of the CNN+LSTM+Attention ECG model.
Step 3: Concatenate the two outputs along the specified axis.
Step 4: Connect the concatenated outputs to a fully connected layer with softmax activation for the final prediction.
Step 5: Create and compile the fusion model using the concatenated structure.
Step 6: Return the compiled CNN+LSTM+Attention **fusion_model**.

Train Fusion Model
Step 7: Compile the **fusion_model** with an optimizer and categorical cross-entropy loss function.
Step 8: Fit the model to the provided text and ECG data along with ground truth labels.
Step 9: Train the model for the specified number of epochs using the given batch size and validation split.

Predict Fusion
Step 10: Use the trained CNN+LSTM+Attention **fusion_model** to predict sentiment jointly.
Step 11: Provide text and ECG input.

Display Result
Step 12: Display **joint_prediction**.

End

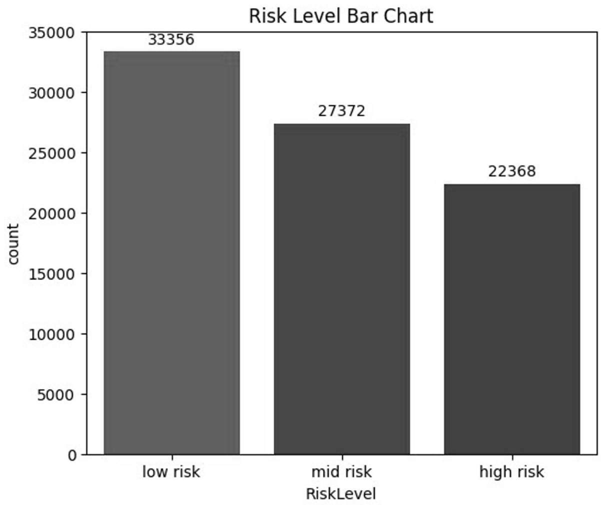

Fig. 2. Risk level bar chart.

pattern detection. Visual data is also integrated, with hidden states f_{j-1} and f_k, and corresponding weights X_1 and X_2, using ϕ for nonlinearity and bias term C_2 for adjustment. This mechanism allows the model to dynamically prioritize modalities, capturing detailed sentiment patterns in social media data related to maternal health risks. Each of the variables in the attention rating equation

$$b_j^i = \frac{\exp\left(\sigma\left(X_u^i \times i_u + V_u^i + W_u^i \times w_j + c_u^i\right)\right)}{\sum_{k=1}^{U}\exp\left(\sigma\left(X_u^i \times i_k + V_u^i + W_u^i \times w_k + c_u^i\right)\right)} \tag{7}$$

is essential to the attention mechanism for multimodal sentiment evaluation. Here b_j^i represents the attention score that the i-th attention unit assigns to the j-th feature within the sequence. The nonlinear activation function σ introduces complex interactions. The model provides an additional degree of flexibility with the inclusion of c_u^i as a bias term for the i-th attention head and U as the total number of segments. The algorithm can focus on pertinent information during sentiment evaluation for social network threat detection related to maternal health because (7) dynamically weighs the importance of each modality and segment based on these factors.

4. Results and discussion

The efficacy of the MHRFD-DLA model is highly dependent on the precision and accuracy of the sentiment analysis and risk factor detection. Using a combination of CNNs, LSTMs, and an attention mechanism, MHRFD-DLA, seeks to understand emotions and detect any dangers to the health of mothers.

4.1. Analysis and discussion of text SA results. The analysis of text sentiment analysis results reveals significant predictive capabilities, with the model accurately classifying maternal health risk levels based on extracted features. This highlights the potential of leveraging textual data for effective risk assessment in maternal healthcare contexts.

Table 1 reflects the textual sentiment analysis which lists variables `Age`, `SystolicBP`, `DiastolicBP`, `BS`, `BodyTemp`, `HeartRate`, and `RiskLevel`. The age limit ranges from 22 to 55 followed by the systolic BP from 90 to 140. Diastolic BP ranges between

Table 1. Textual sentiment analysis EDA.

S.no.	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate	RiskLevel
0	25	130	80	15	98	86	high risk
1	35	140	90	13	98	70	high risk
2	29	90	70	8	100	80	high risk
3	30	140	85	7	98	70	high risk
4	35	120	60	6.1	98	76	low risk
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
83106	30	140	100	15	98	70	high risk
83107	31	120	60	6.1	98	76	mid risk
83108	23	120	90	7.9	98	70	mid risk
83109	29	130	70	7.9	98	78	mid risk
83110	17	85	60	7.9	102	86	low risk

Table 2. Comparison of text sentiment analysis models.

Model	Accu (%)	Prec (%)	Recall (%)	F1 score (%)
CNN	86.45	87.90	86.45	86.17
LSTM	89.54	88.95	89.54	89.24
CNN+ LSTM	91.43	92.02	91.43	91.72
CNN+LSTM+ attention	93.00	92.77	93.00	92.38

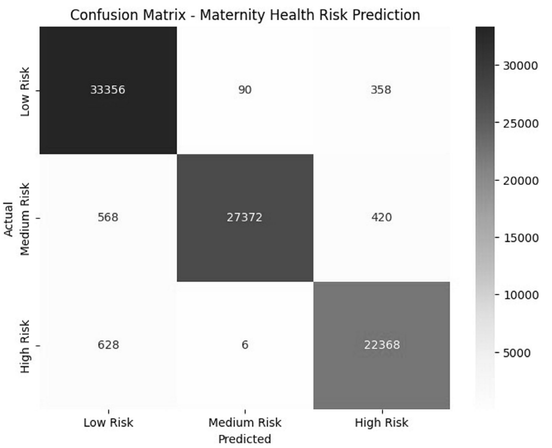

Fig. 3. Confusion matrix for maternity health textual data.

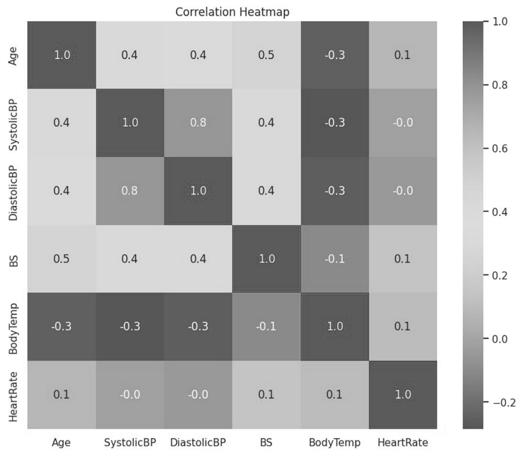

Fig. 4. Correlation heatmap for maternity health textual data.

60 to 90 and risk levels as low, mid, and high risks. Other variables are mentioned in the image.

Figure 2 shows the risk level bar chart illustrating the low, mid, and high-risk levels. The x-and y-axes of the graph represent the count and the risk level. The low risk has a count over 33356, followed by the mid-level risk at 27372. The high-level risk was identified to be lower at 22368, which is less than the high- and mid-level risks. Figure 3 shows the maternity health risk confusion matrix visualized with categories for low, medium, and high-risk stages. Figure 4 presents a correlation heatmap for maternity health data, categorizing risk into low, mid, and high levels. It visually depicts correlations between age, heart rate, body temperature, blood sugar (BS), diastolic blood pressure (BP), and systolic blood pressure (BP) across these risk categories. Colors denote negative to positive correlations, aiding in understanding health parameter relationships across risk levels.

Table 2 compares the accuracy, precision, recall and F1 score of various text SA models. Compared with other models, the CNN + LSTM + attention model performs better on all measures.

4.2. Analysis and discussion of audio SA results. The analysis of audio sentiment analysis results demonstrates robust classification performance,

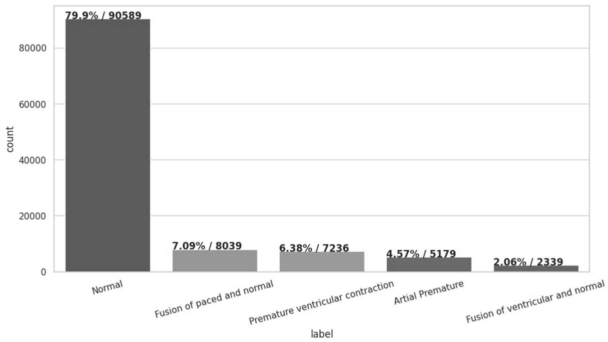

Fig. 5. Bar graph for ECG shapes of different cases.

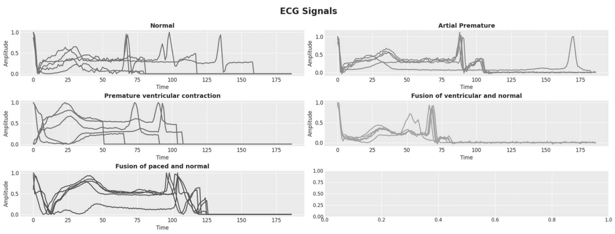

Fig. 6. ECG signals.

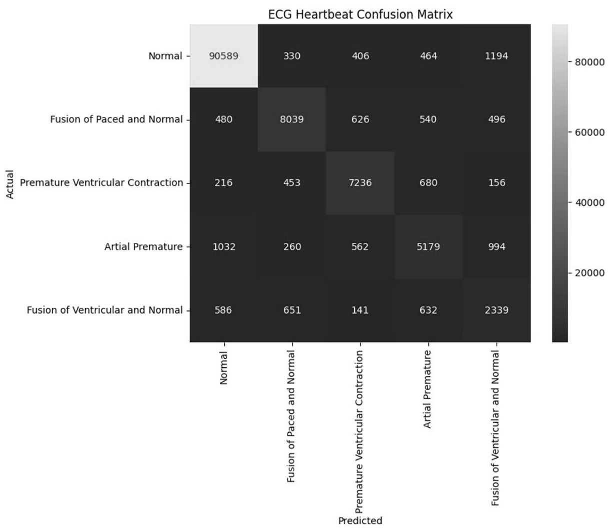

Fig. 7. Confusion matrix for ECG classification.

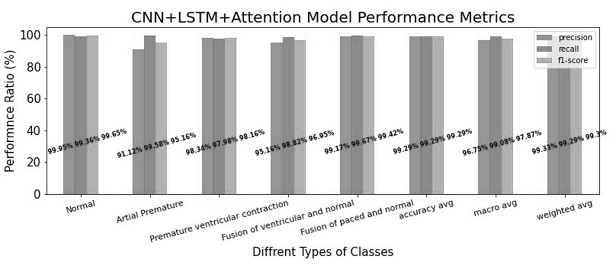

Fig. 8. Audio SA classification report for the proposed combination CNN + LSTM + attention.

effectively categorizing different heartbeat patterns with high accuracy. This underscores the utility of ECG signal analysis for diagnosing cardiac rhythm abnormalities and improving cardiovascular health monitoring.

Figure 5 defines the label and counts for the normal, fusion of paced and normal, premature ventricular contraction, atrial premature, fusion of ventricular and normal. The normal remains the highest of all the labels, while the fusion of ventricular and normal is determined to be the lowest

In Fig. 6 ECG signal analysis, distinguishing between different heartbeat categories is crucial. Here, “normal” heartbeats constitute the majority with 90589 instances, while “fusion of paced and normal” heartbeats are represented by 8039 instances. In turn “premature ventricular contraction” occurrences amount to 7236, followed by “atrial premature” beats with 2779 instances. Lastly, “fusion of ventricular and normal” beats are present in 803 instances. These categories delineate various cardiac rhythm abnormalities, emphasizing the importance of accurate classification for effective diagnosis and treatment planning.

Figure 7 shows the classification performance of ECG heartbeat types, revealing substantial imbalances across classes. Utilizing a comprehensive confusion matrix visualization, this study underscores the importance of addressing class imbalances to enhance the reliability of ECG-based heartbeat classification systems in clinical practice.

Table 3 includes various performance metrics such as true positive (TP), true negative (TN), false positive (FP), false negative (FN), and accuracy (%) for different ECG heartbeat classes.

Figure 8 reports classification results for the CNN + LSTM + attention model. The graphs comprise the percentage and estimators. The bar charts represent the normal, atrial premature, premature ventricular contraction, fusion of parcel and normal, accuracy ring, macro average, and weighted average. Figure 9 represents the ensemble classification report in the form of a matrix. It resembles the values for the precision, recall, and F1-score values. The normal, atrial premature, premature ventricular and normal, accuracy, weighted average, and macro average values mostly represent the values.

Table 4 presents the performance metrics of different audio sentiment analysis models, showcasing the superior performance of the CNN + LSTM + attention model with exceptionally high accuracy, precision, recall, and F1 score compared with other models.

4.3. Analysis and discussion of MHRFD-DLA results. A significant parameter that indicates its efficacy is the accuracy of sentiment analysis and chance element detection in the proposed multimodal sentiment analysis version, specially customized for MHRFD-DLA. The

Table 3. Performance metrics of ECG heartbeat classes.

Class	TP	TN	FP	FN	Acc (%)
Normal	90589	29144	3596	1470	95.94
Fusion of paced and normal	8039	112771	1657	2332	96.80
Premature ventricular contraction	7236	112908	2481	2174	96.27
Atrial premature	5179	114795	1920	2905	96.13
Fusion of ventricular and normal	2339	118161	1763	2536	96.55

Table 4. Comparison of audio sentiment analysis models.

Model	Accu(%)	Prec(%)	Recall(%)	F1 score(%)
CNN	96.45	96.90	96.45	96.67
LSTM	97.54	97.95	97.54	97.74
CNN+ LSTM	98.43	98.02	98.43	98.22
CNN+LSTM+ attention	99.95	99.95	99.36	99.65

Table 5. Comparative analysis of proposed MHRFD-DLA model with fine-tuned MSA models.

Model	Accu (%)	Prec (%)	Recall (%)	F1 score (%)
ASHN	70.6	70.2	82.4	75.8
SSCL	60.8	79.4	74.9	77.1
PM	72.9	70.4	79.4	74.6
SDG4	79.6	73.2	62.3	67.4
MHRFD-DLA	98.4	97.6	95.6	96.6

hybrid version, well-known for its precise blending of CNNs, LSTMs, and an attention mechanism, decodes sentiments related to maternal well-being with outstanding precision. In evaluating sentiment, the model can remarkably distinguish nuanced emotional aspects linked to cerebral health, prenatal care, and vitamins.

Figure 10 illustrates an exemplary implementation of the proposed MHRFD-DLA model. The code snippet showcases the method of generating text and audio facts, preprocessing, defining the version architecture, training, and assessment. Key steps consist of the text and ECG data, incorporating CNN+LSTM+attention mechanisms, concatenating modalities, and compiling the version. The education procedure is conducted on the education facts, accompanied through evaluation using check data to compute performance metrics such as accuracy, precision, recall, and F1-score. This implementation serves as a practical demonstration of the proposed MHRFD-DLA model in a multimodal sentiment analysis context.

Figure 11 illustrates that the model architecture is a multimodal neural network designed for sentiment analysis and health risk detection tasks. It incorporates textual and ECG data, processing them through embedding, convolutional, LSTM, and attention layers to extract relevant features. With a final dense layer for classification, the model predicts one of three classes: 'healthy', 'at risk', or 'uncertain', providing a comprehensive approach to analysing diverse data types for healthcare applications.

In Fig. 12, true positives (TP) represent correctly identified healthy cases, true negatives (TN) denote accurately recognized at-risk instances, false positives (FP) indicate misclassifications of at-risk or uncertain cases as healthy, and false negatives (FN) signify misclassifications of at-risk cases as healthy or uncertain.

The bar plot of Fig. 13 illustrates the distribution of sentiment analysis classes in the proposed multifaceted emotion analysis model. It showcases the counts of 'healthy', 'at risk', and 'uncertain' classes, providing insights into the distribution of maternal health risk levels and emotional states within the analyzed data.

In Table 5, the proposed MHRFD-DLA model

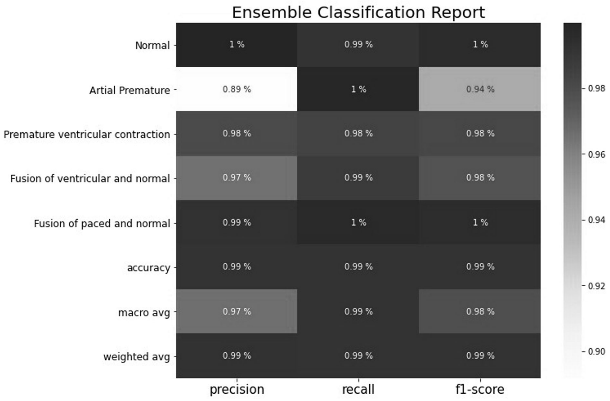

Fig. 9. Ensemble classification report for all models.

```
# Text input
text_input = Input(shape=(max_common_sequence_length,))
embedding_layer = Embedding(vocabulary_size, embedding_dim)(text_input)
cnn_lstm_attention_text = Attention()(
    [LSTM(64, return_sequences=True)(embedding_layer), LSTM(64, return_sequences=True)(embedding_layer)]
)
cnn_lstm_attention_text = Flatten()(cnn_lstm_attention_text)

# ECG input
ecg_input = Input(shape=(max_common_sequence_length, 1))
cnn_lstm_attention_ecg = Attention()(
    [LSTM(32, return_sequences=True)(Conv1D(filters=64, kernel_size=3, activation='relu')(ecg_input)),
     LSTM(32, return_sequences=True)(Conv1D(filters=64, kernel_size=3, activation='relu')(ecg_input))]
)
cnn_lstm_attention_ecg = Flatten()(cnn_lstm_attention_ecg)

# Combine modalities
combined = Concatenate()([cnn_lstm_attention_text, cnn_lstm_attention_ecg])

# Additional Dense layers for processing
dense_layer = Dense(64, activation='relu')(combined)

# Output layer
output_layer = Dense(num_classes, activation='softmax')(dense_layer)
# Create the model
model = Model(inputs=[text_input, ecg_input], outputs=output_layer)

# Compile the model
model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])
```

Fig. 10. Implementation sample of the proposed MHRFD-DLA model.

```
Model: "model"
__________________________________________________________________________________________________
 Layer (type)                Output Shape                 Param #   Connected to
==================================================================================================
 input_1 (InputLayer)        [(None, 150)]                0         []

 input_2 (InputLayer)        [(None, 150, 1)]             0         []

 embedding (Embedding)       (None, 150, 100)             10000     ['input_1[0][0]']

 conv1d (Conv1D)             (None, 148, 64)              256       ['input_2[0][0]']

 conv1d_1 (Conv1D)           (None, 148, 64)              256       ['input_2[0][0]']

 lstm (LSTM)                 (None, 150, 64)              42240     ['embedding[0][0]']

 lstm_1 (LSTM)               (None, 150, 64)              42240     ['embedding[0][0]']

 lstm_2 (LSTM)               (None, 148, 32)              12416     ['conv1d[0][0]']

 lstm_3 (LSTM)               (None, 148, 32)              12416     ['conv1d_1[0][0]']

 attention (Attention)       (None, 150, 64)              0         ['lstm[0][0]',
                                                                     'lstm_1[0][0]']

 attention_1 (Attention)     (None, 148, 32)              0         ['lstm_2[0][0]',
                                                                     'lstm_3[0][0]']

 flatten (Flatten)           (None, 9600)                 0         ['attention[0][0]']

 flatten_1 (Flatten)         (None, 4736)                 0         ['attention_1[0][0]']

 concatenate (Concatenate)   (None, 14336)                0         ['flatten[0][0]',
                                                                     'flatten_1[0][0]']

 dense (Dense)               (None, 64)                   917568    ['concatenate[0][0]']

 dense_1 (Dense)             (None, 3)                    195       ['dense[0][0]']

==================================================================================================
Total params: 1037587 (3.96 MB)
Trainable params: 1037587 (3.96 MB)
Non-trainable params: 0 (0.00 Byte)
```

Fig. 11. Multimodal health risk analyzer model.

outperforms other fine-tuned MSA models throughout all metrics, with notably higher accuracy, precision, recall, and F1-score Its superior performance underscores its effectiveness in multimodal sentiment analysis tasks.

Figure 14 displays performance assessment of proposed multimodal sentiment analysis models with various excellent fine-tuned models. Each plot affords metrics which include accuracy, precision, and F1-rating. Notably, the proposed MHRFD-DLA version demonstrates advanced performance throughout all evaluated metrics, showcasing its effectiveness in sentiment analysis responsibilities.

Figure 15 presents a performance evaluation of various fine-tuned multimodal sentiment analysis models using bar plots for accuracy, precision, recall, and F1-score. The proposed MHRFD-DLA model excels across all metrics, demonstrating its robustness for sentiment analysis applications. The model includes 1,037,587 trainable parameters and 8 hidden layers, comprising CNN, LSTM, dense, and 2 attention layers. These attention layers enhance the model's ability to focus on relevant features and patterns. With a learning rate of 0.001, the model was trained for 100 epochs, showing minimal improvement after around 73 epochs.

Figure 16 illustrates the computational time comparison among various maternal health risk prediction models, including ASHN, SSCL, PM, SDG4, and MHRFD-DLA. Each model is represented by a distinct color, showcasing the computational efficiency of MHRFD-DLA compared with others. MHRFD-DLA demonstrates significantly lower computational time, denoted by the shorter vertical line, indicating its potential for efficient real-time deployment in maternal healthcare settings. Specifically, ASHN takes 1778.57 seconds, SSCL takes 2546.72 seconds, PM takes 1567.16 seconds, SDG4 takes 2298.32 seconds, and MHRFD-DLA takes only 959.24 seconds.

5. Future research directions

Below prospective directions of extending the presented results are listed:

(i) Future research should advance multimodal fusion techniques for better integration of text, images, videos, and user metrics.

(ii) Methods are to be developed for fine-grained emotion analysis to capture subtle emotional nuances and changes.

(iii) More efforts are needed towards cross-domain generalization to enhance model performance across various social media platforms and improve interpretability.

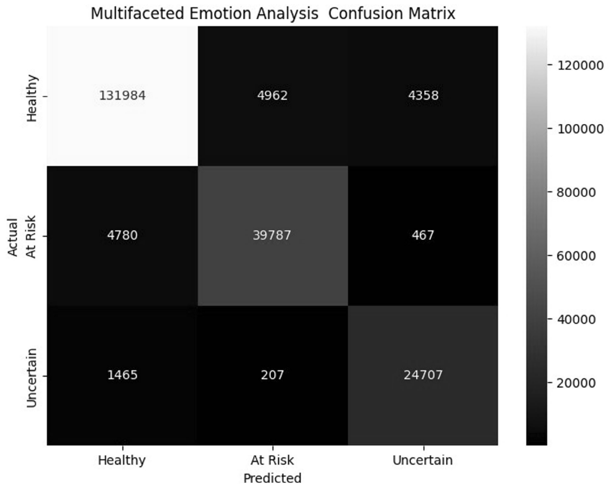

Fig. 12. Confusion matrix for multifaceted emotion analysis.

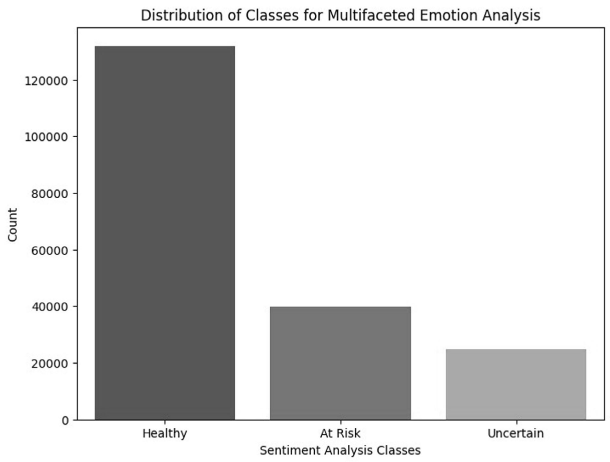

Fig. 13. Distribution of classes for multifaceted emotion analysis.

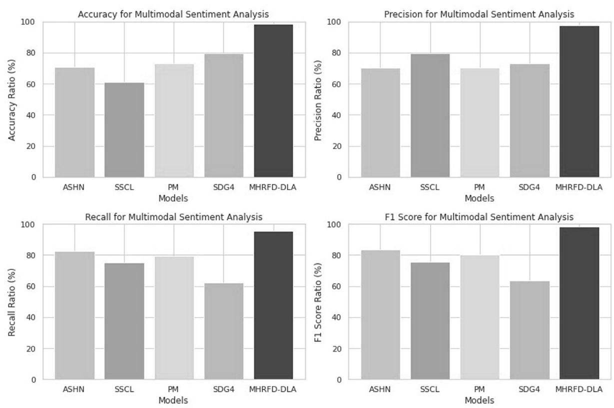

Fig. 14. Individual performance evaluation of the proposed MHRFD-DLA models with various fine-tuned multi-modal approaches.

(iv) Focus on real-time monitoring and longitudinal analysis for tracking health trends is expected.

(v) Ethical considerations and biases to improve public health interventions and maternal and child health outcomes globally should be addressed.

6. Conclusion and significance

The proposed maternal health risk factor detection constitutes a significant advancement in social network multimodal sentiment analysis for maternal health. By effectively handling the complexities of dynamic user-generated content and multiple data modalities, MHRFD-DLA offers a superior understanding of maternal health sentiments compared with traditional methods. This approach enhances risk factor detection and enables timely interventions by analyzing sentiments related to prenatal care, nutrition, and mental well-being. MHRFD-DLA is a valuable tool for shaping public health policies and improving maternal and child health outcomes globally. Simulation results demonstrate its effectiveness, with 98.4% accuracy, 97.6% precision, 95.6% recall, and an F1-score of 96.6%.

Acknowledgment

We sincerely thank Anna University in Chennai for their support and encouragement throughout this research. Special thanks go to our colleagues and mentors for their invaluable insights and encouragement throughout this research.

References

Acharya, A., Ramesh, R., Fathima, T. and Lakhani, T. (2023). Clinical tools to detect postpartum depression based on machine learning and EEG: A review, *in 2nd International Conference on Computational Systems and Communication (ICCSC), Thiruvananthapuram, India,* pp. 1–8.

Ahmed, M., Kashem, M.A., Rahman, M. and Khatun, S. (2020). Review and analysis of risk factor of maternal health in remote area using the internet of things (IoT), *in* A.N. Kasruddin Nasir *et al.* (Eds), *InECCE2019*, Lecture Notes in Electrical Engineering, Vol. 632, Springer, Singapore, pp. 357–365.

Ahmed, M. and Kashem, M.A. (2020). IoT based risk level prediction model for maternal health care in the context of Bangladesh, *Proceedings of the 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh*, pp. 1–6, DOI: 10.1109/STI50764.2020.9350320.

Afyouni, I., Al Aghbari, Z. and Razack, R.A. (2022). Multi-feature, multimodal, and multi-source social event detection: A comprehensive survey, *Information Fusion* **79**: 279–308, DOI: 10.1016/j.inffus.2021.10.013.

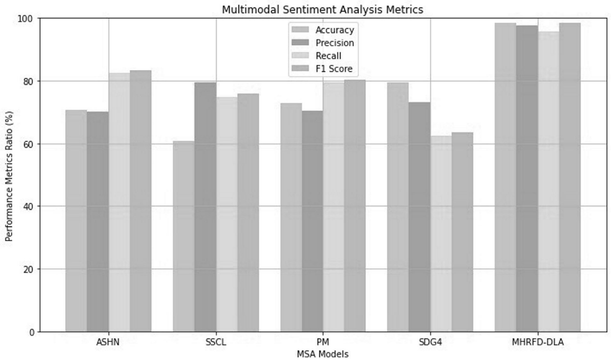

Fig. 15. Comprehensive performance evaluation of the proposed MHRFD-DLA models with various fine-tuned multimodal approaches.

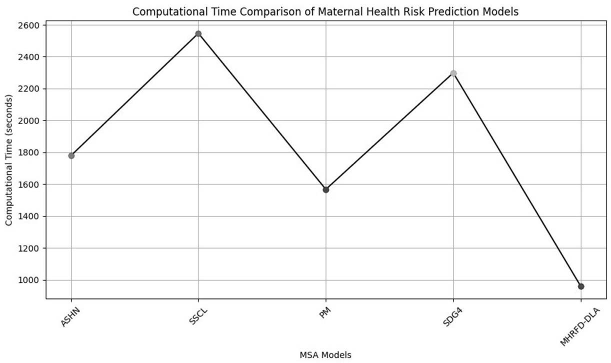

Fig. 16. Computational time of maternal health risk prediction models.

Argyriou, A., Evgeniou, T. and Pontil, M. (2008). Convex multi-task feature learning, *Machine Learning* **73**(3): 243–272.

Byeon, H. (2023). Advances in machine learning and explainable artificial intelligence for depression prediction, *International Journal of Advanced Computer Science and Applications* **14**(6): 896, DOI: 10.14569/IJACSA.2023.0140656.

Cimtay, Y., Ekmekcioglu, E. and Caglar-Ozhan, S. (2020). Cross-subject multimodal emotion recognition based on hybrid fusion, *IEEE Access* **8**: 168865–168878, DOI: 10.14569/IJACSA.2023.0140656.

Fazeli, S. (2024). Heartbeat data, `https://www.kaggle.com/datasets/shayanfazeli/heartbeat`.

García-Díaz, J.A., Cánovas-García, M. and Valencia-García, R. (2020). Ontology-driven aspect-based sentiment analysis classification: An infodemiological case study regarding infectious diseases in Latin America, *Future Generation Computer Systems* **112**: 641–657, DOI: 10.1016/j.future.2020.06.019.

Geethanjali, R. and Valarmathi, A. (2022). Issues and future challenges of sentiment analysis for social networks—A survey, *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India,* pp. 332–339.

Geethanjali, R. (2024). Maternal health risk data, `https://www.kaggle.com/datasets/geethanjaliyokesh/maternal-health-risk-data-1`.

Ghosh, T., Banna, M.H.A., Nahian, M.J.A., Uddin, M.N., Kaiser, M.S. and Mahmud, M. (2023). An attention-based hybrid architecture with explainability for depressive social media text detection in Bangla, *Expert Systems with Applications* **213**: 119007, DOI: 10.1016/j.eswa.2022.119007.

Gopalakrishnan, R. (2023). Attribute selection hybrid network model for risk factors analysis of postpartum depression using social media, *Brain Informatics* **10**(1): 28, DOI: 10.1186/s40708-023-00206-7.

Gupta, G.K. and Sharma, D.K. (2023). Depression detection using semantic representation-based semi-supervised deep learning, *International Journal of Data Analysis Techniques and Strategies* **15**(3): 217–237.

Gupta, V. (2021). An emotion care model using multimodal textual analysis on COVID-19, *Chaos Solitons Fractals* **144**: 110708, DOI: 10.1016/j.chaos.2021.110708.

Kachuee, M., Fazeli, S. and Sarrafzadeh, M. (2018). ECG heartbeat classification: A deep transferable representation, *arXiv:* 1805.00794.

Lilhore, U.K., Dalal, S., Varshney, N., Sharma, Y.K., Rao, K.B.V.B., Rao, V.V.R.M., Alroobaea R., Simaiya, S., Margala, M. and Chakrabarti, P. (2024). Prevalence and risk factors analysis of postpartum depression at early stage using hybrid deep learning model, *Scientific Reports* **14**: 4533, DOI: 10.1038/s41598-024-54927-8.

Nadeem, A., Naveed, M., Islam Satti, M., Afzal, H., Ahmad, T. and Kim, K.-I. (2022). Depression detection based on hybrid deep learning SSCL framework using self-attention mechanism: An application to social networking data, *Sensors* **22**(24): 9775, DOI: 10.3390/s22249775.

Nazir, A., Rao, Y., Wu, L. and Sun, L. (2022). Issues and challenges of aspect-based sentiment analysis: A comprehensive survey, *IEEE Transactions on Affective Computing* **13**(2): 845–863.

Nti, I.K. and Owusu-Boadu, B. (2022). A hybrid boosting ensemble model for predicting maternal mortality and sustaining reproductive health, *Smart Health* **26**: 100325, DOI: 10.1016/j.smhl.2022.100325.

Oueslati, O., Cambria, E., HajHmida, M.B. and Ounelli, H. (2020). A review of sentiment analysis research in Arabic language, *arXiv:* 2009.01360.

Pooja and Bhalla, R. (2022). A review paper on the role of sentiment analysis in quality education, *SN Computer Science* **3**(6): 469.

Rahman, A. and Alam, M.G.R. (2023). Explainable AI based maternal health risk prediction using machine learning and deep learning, *2023 IEEE World AI IoT Congress (AIIoT), Seattle, USA*, pp. 0013–0018, DOI: 10.1109/AIIoT58121.2023.10174540.

Shoumy, N.J., Ang, L.-M., Seng, K.P., Rahaman, D.M.M. and Zia, T. (2020). Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals, *Journal of Network and Computer Applications* **149**: 102447, DOI: 10.1016/j.jnca.2019.102447.

Teisseyre, P. (2022). Joint feature selection and classification for positive unlabelled multi-label data using weighted penalized empirical risk minimization, *International Journal of Applied Mathematics and Computer Science* **32**(2): 311–322, DOI: 10.34768/amcs-2022-0023.

Titla-Tlatelpa, J. de J., Ortega-Mendoza, R.M., Montes-y-Gómez, M. and Villaseñor-Pineda, L. (2021). A profile-based sentiment-aware approach for depression detection in social media, *EPJ Data Science* **10**, Article no. 54, DOI: 10.1140/epjds/s13688-021-00309-3.

Togunwa, T.O., Babatunde, A.O. and Abdullah, K.U. (2023). Deep hybrid model for maternal health risk classification in pregnancy: Synergy of ANN and random forest, *Frontiers in Artificial Intelligence* **6**: 1213436, DOI: 10.3389/frai.2023.1213436.

Wang, S., Chen, J. and Lu, Z. (2021). Multimodal sentiment analysis using transformer-based approaches, *IEEE Transactions on Affective Computing* **12**(1): 82–93.

Xue, Y., Wang, Y. and Yu, M. (2020). Cross-modal sentiment analysis with sentiment-enhanced multimodal fusion, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 5677–5687, (online).

Zhang, J., Yin, Z., Chen, P. and Nichele, S. (2020). Emotion recognition using multimodal data and machine learning techniques: A tutorial and review, *Information Fusion* **59**: 103–126, DOI: 10.1016/j.inffus.2020.01.011.

R. Geethanjali is currently pursuing her PhD on deep learning based multi-lingual and multi-modal sentiment analysis model for social networks with applications in public health. Her scientific interest include machine learning and data science.

A. Valarmathi is an assistant professor at Anna University in Chennai, specializing in geoinformatics, machine learning, and wireless networks. With a distinguished career spanning over a decade, she has held various administrative positions and received numerous awards for her contributions to education and research. Her scientific interest include spatial mining and soil contamination.

Received: 8 February 2024
Revised: 28 March 2024
Re-revised: 25 June 2024
Accepted: 23 August 2024