International Journal of Advanced Trends in Engineering and Management
(ISSN 2583-7052)

Vol. V, No. 09, pp. 52-60, September

# Optimized Mamba-KAN with Residual Attention for Social Media Stress Detection

P. Kavitha[3], A. Anish Rama[4]

[3]Research Scholar, Department of Electrical and Electronics Engineering,
Government College of Engineering, Tirunelveli, Tamilnadu, India. kavitha.paulsamy6@gmail.com

[4]Assistant Professor, Department of Electronics and Communication Engineering
Arunachala College of Engineering for Women, Tamil Nadu, India. ramaanish86@gmail.com

**ABSTRACT:** Stress is a prevalent mental health condition that reveals in various situations and varies its degrees of severity. Conventional frameworks for identifying psychological stress are mainly heuristic, primarily binary classed, and fail to recognize the complexity and multidimensionality of stress. Hence, in this paper, a novel Mountain Gazelle Optimizer (MGO)-enhanced Mamba-Kolmogorov-Arnold Network (KAN) framework which is integrated with a Residual Attention mechanism is presented for precise social media classification. Initially, to increase data quality, social media textual data is first pre-processed using techniques including data visualization, text normalization, and noise reduction. Following that a Mamba layout that use selective state-space modelling is utilized to identify sequential and contextual connections from textual representations. Residual Attention mechanism is incorporated to preserve important contextual information. Furthermore, to classify stress and capture intricate correlations between retrieved variables, KAN architecture makes use of learnable nonlinear functions. In order to enhance classification accuracy, convergence properties, and generalization performance, specific model hyperparameters are optimized using MGO. Effectiveness of proposed framework is assessed through python software with classification accuracy, precision, recall, F1-score, Receiver Operating Characteristics-Area Under Curve (ROC-AUC), Matthews Corelation Coefficient (MCC), specificity and sensitivity. Values of performance metrics are 0.9853, 0.9854, 0.9853, 0.9853, 0.9984, 0.9647, 0.9962 and 0.9594 respectively. Proposed integration aims to enhance contextual feature representation, lower classification mistakes, and facilitate automated stress identification.

**Keywords:** Stress classification, Mountain Gazelle Optimizer, Mamba architecture, Kolmogorov-Arnold Network, Residual Attention Mechanism.

## 1. Introduction

Social networking becomes a vital tool for people who are looking for mental health assistance. Researchers demonstrated that utilizing social media data analysis is capable of showing signs of anxiety, long-term stress, and emotional strain [1, 2]. World Health Organization (WHO) states that psychological stress is a major contributor to mental health issues that impair people's general quality of life and productivity in work environments. Stress detection based on text provides non-intrusive alternative to more conventional approaches like surveys, bio-signals monitoring, or clinical interviews [3, 4]. Social networking sites like Reddit and Twitter are used by people who are depressed and under stress to share their ideas, feelings, and other things with other users through posts and comments. As a result, social media is a useful source of information where linguistic patterns of posts that

are stressful or depressing are easily examined [5]. Deep learning (DL) techniques have emerged recently because they outperform conventional Machine Learning (ML) algorithms and frequently eliminate time-consuming feature extraction process. In [6], Hierarchical convolutional attention network is presented for depression detection in which Convolutional Neural Network (CNN) aids in detecting stress. In [7], hierarchical attention model is presented for depression detection but its classification model consumes more time in convergence with poor interpretability. Hence to achieve higher classification accuracy with less iteration, Mamba-KAN framework with Residual attention network is proposed in this work to enhance robustness and model convergence. Optimizing attention networks improves model convergence thereby improving classification accuracy. In [8], Grey Wolf Optimization (GWO) is utilized to optimized automatic stress detecting model that improves success rate by including objective fitness function. Particle Swarm Optimization (PSO) algorithm in [9], optimizes CNN model to enhance classification accuracy and model interpretability in social media stress classification but appropriate weight parameters is required for feature extraction and classification. Hence to address aforementioned challenges, MGO algorithm is proposed in this work to improve classification accuracy thereby reducing computational complexity. Main contribution of proposed work includes,

- Developing effective stress classification framework by integrating Mamba-KAN framework that capture complex non-linear relationship and behavioural pattern.
- To enhance feature extraction and representation using Residual Attention network while reducing influence of less informative features.
- To optimize model performance using MGO thereby improving classification accuracy and model convergence.

Next section details about work related to proposed system with its advantages and limitations.

## 2. Recent Works

**Kalyani Pendke et al [2026] [10]** have proposed a novel hybrid DL technique integrating Bidirectional Encoder Representations from Transformer (BERT), Robustly optimized BERT Pretraining Approach (RoBERTa), CNN, and BiLSTM models for stress classification. Three emotional states such as stressed, not stressed and highly stressed are detected using this approach. Despite of optimum F1-score and accuracy rate of 97.84 percent is obtained, misclassification and overlapping of datasets occurs that influence performance of system.

**Nazzere Oryngozha et al [2024] [11]** have proposed automatic stress classification approach based on online education and remote work. Using Dreaddit as training dataset, overall stress level is detected as 29% and accuracy is 72%. Limitations occur due to poor classification accuracy, absence of hyperparameters optimization and lack of DL approach integrations.

**Sathiya et al [2025][12]** have proposed Quantum Neural Network (QNN) based novel framework for stress classification in mood-disordered patients. Using QNN, accuracy and precision are obtained as 94.5% and 94% respectively in identifying stress percentage. Uncertainty occurs in personality traits estimation, high computational complexity and lack of real-time deployment evaluations.

**Luna Ansari et al [2023] [13]** have proposed a novel method improving stress recognition by comparing two methods such as hybrid and ensemble. Accuracy and F1-score are obtained as 75% and 0.77 respectively, in detecting depression using Reddit social media dataset. Limitation of this work are impacts on mental health indicators and assessment due to misclassification, poor data privacy, interpretability and uncertainty issues while gathering data and model training.

## 3. Proposed Work Explanation

In this proposed work, effective stress classification from social media textual dataset is presented. Schematic block diagram for social media stress classification is shown in Figure 1. Initially, social media stress dataset is subjected to data processing approach to reduce noise thereby enhancing quality of input information. Dataset properties, class distribution, and relevant textual patterns linked to stress-related material are then analysed through data visualization.

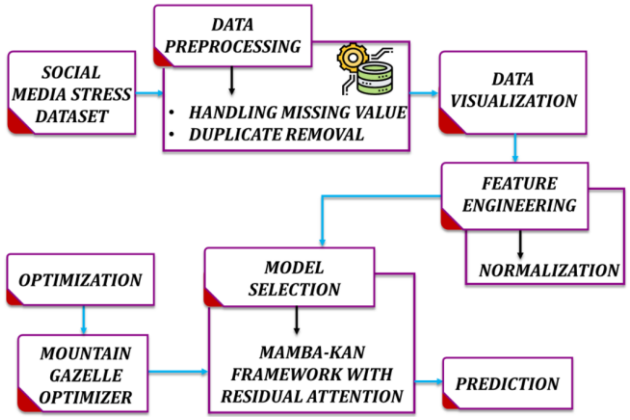



**Figure 1:** *Block diagram for social media stress classification*

Processed data are then subjected to feature engineering stage, where textual representations are prepared for classification. Feature normalization is used to facilitate steady learning and preserve consistency in extracted representations. Model selection stage is utilized to classify stress using Mamba-KAN framework with Residual Attention Network in which Mamba architecture uses selective state-space modelling to capture contextual dependencies and sequential relationships in social media textual information. Tuning hyperparameters of selection model, MGO is utilized thereby enhancing classification accuracy. Finally optimized framework provides predicted output, classifying social media textual datasets into stress and non-stress categories.

### 3.1 Data processing

Initially, social media dataset is subjected to pre-processing technique. In this process, missing values are handled precisely by removing unwanted noise and removing duplicate information form input dataset. Preprocessing is followed by data visualization where related textual patterns linked with stress detection is analysed. In data normalization, steady state learning is performed that prepares the model for classification purpose.

### 3.2 Model Selection

After normalization process, dataset under goes model selection which is performed using Mamba-KAN framework with Residual Attention Network. Mamba block creates a selective state space model by combining linear and convolutional layers with element-wise multiplication and SiLU activation function. KAN branch uses learnable activation functions across nodes in place of traditional fixed activation functions. Mamba block is used as temporal backbone, which achieves discretization through Zero-Order Hold (ZOH) which is expressed as,

$$h_t = Ah_{t-1} + Bx_t \quad (1)$$

$$y_t = Ch_t \quad (2)$$

Where input and output are denoted as $x_t$ and $y_t$, hidden state is denoted as $h_t$. Mamba's primary uniqueness is its selective scan approach, which dynamically generates parameters $B$ and $C$ based on current input $x_t$:

$$B = s_B(x_t) \quad (3)$$

$$C = s_c(x_t) \quad (4)$$

$$\Delta = s_\Delta(x_t) \quad (5)$$

Mamba-KAN Network model adaptively determines whether to retain or forget historical data based on social media stress classification because of its intrinsic input dependency. KAN functions as a nonlinear semantic improvement branch to improve model's capacity to capture complex nonlinear process couplings. Output is calculated as given below,

$$KAN(x) = \sum_{j=1}^{n} \phi_{q,j}(x_p) \quad (6)$$

Where $x_p$ indicates $p$-th scalar elements inside $x$ and $x = [x_1, x_2, \dots, x_d]$ denotes multidimensional input features vector. Further enhancing strength and flexibility of stress detecting system, a residual attention module is added to proposed selection model. Further, selected model is optimized to enhance its classification accuracy and following section provides details on optimization algorithm utilized in this proposed work.

### 3.3 Fine tuning using MGO

To optimize hyperparameters of Mamba-KAN framework with residual attention network, MGO is employed, in which complex behaviour of natural mountain gazelle herds serves as an inspiration for MGO, where different responsibilities contribute to herd's survival and success. Figure 2 represents flowchart of MGO for optimization process of Mamba-KAN framework with residual attention network.

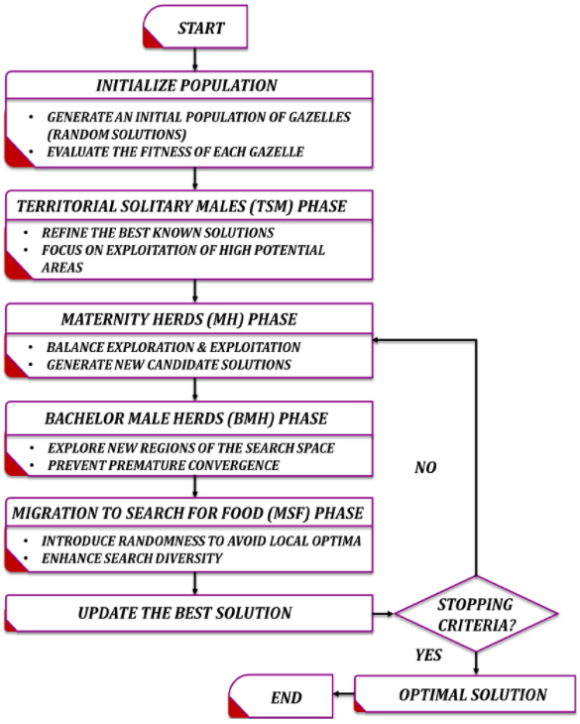



**Figure 2:** *Flowchart for MGO*

Territorial Solitary Males (TSM) is dominating individuals in charge and protecting particular areas. Refinement process is supported by Maternity Herds (MH) that consists of female gazelles and its smaller ones. Bachelor Male Herds (BHM) consists of younger male which is responsible for wide-ranging exploration. Migration to Search for Food (MSF) phase reflects gazelle tendency to travel long distance in quest of food.

#### *3.3.1 TSM phase*

The optimal region is represented by TSM mechanism, which is best-performing solution and it is expressed as,

$$TSM = male_{gazelle} - \left| \left(r_{i1} * BH - r_{i2} * X(t)\right) * F \right| * Cofr \quad (7)$$

Here, present best solution position vector is denoted as $male_{gazelle}$, Coefficient vector that represent bachelor male herd is denoted as $BH$ and it is calculated using;

$$BH = X_{ra} * \lfloor r_1 \rfloor + Mpr * \lceil r_2 \rceil \quad (8)$$

Here integers that are randomly chosen is denoted as $r_{i1}$ $and$ $r_{i2}$, adaptive adjustment factors is denoted as F and expressed as,

$$F = N_1(D) * \exp\left(2 - Iter * \frac{2}{MaxIter}\right) \quad (9)$$

$$a = -1 + Iter \times \left(\frac{-1}{MaxIter}\right) \quad (10)$$

Where $Iter$ denotes present iteration and summation of iteration is denoted as $MaxIter$. Random coefficient vector is expressed as,

$$Cofr = (a + 1) + r_3 \quad (11)$$

Where $r_3$ is chosen randomly from range between [0, 1]. Each coefficient plays major role in guiding optimization process.

#### *3.3.2 MH phase*

Focusing on providing up with fresh, viable solutions while utilizing knowledge of best ones that already exist and it is expressed as,

$$MH = (BH + Cof\ 1) + \left(r_{i3} * male_{gazelle} - r_{i4} * X_{rand}\right) * Cof\ 1 \quad (12)$$

Where random integers are $r_{i3}$ and $r_{i4}$, Random coefficient vector is denoted as $Cof1$ and gazelles' position that is randomly selected is denoted as $X_{rand}$. By preserving diversity and improving present solutions, MH phase supports both exploration and exploitation.

*3.3.3 BMH phase*

Through promoting exploration, bachelor herds validate that the algorithm is validated in variety of areas.

$$BMH = (X(t) - D) + \left(r_{i5} * male_{gazelle} - r_{i6} * BH\right) * Cofr \quad (13)$$

Here displacement factor is denoted as $D$, present position of gazelle is denoted as $X(T)$.

$$D = (\left|X(t)\right| + \left|male_{gazelle}\right|) * (2 * r_6 - 1) \quad (14)$$

Where random integers are $r_{i5}$ and $r_{i6}$. By increasing search diversity, this stage lessens possibility that the algorithm may become constrained in local optima.

*3.3.4 MSF phase*

Migration phase allows algorithm to become free from local optima by ensuring random but strategic exploration of search space and it is expressed as,

$$MSF = (ub - lb) * r_7 + lb \quad (15)$$

Random number is denoted as $r_7$ which is chosen from 0 and 1, lower and upper boundaries search space is denoted as $lb$ and $ub$. While TSM and MH phase refine ideal regions to improve convergence accuracy during exploitation phase, BMH and MSF mechanisms allow glance over of large areas in search space during exploration phase. Following section provides details about mathematical equations utilized in this proposed work.

### 3.4 Mathematical Expressions and Symbols

This section details with numerical expression and symbols related to proposed work. Accuracy is evaluated by considering ratio of accurately defined parameters to sum of cases.

$$Accuracy = \frac{TP+TN}{\text{TP+TN+FP+FN}} \quad (16)$$

Precision is ratio of accurate positives among positive cases predicted to false positive cases.

$$Precision = \frac{TP}{\text{TP+FP}} \quad (17)$$

Recall is ratio measure of actual positive cases to cases detected with pneumonia.

$$Recall = \frac{TP}{\text{TP+FN}} \quad (18)$$

Specificity is ratio of negative cases that is detected perfectly by avoiding false alarms.

$$Specificity = \frac{TN}{\text{TN+FP}} \quad (19)$$

F1-score generates harmonic average of recall and precision.

$$F1 - score = \frac{2 \times Precision \times Recall}{\text{Precision+Recall}} \quad (20)$$

The next section details with results and discussion for proposed work.

## 4. Results and Discussion

In this section, results for proposed social media stress detection are explained in detail using dataset provided as social media usage, its sleep and stress impact. Results are validated using python software and dataset is accessed using link https://www.kaggle.com/datasets/onlinemad/social-media-usage-its-sleep-and-stress-impact?utm_source=chatgpt.com&select=Smartphone_Usage_And_Addiction_Analysis.xlsx. Structure of dataset are 7500 rows and 12 columns with 0 missing values. Ratio values under train and test samples are 80% and 20% respectively with train and test counts 6000 and 1500 respectively. High and medium class counts are 5284 and 2216 respectively. Parameters specified for proposed work is tabulated in Table 1.

**Table 1:** *Parameters Specification*

| Parameters | Values |
|---|---|
| Batch Size | 64 |
| Learning Rate | 0.001 |

| | |
|---|---|
| Epochs | 50 |
| Number of layers | 2 |
| Hidden units | 64 |
| Attention heads | 4 |
| Dropout rate | 0.3 |
| Momentum | 0.9 |
| Sequence Length | 1 |
| Hidden Dimension | 64 |

The features of dataset, such as transaction ID, user ID, age, age group, gender, daily screen time, social media usage, gaming hours, screen time category, sleep hours, sleep health, and stress level, were correlated as shown in Figure 3(a). Diagonal elements show correlation coefficient of 1.0 that indicates a positive correlation of each feature. Gender distribution among male, female and other are illustrated in Figure 3(b). The results shows that 34% male participants, 32.8% female participants and 33.1% is classified under other categories.

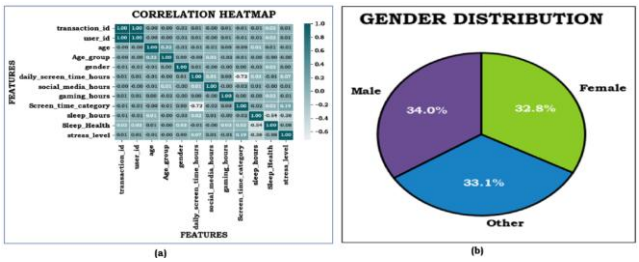



**Figure 3:** *(a) Correlation Heatmap (b) Gender distribution among male, Female and other*

The behavioural and population-based distributions offer important information about features of the dataset used to classify stress. Participants were divided into five age groups: children, teens, young adults, adults, and elders as depicted in Figure 4(a). Bar chart represents adult, seniors, young adult, children and teenager categories with 2994, 2073, 1009, 925 and 499 participants respectively. Frequency distribution of participants across different ages are illustrated in Figure 4(b), in which observed counts varies between 70 and 120. These results provide strength to proposed stress classification framework's consideration of age as a possible demographic characteristic. Figure 4(c) illustrates distribution of screen time used daily among participants covering almost 3 to 12 hours per day. Figure 4(d) illustrates variations in social media usage hours across ordered observations. Social media usage rises from about 0.5 hours to 6 hours as observation index rises, according to graph's almost linear increasing pattern.

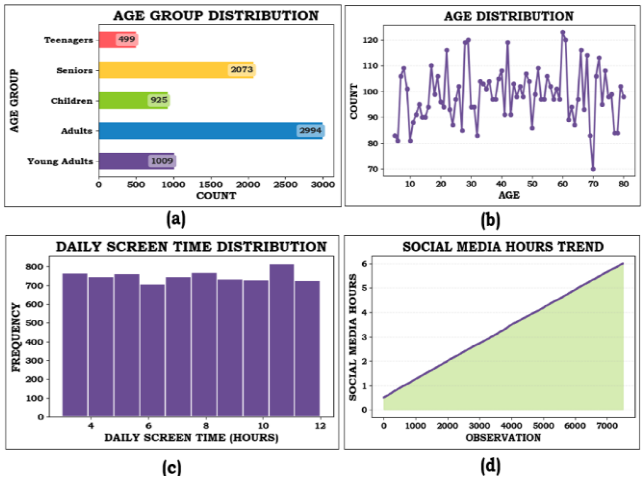



**Figure 4:** *Outcomes based on different distributions (a) Age group (b) Age (c) Daily access time (d) Social media hours trends*

Gaming hours, screen time categories, sleep length, and sleep health distributions offer crucial information about dataset's behavioural traits. Figure 5(a) provides details about density distribution of daily gaming hours among participants. The findings show that range of game time is from 0 to 4.5 hours, with highest density recorded between 0.5 and 3.5 hours. Figure 5(b) illustrates frequency distribution of participants across high, moderate, low and extreme. High and moderate screen time categories have higher number of observations with 3000 participants each. Low screen participants are with 900 observations whereas extreme categories with 200 observations in stress classification. Figure 5(c) represents daily sleep duration of participants that covers approximately 5 to 9 hours. Figure 5(d) represents distribution of participants of sleep health categories namely, healthy sleep, overlapping and poor sleep. Results highlights that healthy sleep contains highest number of observations of 3409, 2490 observations for poor sleep category, and overlapping sleep category contains 1600 observation.

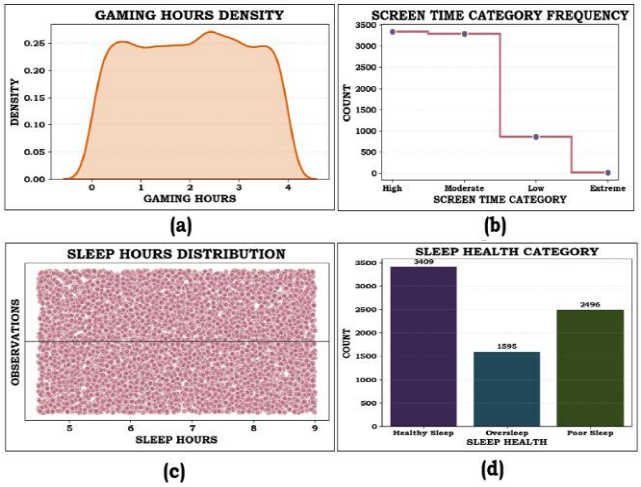



**Figure 5:** *(a) Density of gaming hours (b) Screen time category frequency (c) Sleep hours distribution (d) Sleep health category*

The outcomes of simulation show learning properties and initial classification efficacy of proposed Mountain Gazelle Optimizer-Enhanced Mamba-KAN Framework with Residual Attention for Social Media Stress Classification. Figure 6(a) represents distribution of stress levels among participants with high and medium stress levels. The findings show that 2,216 participants (29.55%) belong to medium-stress category and 5,284 participants (70.45%) belongs to high-stress category. Figure 6(b) denotes division of dataset into training and testing subsets for evaluating proposed stress classification framework. Among 7,500 observations in entire dataset, 1,500 samples (20%) are set aside for testing and 6,000 samples (80%) are used for model training. Figure 6(c) defines model accuracy based on training and validation curve for proposed model over 30 epochs. Training curve increases from 78% to 90% within initial epochs and both training and validation curves stabilizes at 93% after 25 epochs. Figure 6(d) represents variation in training and validation loss across 30 epochs. In the first few epochs, training loss drops dramatically to almost 0.20 from its starting point of almost 0.48. Throughout the training phase, a further progressive decrease is observed, with the loss approaching 0.05–0.07 in the last epochs.

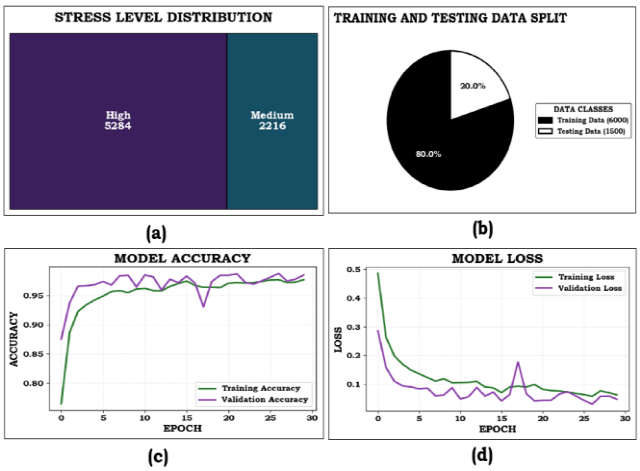



**Figure 6:** *(a) Distribution among stress level (b) Training and testing dataset split (c) Model accuracy (d) Model loss*

Predictive performance of proposed MGO-enhanced Mamba-KAN framework with Residual Attention for social media stress classification is represented in Figure 7. Performance metrics such as accuracy, recall precision, specificity, sensitivity, MCC, F1-score and ROC-AUC is depicted in Figure 7(a) with values of 0.9853, 0.9853, 0.9854, 0.9962, 0.9594, 0.9547, 0.9853 and 0.9984 respectively. Figure 7(b) defines confusion matrix of proposed framework using 1500 testing samples. The findings indicate that 425 medium-stress samples and 1,053 high-stress samples were appropriately identified. While 18 medium-stress samples are mistakenly classified as high stress and 4 high-stress samples are mis predicted as medium stress. Figure 7(c) illustrates ROC-AUC curve that indicates relationship between positive and negative rate across different classification threshold. ROC curve rises steeply toward upper-left corner and stays near maximum true positive rate and AUC is nearly 0.9984.

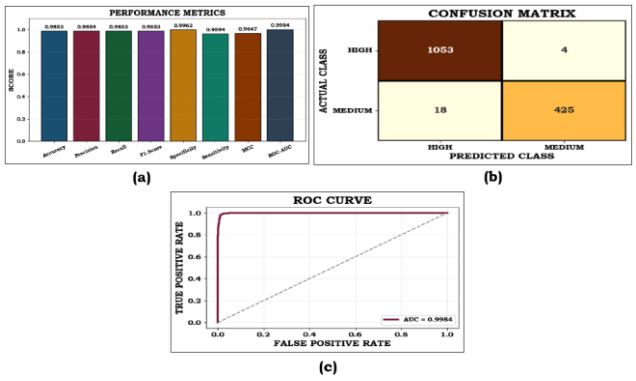



**Figure 7:** *(a) Performance Metrics (b) Confusion Matrix (c) ROC curve*

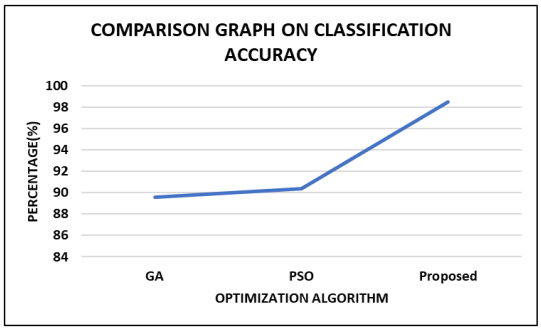



**Figure 8:** *Comparison graph for different classifiers*

Comparison graph for different optimization algorithm such as Genetic Algorithm (GA) [14], PSO [15], and proposed is depicted in Figure 8, in which classification accuracy is compared and it details that proposed algorithm performs with higher classification accuracy of 98.5%.

## 5. Conclusion

In this paper, social media stress among different participants such as teenagers, adults, young adults and children is determined. Framework incorporates Mamba-KAN with Residual Attention Network and hyperparameters are finely tuned using MGO. Proposed system is validated using Python software and outcomes are obtained based classification accuracy, recall, precision, specificity, sensitivity, MCC, F1-score and ROC-AUC values of 0.9853, 0.9853, 0.9854, 0.9962, 0.9594, 0.9547, 0.9853 and 0.9984 respectively. Future research focus on detecting multiclass stress using explainable Artificial Intelligence (XAI), real-time stress monitoring and computational optimization to improve robustness, interpretability and practical application.

## References

1. Rohan Sharma Sitoula; Moumita Pramanik; Ranjit Panigrahi, Year: 2024, “Fine-grained classification for emotion detection using advanced neural models and GoEmotions dataset”, Journal of Soft Computing and Data Mining, Vol. 5, No: 2, pp. 62–71.
2. Jyoti Kandpal; A. Jonitha; Janaki Sivakumar; Rejon Kumar Ray, Year: 2023, “Multi-Modal Data Fusion Approach Using the Structured and Sparse Principal Component Analysis (MMSSCA)”, In 2023 IEEE International Conference on Paradigm Shift in Information Technologies with Innovative Applications in Global Scenario (ICPSITIAGS), pp. 386-391.
3. Janaki Sivakumar; Khoula Al Harthy, Year: 2024, “AI-Powered Teaching Plan: Present and Post Evaluation”, In Utilizing AI for Assessment, Grading, and Feedback in Higher Education, pp. 162-188.
4. V. Thamilarasi, Year: 2026, “Neuro-Symbolic AI with Edge Computing and Reinforcement Learning Optimizing Autonomous Engineering Design Workflows”.
5. Srinivasarao Gadu; Gadi Himaja; Yenumula Lakshmi Vinanya; Kosuri Durga Sravya; Kosuri SS Durga Aryan; Tasleem; Chenjeri Aditya Sai, Year: 2025, “Detection of early-stage of lung cancer using machine learning classifiers”, In AIP Conference Proceedings, AIP Publishing LLC, Vol: 3306, No: 1, pp. 040001.
6. Hamad Zogan; Imran Razzak; Shoaib Jameel; Guandong Xu, Year: 2023, “Hierarchical convolutional attention network for depression detection on social media and its impact during

pandemic", IEEE Journal of Biomedical and Health Informatics, Vol: 28, No: 4, pp. 1815-1823.

7. Zepeng Li; Zhengyi An; Wenchuan Cheng; Jiawei Zhou; Fang Zheng; Bin Hu, Year: 2023, "MHA: A multimodal hierarchical attention model for depression detection in social media", Health information science and systems, Vol: 11, No: 1, pp. 6.
8. Sinem Akyol, Year: 2023, "New chaos-integrated improved grey wolf optimization-based models for automatic detection of depression in online social media and networks", PeerJ Computer Science, Vol: 9, pp. e1661.
9. Adebiyi Bukola Taibat; Ihekwoaba Ezinne Olachi; Gideon O. Aleburu; Femi Johnson, Year: 2026, "Particle Swarm Optimization-Convolutional Neural Network (PSO-CNN): An Optimized Deep Learning Model for Personality Recognition from Social Media Profiles", Cureus Journals, Vol: 3, No: 1.
10. Kalyani Pendke; Mayuri Digalwar; Charu Goel, Year: 2026, "Multiclass Stress Detection Using Psychological Emotion Mapping and Transformer Architectures", in IEEE Access, Vol: 14, pp. 38503-38516.
11. Nazzere Oryngozha; Pakizar Shamoi; Ayan Igali, Year: 2024, "Detection and Analysis of Stress-Related Posts in Reddit's Academic Communities", in IEEE Access, Vol: 12, pp. 14932-14948.
12. B. Sathiya; J. Arokia Renjit, Year: 2025, "Mental Health Prediction from Twitter Content Data Based on Stress Inducing Themes and Quantum Neural Network Classifier", in IEEE Access, Vol: 13, pp. 211476-211496.
13. Luna Ansari; Shaoxiong Ji; Qian Chen; Erik Cambria, Year: 2023, "Ensemble Hybrid Learning Methods for Automated Depression Detection", in IEEE Transactions on Computational Social Systems, Vol: 10, No: 1, pp. 211-219.
14. Ensar Arif Sağbaş; Serdar Korukoglu; Serkan Ballı, Year: 2024, "Real-time stress detection from smartphone sensor data using genetic algorithm-based feature subset optimization and k-nearest neighbor algorithm" Multimedia Tools and Applications, Vol: 83, No: 1, pp. 1–32.
15. Muhammad Nadzree Mohd Yamin; Kamarulzaman Ab. Aziz; Tan Gek Siang; Nor Azlina Ab. Aziz, Year: 2023, "Particle swarm optimisation for emotion recognition systems: a decade review of the literature", Applied Sciences, Vol: 13, No: 12, pp. 7054.