

Teen Depression Level Classification Using a Crayfish Optimization Driven KAN-Residual Attention Network Based on Social Media Usage Patterns

Pradheep T Rajan B³

³Assistant Professor, Department of Electronics and Communication Engineering,
Francis Xavier Engineering College, Tirunelveli, Tamil Nadu, India. pradheeptrajan@gmail.com

ABSTRACT: Classifying teen depression levels using social media data is vital because it enables early detection, targeted interventions, and personalized mental health support for adolescents at risk. Therefore, in this paper, a hybrid Kolmogorov-Arnold Network (KAN)- Residual Attention Network (RAN) with Crayfish Optimization (CFO) is proposed for teen depression level classification based on social media usage. Data visualization in depression level classification by transforming complex behavioral, sensor, and psychological datasets, interpretable insights for machine learning models and clinicians. A KAN-RAN uses the feature-preserving nature of residual links and the contextual prioritization of attention mechanisms, this framework significantly enhances the detection of complex, non-linear biomarkers associated with depression. The CFO improves teen depression level classification by smart feature selector and hyperparameter tuner. The performance analysis is carried out using python software and assesses the model's performance using key metrics such as precision, recall, F1-score, confusion matrix, and ROC-AUC. The proposed KAN-RAN-CFO uses nine classes, they are relaxed, mild stress, low stress, moderate stress, elevated stress, high stress, severe stress, extreme stress and critical stress. The proposed KAN-RAN-CFO achieves high accuracy of 98% and better convergence speed. The proposed method outperforms conventional techniques, contributing a robust and reliable tool for teen depression level classification based on social media usage.

Keywords: KAN-RAN, CFO, Teen depression level classification, social media and accuracy.

1. Introduction

Classifying teen depression levels based on social media usage involves analysing time spent, behavioural patterns, and content engagement to identify risk factors for depressive symptoms [1, 2]. Spending more than 2 hours per day on social networks strongly correlates with higher depressive symptoms, anxiety, and emotional distress in adolescents [3]. The risk often compounds for younger teens (ages 12–13). High usage frequently disrupts critical protective factors like sleep quality, physical activity, and face-to-face social connection [4]. The traditional

method of depression level classification relies on standardized clinical diagnostic criteria and structured rating scales that evaluate symptom nature, severity, duration, and functional impact [5,6]. These methods require a patient to actively seek help and schedule an appointment. This often delays classification and treatment until the depression has already escalated to a severe stage [7]. Deep Learning (DL) is essential for depression level classification because it automatically extracts complex, nuanced patterns from diverse data types without manual feature engineering. The conventional techniques such as Adaptive Graph Convolutional Recurrent

Network with Transformer and Whale Optimization Algorithm (AGCRN-T-WOA) [8], Deep Neural Network with Feedback Mechanism (DNN-FM) [9] and Generative Adversarial Network with Graph Neural Network (GAN-GNN) [10] have been used for teen depression level classification based on social media usage. These techniques have some disadvantages like low convergence speed, less reliability and High response time. To avoid these issues, a KAN-RAN-CFO is proposed for teen depression level classification based on social media usage. The contribution of the paper is as follows,

- Data visualization is used to transform complex numerical metrics and categorical model predictions into intuitive visual formats.
- KAN-RAN is proposed to achieve highly accurate, interpretable, and computationally efficient DL framework capable of differentiating multiple levels of depression severity (such as mild, moderate, or severe) from clinical data.
- CFO is proposed to optimize complex model hyperparameters and to maximize classification accuracy while minimizing computational complexity by identifying the most critical behavioural, physiological, or text-based features

2. Recent Works

Depression level classification using AGCRN-T-WOA networks have been presented in [8] to analyse social media usage. It combines graph learning, sequence modeling, and smart tuning to capture complex user interaction patterns. captures complex user interactions over time while automatically finding the best model settings to boost prediction accuracy and speed. However, this mix brings high computing costs, slow training times, and a high risk of overfitting on noisy data. The hybrid DNN-FM have been proposed in [9] for depression level classification. DNN-FM struggle with social media usage due to

high computational costs, slow training times from sequential processing, difficulty handling sudden slang shifts, and vanishing gradients that lose long-term context in fast-moving text threads. [10] presented the GAN-GNN for social media analysis, which improves teen depression level classification by generating realistic synthetic data to fix class imbalances, modeling complex user-connection structures, and enhancing resistance against adversarial attacks and fake account manipulation. However, it produces high computational costs, severe training instability, data sparsity issues, and vulnerability to adversarial attacks, making real-time deployment difficult and model optimization complex. To overcome these drawbacks, a KAN-RAN-CFO is proposed in this paper for teen depression level classification based on social media usage.

3. Proposed Description

Classifying teen depression levels based on social media usage is crucial for early intervention, targeted support, and preventing severe mental health crises. The proposed block diagram of KAN-RAN-CFO for teen depression level classification based on social media usage is represented in figure 1. The process begins with the raw input named as teen social media dataset. To ensure the data is suitable for modelling, it is given to the data quality verification. This stage involves two parallel steps they are data preparation and a consistency check. Once the initial preparation is complete, the data is cleansed through missing value imputation and outlier suppression. Then, the dataset moves to data visualization, allowing to visually inspect patterns and distributions and the data is sent to feature normalization to rescale the variables. This phase employs min-max scaling alongside a feature correlation analysis to determine dependencies between variables and reduce redundant features.

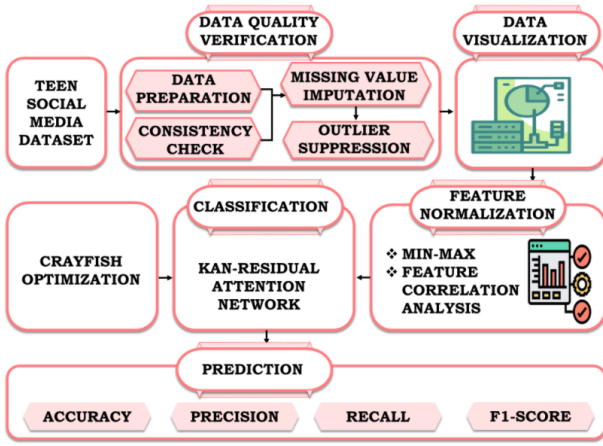


Figure 1: Proposed block diagram of KAN-RAN-CFO for teen depression level classification

The normalized features are fed into the classification block such as KAN-RAN, which is used to classify the teen depression level based on social media usage. To maximize this neural network's performance, a metaheuristic optimization technique called CFO is applied to fine-tune the network's parameters or select the best feature subsets. Finally, the optimized model generates a prediction and the success and effectiveness of the entire architecture are evaluated using four standard performance metrics: accuracy, precision, recall, and F1-Score.

3.1 Modelling of Data Visualization

The process of turning raw numbers into clear data like charts and graphs is named as data visualization.

Mathematical expression for mean (average) to check data center:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

Mathematical expression for min-max feature scaling:

$$x_{normalized} = \frac{x_i - \min(X_{train})}{\max(X_{train}) - \min(X_{train})} \quad (2)$$

Crucially, the min and max values are calculated only from the training data (X_{train}). The same parameters are then used to transform the test data.

A KAN-RAN is used after initial visualization or feature extraction to dynamically capture global

context and complex long-range dependencies. It refines visual or spatial data by adjusting to varying shapes, noise, or multi-scale structures.

3.2 Mathematical modelling of KAN-RAN

A KAN-RAN is an advanced hybrid DL architecture that merges KANs with residual connections and attention modules to capture complex non-linear patterns with high parameter efficiency. Figure 2 shows the architecture of KAN-RAN for depression level classification.

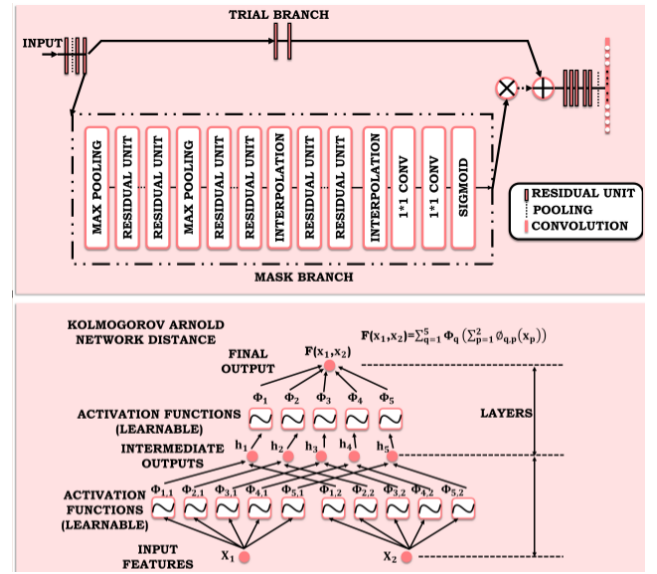


Figure 2: Architecture of KAN-RAN

3.2.1 Input Representation and Local Feature Extraction

Let the input feature map be $X \in R^{B \times C \times H \times W}$, where B is the size of batch, C is the quantity of networks, H is the height, and W is the width.

3.2.2 Channel and Spatial Attention (Attention Module)

The attention module calculates weights to focus on important features.

- **Channel Attention:** Calculates channel weights using average and max pooling followed by multi-layer perceptron or linear layers.
- **Spatial Attention:** Computes spatial weights using pooling across channels.

3.2.3 KAN Layer

Instead of fixed activation functions on nodes (like standard MLPs), KAN applies learnable 1D functions on the edges (weights).

For an input vector z , a KAN layer transforms it as:

$$y_i = \sum_{j=1}^{d_{in}} \phi_{i,j}(z_j) \quad (4)$$

where $\phi_j(x) = w_i(b(x)) + \text{split}(x)$ with $b(x)$ being a base activation function (like SiLU) and $\text{split}(x)$ being a linear combination of B-splines. Applying the KAN transformation to the attention-refined features X' gives:

$$F_{kan} = \text{KAN}(X') \quad (5)$$

3.2.4 Residual Connection

The residual connection adds the original input X directly to the output of the KAN transformation to prevent vanishing gradients and retain low-level details.

The final output Y is:

$$Y = X + F_{kan} \quad (6)$$

By capturing long-range spatial and temporal dependencies, the KAN-RAN ignores conversational or physiological background noise and focuses only on clinical indicators of depression severity. Optimization is required after classification to adjust model weights and minimize error and for tuning hyperparameters.

3.3 Modelling of CFO

The CFO is a meta-heuristic optimization algorithm inspired by the summer behaviour, foraging, and predator-avoidance strategies of crayfish.

Here is the step-by-step process of using COA for hyperparameter tuning, complete with the mathematical equations that drive each phase. Flow chart of CFO for teen depression level classification is represented in figure 3.

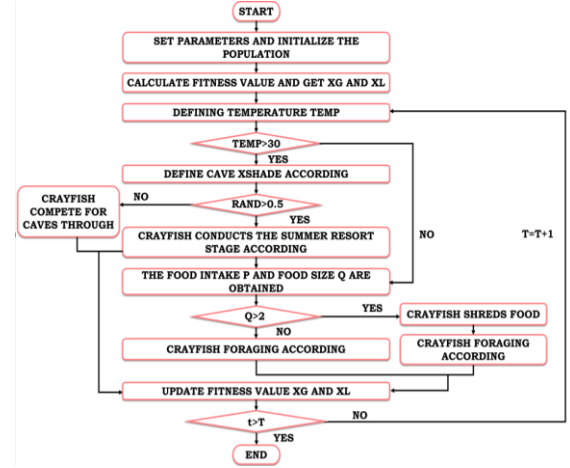


Figure 3: Flowchart of CFO

3.3.1 Initialize the Population and Parameters

Define the search space based on the hyperparameters. The algorithm initializes a population of crayfish, where each crayfish represents a specific combination of hyperparameters.

$$X_i = LB + rand(0,1) \times (UB - LB) \quad (7)$$

The X_i is i^{th} crayfish vector (a set of hyperparameters). UB and LB are the Lower and upper bounds of your hyperparameters.

3.3.2 Evaluate Fitness (Model Training)

For each crayfish (hyperparameter set), train DL model and calculate its performance. The fitness value is typically the validation error, loss, or 1 - accuracy and find the best crayfish position X_{best} which yields the lowest validation error.

3.3.3 Compute Environmental Factors

Crayfish behavior changes based on the surrounding temperature (T) and food availability (F). These factors dictate whether the algorithm explores new hyperparameter combinations or exploits known good ones.

$$T = \exp\left(-\frac{t}{T_{Max}}\right) \quad (8)$$

$$F = C_1 \times rand(0,1) \times T \quad (9)$$

T is the current iteration, T_{Max} is the maximum number of iterations, and C_i is a constant parameter to control food abundance.

3.3.4 Iterative Optimization Phases

In every iteration, the algorithm checks the temperature (T) to decide which behavioural-phase to execute. A penalty profile that adaptively reduces search steps as iterations increase, calculated as:

$$p(t) = C_3 \times \left(1 - \frac{t}{T_{Max}}\right) \quad (10)$$

3.3.5 Boundary Control and Selection

After updating the hyperparameters, ensure iterations do not exceed defined limits (LB and UB). If a new hyperparameter set yields a lower validation error than its previous configuration, update that crayfish's memory with the new position.

3.3.6 Termination

Repeat Steps 3.3.3 through 3.3.5 until the maximum iterations T_{Max} are reached or the validation loss stops improving. The final X_{best} outputs the optimum hyperparameters for your model.

KAN-RAN-CFO models are accurately classifying depression across multiple distinct levels of severity. This allows care providers to instantly distinguish between a teenager experiencing mild, transient distress and one requiring immediate, crisis-level suicide prevention.

4. Results and Discussion

The KAN-RAN-CFO is proposed for teen depression level classification, which is evaluated using python software. The name of dataset used for simulation is student depression and lifestyle and includes nine classes, such as relaxed, mild stress, low stress, moderate stress, elevated stress, high stress, severe stress, extreme stress and critical stress, which are shown in table 1.

Table 1: Stress level type

Number	Level type
1	Relaxed
2	Mild stress

3	Low stress
4	Moderate stress
5	Elevated stress
6	High stress
7	Severe stress
8	Extreme stress
9	Critical stress

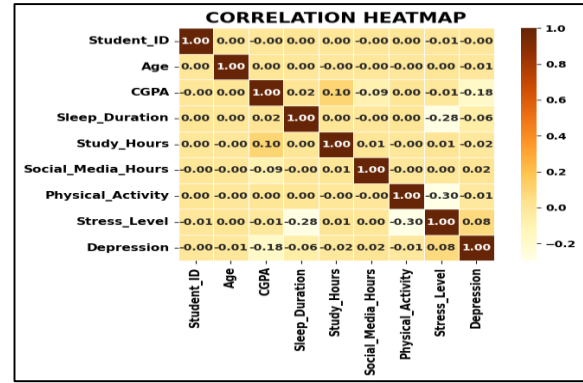


Figure 4: Correlation Heatmap

Figure 4 displays a correlation heatmap that illustrates the linear relationships between various student metrics, such as study habits, sleep, and mental health. The values in the grid range from -0.30 to 1.00, which is strong positive correlation and lighter shades signify weak or negative relationships. Based on KAN-RAN classification, Stress_Level shares the strongest negative correlations in the dataset with both Physical_Activity and Sleep_Duration.

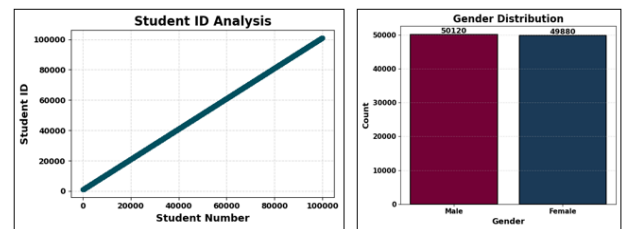


Figure 5: Student ID analysis and gender distribution

Figure 5 shows the student ID analysis and gender distribution. Student ID analysis is indicating that IDs are assigned sequentially from 1 to 100,000. In gender distribution, there are 50,120 male students and 49,880 female students, showing a highly balanced demographic distribution across the population based on social media usage.

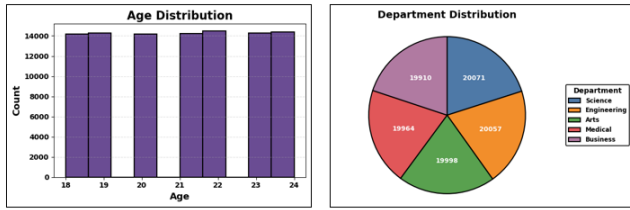


Figure 6: Age and department distribution

Figure 6 shows the distribution of age and department distribution. The age distribution: shows a balanced, uniform distribution across seven age groups from 18 to 24, with each individual age bar consistently reaching a count of approximately 14,000. The department distribution divides a total population of 100,000 individuals into five nearly equal sectors.

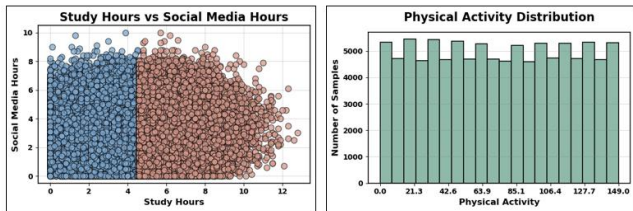


Figure 7: Study hours vs social media hours and physical activity distribution

Figure 7 shows the study hours vs social media hours and physical activity distribution, The study hours vs social media hours scatter displays study times ranging from 0 to 13 hours against social media usage that peaks near 10 hours. As study hours surpass 10 hours, the maximum time spent on social media begins to decrease. Meanwhile, the physical activity distribution histogram displays a uniform spread of data across a physical activity scale from 0.0 to 149.0. The number of samples remains consistent throughout the range, fluctuating tightly between 4,600 and 5,500 samples.

Table 2: Classification report

Classes	Precision	Recall	F1-Score	support
Relaxed	0.9863	0.9794	0.9828	2718
Mild stress	0.9911	0.9790	0.9850	4533
Low stress	0.9915	0.9788	0.9851	4665

Moderate stress	0.9948	0.9812	0.9880	4905
Elevated stress	0.9803	0.9824	0.9814	2331
High stress	0.8959	0.9734	0.9330	451
Severe stress	0.8559	0.9867	0.9167	301
Extreme stress	0.6351	0.9895	0.7737	95
Critical stress	0.0169	1.0000	0.0333	1
Accuracy			0.9800	20000
Macro avg	0.8164	0.9834	0.8421	20000
Weighted avg	0.9843	0.9800	0.9818	20000

Table 2 shows the classification report, which includes nine classes, they are relaxed, mild stress, low stress, moderate stress, elevated stress, high stress, severe stress, extreme stress and critical stress. The class moderate stress has highest precision of 0.9948, the class critical stress has highest recall of 1.0000 and the class moderate stress has highest F1-Score of 0.9880. The classification report shows the accuracy of proposed KAN-RAN-CFO is 98%.

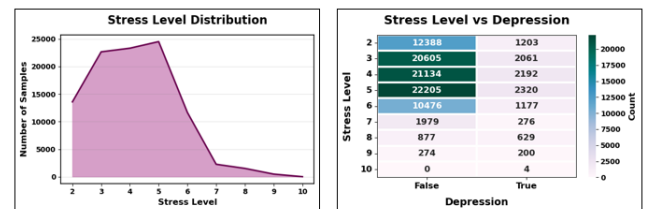


Figure 8: Stress level distribution and stress level vs depression

Stress level distribution and stress level vs depression are shown in figure 8. The stress level distribution area shows that the sample size peaks significantly between stress levels 3 and 5, with level 5 containing the highest frequency of nearly 25,000 samples before sharply declining toward level 10. In the stress level vs depression heatmap, the vast majority of individuals do not experience depression, as evidenced by a peak count of 22,205 non-depressed individuals at stress level 5.

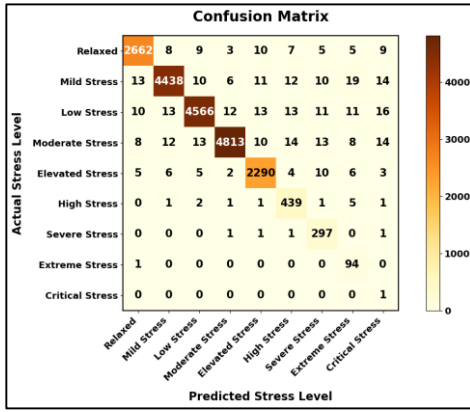


Figure 10: Confusion matrix

Figure 10 shows the confusion matrix. The main diagonal highlights the true positive counts where the model predicted correctly, such as 4,566 correct predictions for Low Stress and 4,813 for Moderate Stress. Off-diagonal numbers represent misclassifications, showing that the model occasionally confuses adjacent states, like predicting Mild Stress 13 times when the actual state was Relaxed. Overall, the high values concentrated along the diagonal indicate strong predictive accuracy across most categories, particularly for Low, Mild, and Moderate stress classes. The sample sizes vary significantly across classes, ranging from a high of 4,813 correct instances for Moderate Stress down to just 1 correct instance for Critical Stress.

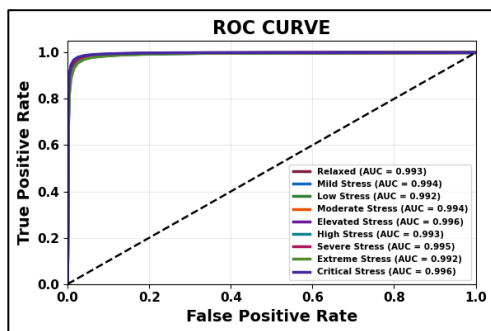


Figure 11: ROC Curve

Figure 11 shows the ROC Curve of proposed KAN-RAN-CFO. The AUC values of class relaxed is 0.993, mild stress is 0.994, low stress is 0.992, moderate stress is 0.994, elevated stress is 0.996, high stress is 0.993, severe stress is 0.995, extreme stress is 0.992 and critical stress is 0.996. The AUC values shows that the effectiveness of

proposed KAN-RAN-CFO in teen depression level classification based on social media usage.

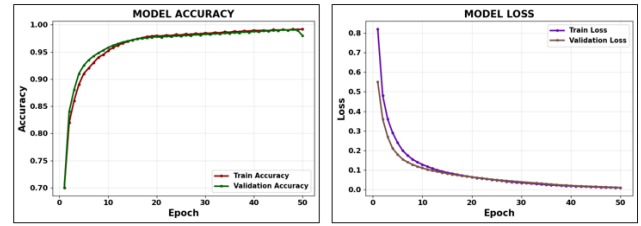


Figure 12: Accuracy and loss of proposed KAN-RAN-CFO

The accuracy and loss of proposed KAN-RAN-CFO are represented in figure 11. The accuracy of proposed KAN-RAN-CFO is 98% and the loss of proposed KAN-RAN-CFO is 10%, which demonstrates the effectiveness of proposed system.

Table 3: Comparison of response time and accuracy

Method	Response time	Accuracy
AGCRN-T-WSO [8]	0.7s	97%
DNN-FM [9]	0.6s	97.34%
GAN-GNN [10]	0.5s	96%
Proposed KAN-RAN-CFO	0.1s	98%

Table 3 shows the comparison of response time and accuracy. The response time of proposed KAN-RAN-CFO is 0.1s. Lower response time in a KAN-RAN-CFO classification model enables real-time edge deployment, reduced compute costs, and instant decision-making. The accuracy of proposed KAN-RAN-CFO is 98%. High accuracy in a KAN-RAN-CFO for teen depression level classification ensures precise predictions of depression level. It reduces false results, severe depression level, less social media usage, and improves content recommendations by understanding complex user behavior patterns.

5. Conclusion

In this paper, a hybrid KAN-RAN-CFO has been proposed for teen depression level classification

based on social media usage. The KAN-RAN-CFO are effective at classifying teen depression level, utilizing optimized feature sets allows for efficient, real-time detection without requiring massive, computationally expensive models. The developed method is compared with existing methods AGCRN-T-WSO, DNN-FM and GAN-GNN in terms of precision, recall and F1-Score. The performance analysis is carried out using python software. The DL models, KAN-RAN-CFO offer high accuracy of 98%, and less computational time of 0.1s in teen depression level classification. In future, this work will be implemented using advanced AI technologies and hybrid optimization techniques.

References

1. Rozi Karimi; Milad Baghalzadeh Shishehgarkhaneh; Robert C. Moehler; Yihai Fang; Saima Ahmad, Year: 2025, "Social Media Practices on Construction Sites: A New Conceptual Model of Social Media Impact on Team Performance", Vol: 53, No: 6, pp. 147 – 167.
2. Xinyi Wang; Loo-See Beh, Year: 2025, "Navigating New Terrain: Diverse Effects of Social Media on Employee Performance in China's Social Commerce Sector", IEEE Transactions on Professional Communication, Vol: 68, No: 2, pp. 209 – 227.
3. Thomas F. Stafford; B. Q. Duong, Year: 2023, "Social Media in Emerging Economies: A Cross-Cultural Comparison", IEEE Transactions on Computational Social Systems, Vol: 10, No: 3, pp. 1160-1178.
4. Praboda Rajapaksha; Reza Farahbakhsh, Year: 2019, "Scrutinizing News Media Cooperation in Facebook and Twitter", IEEE Access, Vol: 13, pp. 123355-123368.
5. Zixing Shen; S. Tan; M. J. Pritchard, Year: 2022, "Understanding the Effects of Visual Cueing on Social Media Engagement with YouTube Educational Videos", IEEE Transactions on Professional Communication, Vol: 65, No: 2, pp. 337-350.
6. Wei Wang; S. X. Zhang; Q. Liang; W. Deng, Year: 2024, "Social Media Levels the Playing Field of Entrepreneurial Entry for Disadvantaged Populations", IEEE Transactions on Engineering Management, Vol: 71, pp. 9776-9789.
7. Ujun Jeong; Bohan Jiang; Zhen Tan; H. Russell Bernard; Huan Liu, Year: 2024, "Descriptor: A Temporal Multi-network Dataset of Social Interactions in Bluesky Social (BlueTempNet)", IEEE Data Descriptions, Vol: 1, pp. 71-79.
8. Chen Zhang; Yue Wu; Ya Shen; Shengzhao Wang; Xuhui Zhu; Wei Shen, Year: 2024, "Adaptive Graph Convolutional Recurrent Network with Transformer and Whale Optimization Algorithm for Traffic Flow Prediction", Mathematics, Vol: 12, No: 10, pp. 1493.
9. Wanlun Ma; Xiangyu Hu; Chao Chen; Sheng Wen, Year: 2022, "Social Media Event Prediction using DNN with Feedback Mechanism", ACM Transactions on Management Information Systems, Vol: 13, No: 3, pp. 1-24.
10. Ash Mohammad Abbas, Year: 2021, "Social network analysis using deep learning: applications and schemes", Social Network Analysis and Mining, Vol: 11, No: 1, pp. 106.
11. Jeya Bright Pankiraj; Govindaraj Vishnuvarthanan; Yudong Zhang; Murugan Pallikonda Rajasekaran; Ramaraj Kottamalai, Year: 2023, "Development of scalable coding on Encrypted Images using BTC for different Block Size", Journal of Signal, Image and Video Processing, Vol: 17, No: 7, pp. 3821-3828.

12. Kulandaivel Maruthamuthu Paramasivam; Prabakaran MP; Jeya Bright Pankiraj, Year: 2025, "CycleGAN algorithm based enhancement of night vision halation images from multiple sources", AIP Advances, Vol: 15, No: 7, pp. 075031-1 - 075031-6.
13. Dilip Kumar Sharma; Ningthoujam Chidananda Singh; Daneshwari A; Noola; Amala Nirmal Doss; Janaki Sivakumar, Year: 2025, "A review on various cryptographic techniques & algorithms", Materials Today: Proceedings, Vol: 51, pp. 104-109.
14. Thamilarasi V, Year: 2024, "Artificial intelligence-driven smart scenic management: Automated decision making and optimization", Int. J. Intell. Syst. Appl. Eng, Vol: 12, pp. 2731.