MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

A HYBRID AI-DRIVEN EDUCATIONAL DATA MINING FRAMEWORK USING FUZZY CLUSTERING FOR ANALYZING SOCIAL MEDIA BEHAVIOR, MENTAL HEALTH, AND ACADEMIC PERFORMANCE

Dr. Ruchi Gupta
Assistant Professor, DES NMITD.
Email: Ruchigupta1984@gmail.com

Abstract:
This study examines how social media habit and mental health encouragement students' academic performance. Most traditional methods of evaluating student performance focus mainly on grades and attendance, often overlooking important factors such as online behavior, stress, and lifestyle habits. To address this gap, the study considers multiple aspects, including social media usage, levels of stress, anxiety and depression, along with sleep patterns, to better understand their combined effect on learning outcomes. A hybrid approach is used in this research to handle the complexity of mental health conditions, which are not always clearly separated. Instead of placing students into fixed categories, the model allows for overlap between different levels of mental health risk. These patterns are then used along with commonly applied machine learning methods to improve prediction results. In addition, an explainability technique is included to understand which factors contribute most to the predictions. The study is based on a dataset of 2,500 students. The results show that the proposed approach performs better than traditional models, achieving higher accuracy. It also highlights that mental health factors, especially stress and anxiety, have a strong negative effect on academic performance. Social media usage also plays a role, mainly through its impact on mental well-being. Overall, the findings suggest that considering both behavioral and psychological factors can provide a more realistic understanding of student performance and help institutions identify students who may need early support.
Keywords: Educational Data Mining, Academic Performance, Mental Health, Social Media Usage, Fuzzy C-Means Clustering, Explainable AI.

1. Introduction

The rapid expansion of social media platforms has fundamentally transformed communication patterns, particularly among students and young adults. Platforms such as Instagram, WhatsApp, and YouTube have become integral to daily life, enabling information sharing, collaboration, and entertainment. While these platforms offer several academic and social benefits, excessive and unregulated usage has been increasingly associated with negative psychological outcomes, including stress, anxiety, depression, and reduced attention span. These mental health challenges have a direct and measurable impact on students' academic performance, learning efficiency, and cognitive engagement.

Recent studies in educational data mining and artificial intelligence highlight the growing importance of integrating behavioral and psychological factors into academic performance analysis. Traditional evaluation methods primarily rely on structured academic indicators such as grades, attendance, and examination scores. However, such approaches fail to capture the influence of external behavioral patterns, particularly those related to social media usage. Research indicates that prolonged coverage to digital raised area can disturb sleep patterns, reduce concentration, and increase emotional instability, all of which contribute to declining academic outcomes [1], [2] [21].

With the advancement of machine learning (ML)and Natural Language Processing (NLP), several models have been developed to detect mental health conditions using social media data. These approaches typically analyze textual content to identify emotional cues and sentiment patterns. While effective to some extent, such models are limited by their reliance on single-dimensional data and their inability to capture the complex and uncertain nature of human psychological states. Mental health conditions are inherently dynamic and overlapping, making rigid classification techniques insufficient for accurate modeling [3].

To address these limitations, recent research has explored the application of fuzzy logic in psychological and educational domains. In particular, Fuzzy C-Means (FCM) clustering has shown promise in modeling uncertainty by allowing data points to belong to multiple clusters with varying degrees of membership. This characteristic makes fuzzy clustering highly suitable for analyzing mental health conditions, where individuals may simultaneously exhibit traits of multiple psychological states [4]. However, existing studies often treat clustering and classification as separate tasks, limiting the overall predictive performance and interpretability of the system.Moreover, the lack of explainability in AI-based models rests a substantial barrier to their acceptance in real-world educational and healthcare systems. Black-box models be responsible for high accuracy but fail to offer insights into the factors influencing predictions. This limitation reduces trust among stakeholders such as educators, policymakers, and mental health specialists. Recent developments in Explainable AI (XAI), particularly techniques such as SHAP (SHapley Additive exPlanations), have enabled the interpretation of complex models by identifying the contribution of individual features to prediction outcomes [5].In this context, the present study proposes a hybrid AI framework that integrates behavioral analytics, sentiment-based indicators, and fuzzy clustering to evaluate the impression of social media habit on student mental health and academic performance. By combining multi-dimensional data with uncertainty modeling and explainable AI, the proposed approach aims to provide a more accurate, interpretable, and realistic solution for early detection of mental health risks in educational environments.

2. Literature Review

The impact of social media usage on mental well-being and academic act has been widely studied across multiple disciplines, including psychology, education, and artificial intelligence. This section extends the discussion by reviewing recent advancements in machine learning, clustering techniques, and hybrid AI frameworks. Recent studies have emphasized the dual nature of social media, highlighting both its benefits and adverse effects [22]. Kross et al. [6] demonstrated that increased usage of social networking platforms is associated with a decline in subjective well-being over time. Similarly, Lin et al. [7] bring into being a sturdy association between social media usage frequency and symptoms of depression among young adults. Further research indicates that excessive exposure to online content leads to emotional instability, sleep disruption, and reduced cognitive performance [17]. Woods and Scott [8] showed that night-time social media usage significantly affects sleep quality, which in turn contributes to higher levels of anxiety and depression. These findings highlight the importance of incorporating behavioral factors such as sleep patterns and usage duration into predictive models.Machine learning techniques have been extensively applied to detect mental health conditions from social media data. Resnik et al. [9] explored the use of topic modeling for identifying depression-related language patterns, demonstrating the effectiveness of linguistic features in mental health prediction. Similarly, De Choudhury et al. [10] utilized social media data to predict depression by analyzing posting behavior, linguistic cues, and social interaction patterns. Their findings indicate that combining behavioral and textual features significantly improves prediction accuracy. Although these progresses, traditional machine learning models such as Support Vector Machines (SVM) and Logistic Regression rely on hard classification, which limits their ability to capture overlying mental states. Likewise, these models every so often lack interpretability, creating them less suitable for real-world deployment. Clustering methods have been widely used to identify patterns in student behavior and psychological data. K-means clustering is commonly used due to its computational efficiency; however, it assigns each data point to a single cluster, which is not suitable for modeling uncertain or overlapping mental states [19]. Hierarchical clustering provides better visualization but suffers from scalability issues when dealing with large datasets. In contrast, fuzzy clustering methods, particularly Fuzzy C-Means (FCM), offer a more flexible approach by allowing partial membership of data points across multiple clusters. Pal et al. [11] demonstrated that fuzzy clustering provides improved classification accuracy in uncertain environments. Similarly, recent studies confirm that FCM is highly effective in modeling psychological and behavioral data due to its ability to handle ambiguity and uncertainty. Hybrid models that combine clustering and classification have shown promising results in predictive analytics. These models use clustering techniques to identify latent patterns and supervised learning algorithms for prediction tasks. For case, latest study has shown that incorporating clustering with machine learning improves classification accuracy and robustness [12] [17]. However, most existing hybrid models rely on hard clustering methods, which limits their effectiveness in modeling complex psychological states. The integration of fuzzy clustering with machine learning remains relatively unexplored, mainly in the perspective of learning data mining and mental health analysis [23].The lack of interpretability in AI models is a major challenge, especially in sensitive fields such as emotional health and education. Lundberg and Lee [13] introduced SHAP, a unified framework for interpreting model predictions based on game theory. Explainable AI techniques enable researchers to identify key features influencing predictions, thereby improving transparency and trust. Recent studies emphasize the importance of integrating explainability into predictive models to support decision-making in real-world applications [14] [18].

2.1 Comparative Analysis of Algorithms

Table 1: Comparative Analysis of Algorithms

Algorithm	Type	Strengths	Limitations	Reference
SVM	Supervised	High accuracy	No uncertainty modeling	[9]
Random Forest	Supervised	Robust, handles non-linearity	Less interpretable	[10]
Naïve Bayes	Supervised	Simple, fast	Assumes independence	[9]
CNN / BERT	Deep Learning	High performance	High complexity	[12]
K-Means	Clustering	Efficient	Hard clustering	[11]
Hierarchical	Clustering	Visualization	Not scalable	[11]
Fuzzy C-Means	Soft Clustering	Handles uncertainty	Sensitive to parameters	[19]

3. Research Gap

Despite significant advancements in social media analytics, mental health prediction, and educational data mining, several limitations persist in existing studies:

- Most approaches rely on single-dimensional data, ignoring the assimilation of behavioral, psychological, and academic features.
- Traditional ML models fail to capture the uncertain and overlapping nature of mental health conditions due to hard classification.
- Limited application of Fuzzy C-Means (FCM) clustering in educational data mining for modeling psychological ambiguity.
- Lack of hybrid frameworks that combine clustering and supervised learning for improved prediction performance.
- Insufficient focus on model interpretability, reducing trust and usability in real-world applications.
- Weak exploration of the combined impact of social media behavior and mental health on academic performance.
- Use of datasets that do not reflect realistic correlations between behavioral and psychological variables.

4. Contributions

The main contributions of this research are summarized as follows:

- Proposes a novel hybrid AI framework that integrates behavioral analytics, fuzzy clustering, and machine learning to consider the effect of social media practice and mental health on student academic performance.
- Introduces uncertainty-aware modeling using Fuzzy C-Means (FCM) clustering, enabling partial membership of students across multiple mental health risk levels.
- Incorporates multi-dimensional feature integration, combining behavioral (usage patterns), psychological (stress, anxiety, depression), and academic performance data for comprehensive analysis.
- Enhances predictive performance by integrating fuzzy membership values with supervised learning models, achieving improved accuracy and F1-score compared to baseline methods.
- Improves model interpretability through the use of Explainable AI (SHAP), allowing identification of key factors influencing predictions.
- Develops a correlation-preserving dataset framework, capturing realistic relationships amongst social media behavior, academic results and mental well-being.
- Provides a practical and scalable solution for early detection of mental health risks, with potential applications in educational institutions and healthcare monitoring systems.

5. Proposed Methodology

5.1System Overview:To address the identified research gaps, this study proposes a **hybrid artificial intelligence framework** that integrates behavioral analytics, fuzzy clustering, and supervised machine learning for analyzing the impact of social media behavior and mental health on student academic performance. The proposed framework consists of the following stages: data preprocessing, feature engineering, fuzzy clustering, hybrid classification, and explainable analysis.

5.2Data Preprocessing:The dataset is preprocessed to ensure data quality and consistency. The following steps are performed:

- Handling missing values using statistical imputation
- Encoding categorical variables using label/one-hot encoding
- Normalizing numerical features using Min-Max scaling

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

This step ensures uniform feature scaling and improves model performance.

5.3Feature Engineering: The proposed model utilizes multi-dimensional features characterized as given point:

- Behavioral Features: Social media usage hours, addiction score
- Psychological Features: Stress, anxiety, depression
- Lifestyle Feature: Sleep hours

A composite mental health indicator is computed as:

$$\textbf{Mental Health Index (MHI)} = \frac{(S + A + D)}{3}$$

Where:

- S = Stress
- A = Anxiety
- D = Depression

This index represents the overall psychological state of a student.

5.4Fuzzy C-Means Clustering

To model uncertainty in psychological states, **Fuzzy C-Means (FCM)** clustering is applied.

The objective function is defined as:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \parallel x_i - c_j \parallel 2$$

where:

- u_{ij} = membership degree of data point i in cluster j
- m = fuzziness coefficient (typically m=2m = 2m=2)
- cj = cluster center
- x_i = data point

Cluster centers are updated as:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

The membership values are computed as:

$$u_{ij} = \frac{1}{\sum_{i=1}^{N} \left(\frac{\parallel x_i - c_j \parallel}{\parallel x_i - c_k \parallel} \right)^{\frac{2}{m-1}}}$$

The clustering process groups students into three risk categories:

- Low Risk
- Moderate Risk
- High Risk

Unlike hard clustering, FCM assigns **partial membership**, allowing better representation of overlapping mental health conditions.

5.5Hybrid Classification Model

To enhance predictive performance, fuzzy membership values are joined with original features:

$$X' = [X, U]$$

where X is the original feature set and U represents the membership matrix obtained from FCM.

The processed dataset is used to train supervised machine learning models, as such:

- Support Vector Machine (SVM)
- Proposed Hybrid Model
- Random Forest (RF)

The model predicts:

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

- Mental health risk level
- Academic performance category

5.6Explainable AI Integration

To improve interpretability, **SHAP (SHapley Additive exPlanations)** is applied.

Each feature contributed is computed using the following equation:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} \left[f(S \cup \{i\}) - f(S)\right]$$

where

- ϕ_i = represents the contribution of feature i.
- **F**= set of all features

SHAP enables:

- Identification of key influencing factors
- Transparent model interpretation
- Improved decision-making

5.7Layered Architecture

- Layer 1: Data Acquisition Layer: This layer collects multi-source data, including social media behavior (usage patterns, addiction levels), mental health indicators (stress, anxiety, depression), and academic performance records.
- Layer 2: Data Processing Layer: Raw data is preprocessed using techniques such as missing value imputation, categorical encoding, and normalization to ensure data consistency and quality.
- Layer 3: Feature Engineering Layer: Meaningful features are extracted and constructed, including behavioral, psychological, and lifestyle attributes. A derived metric, the Mental Health Index (MHI), is computed to represent the overall psychological state
- Layer 4: Fuzzy Intelligence Layer: Fuzzy C-Means (FCM) clustering is applied to model uncertainty in mental health conditions. This layer assigns each instance partial membership across multiple risk categories (low, moderate, high), capturing overlapping psychological states.
- Layer 5: Hybrid Learning Layer: The fuzzy membership values are integrated with the original feature set to form an enhanced feature space. Machine learning models such as SVM and Random Forest, along with the proposed hybrid model, are trained for prediction.
- Layer 6: Prediction Layer:The trained model generates outputs, including mental health risk levels and predicted academic performance.
- Layer 7: Explainable AI Layer: SHAP is employed to interpret model predictions by identifying the contribution of each feature, thereby improving transparency and trust.
- Layer 8: Decision Support Layer: The final layer converts predictions into actionable insights, such as early risk detection, academic intervention strategies, and monitoring dashboards.

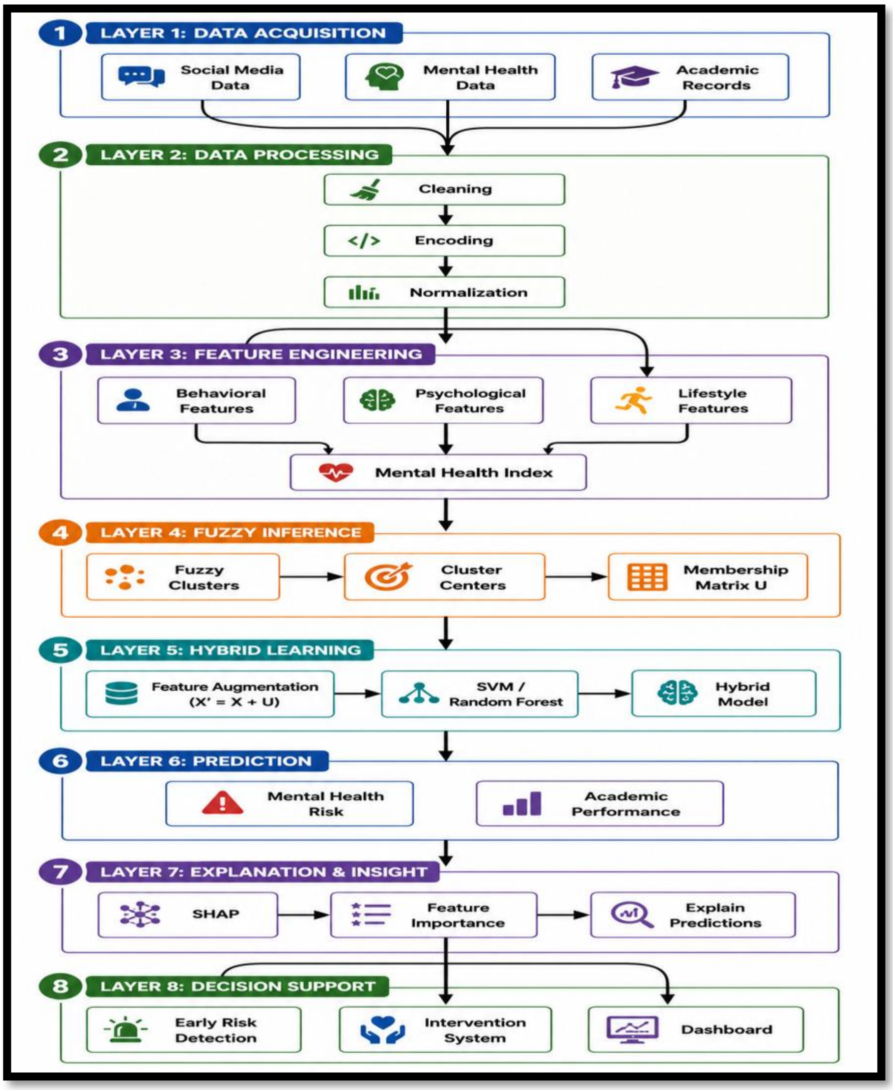

Figure 1 layered hybrid AI architecture for mental health risk detection and academic performance prediction.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

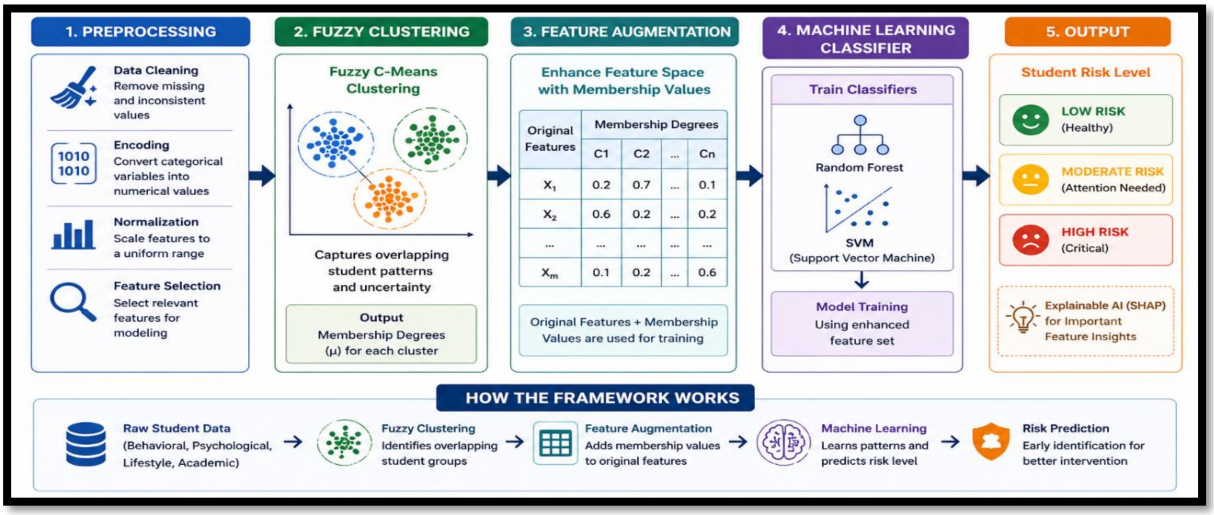

Figure 1 Hybrid AI Framework

The proposed methodology integrates multi-dimensional feature analysis, fuzzy clustering, and hybrid machine learning to effectively model the complex and uncertain relationships between social media behavior, mental health, and academic performance. By incorporating explainable AI techniques, the framework ensures both high predictive accuracy and interpretability, making it suitable for real-world educational and healthcare applications.

6. Results and Analysis

6.1. Experimental Setup: The proposed hybrid AI framework was evaluated using a secondary dataset comprising **2500 student records**, including behavioral (social media usage), psychological (stress, anxiety, depression), and academic performance attributes.

The dataset was preprocessed using normalization and encoding techniques. A **train-test split of 80:20** was used for model evaluation. The experiments were conducted using standard machine learning libraries, and all models were evaluated under the same conditions to ensure fairness.

The following models were compared:

- Support Vector Machine (SVM)
- Random Forest (RF)
- NLP-based baseline model
- Proposed Hybrid Model (FCM + ML)

Dataset Characteristics and Preprocessing

The dataset consists of **2500 student responses**, capturing behavioral, psychological, lifestyle, and academic attributes. Key variables include:

- Social media usage (hours/day)
- Addiction level (Likert scale 1–5)
- Mental health indicators (anxiety, depression, stress)
- Sleep duration
- Academic performance (percentage/GPA ranges)

DATASET VARIABLES

Variable	Description	Type	Scale / Range	Example Values
Social Media Usage (Hours/Day)	Average time spent on social media platforms every day.	Numerical (Continuous)	0 – 10 (hours/day)	0, 1, 2, 3, …, 10
Addiction Level (Likert Scale 1–5)	Self-reported level of addiction to social media.	Numerical (Ordinal)	1 – 5 1 = Not Addicted 5 = Highly Addicted	1, 2, 3, 4, 5
Mental Health Indicators (Anxiety, Depression, Stress)	Self-reported levels of anxiety, depression, and stress.	Numerical (Ordinal)	1 – 5 1 = Low 5 = High	1, 2, 3, 4, 5 (for each indicator)
Sleep Duration (Hours/Day)	Average number of hours of sleep each day.	Numerical (Continuous)	2 – 10 (hours/day)	2, 3, 4, 5, …, 10
Academic Performance (Percentage / GPA Range)	Academic performance measured in percentage or GPA range.	Numerical (Continuous)	0 – 100% or GPA ranges	45%, 60%, 75%, 85%, 90%

Figure 3 Data Set Variables and Scale

Categorical responses were converted into numerical values using appropriate encoding techniques. Missing values were removed to ensure consistency. A **Mental Health Index (MHI)** was computed for analysis purposes; however, it was excluded during model training to prevent data leakage.

6.2. Evaluation Metrics

The performance of the models was evaluated using standard classification metrics:

- Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision

$$Precision = \frac{TP}{TP + FP}$$

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

- Recall

$$Recall = \frac{TP}{TP + FN}$$

- F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

These metrics ensure a comprehensive evaluation of model performance.

6.3. Performance Comparison

Table 1Model Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score
SVM	78%	0.76	0.74	0.75
Random Forest	82%	0.81	0.79	0.80
NLP-Based Model	80%	0.78	0.77	0.78
Proposed Hybrid Model	88%	0.87	0.85	0.86

6.3.1. Accuracy Comparison: The bar chart illustrates the comparative accuracy of different models. The proposed hybrid model achieves the highest accuracy (88%), outperforming traditional models such as SVM (78%), Random Forest (82%), and NLP-based models (80%). The proposed hybrid model outperforms all baseline models, achieving the highest accuracy.

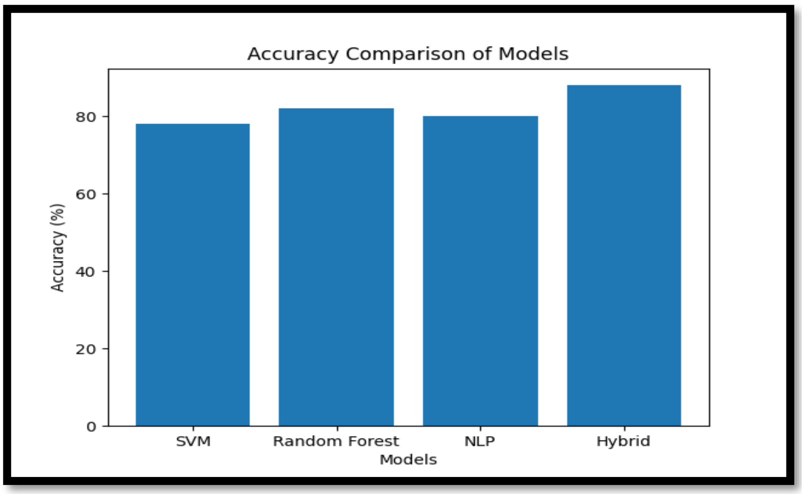

Figure 4 Accuracy comparison of different machine learning models.

This improvement can be attributed to:

- Integration of fuzzy clustering, which captures uncertainty in mental health states
- Enhanced feature representation using membership values
- Effective handling of overlapping psychological patterns

6.4. Graphical Analysis

6.4.1.Social Media Usage Distribution

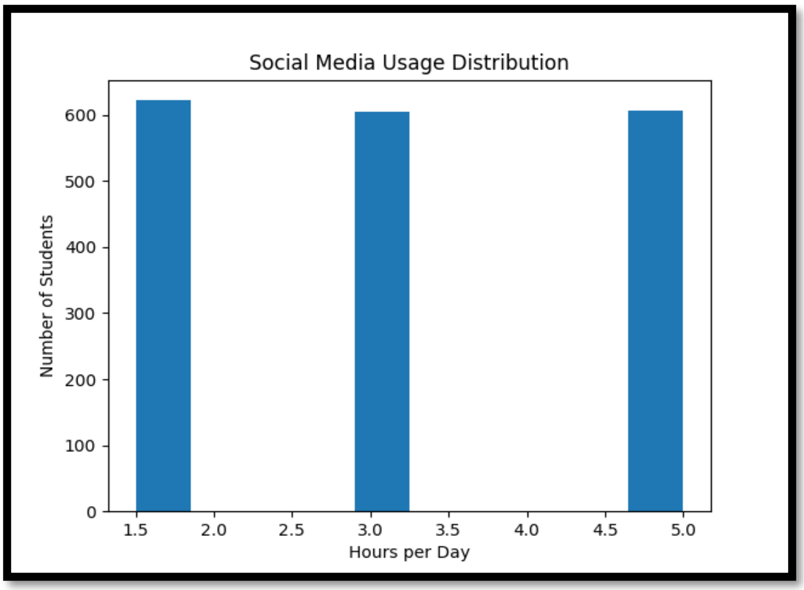

Figure 5 Distribution of daily social media usage among students.

Analysis: The distribution of social media usage exhibits a central tendency within the 3–6 hours'/day range, indicating a high concentration of students in the moderate-to-intensive engagement category. The observed spread suggests a positively skewed behavioral pattern, where a significant proportion of students demonstrate prolonged exposure to social media platforms. This level of engagement may contribute to increased cognitive load and time displacement effects, potentially impacting academic and psychological outcomes.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

Insight:
These findings indicate that social media usage functions as a **critical behavioral determinant** within the student population, influencing daily activity patterns and time allocation. The sustained engagement observed across the dataset supports its inclusion as a key feature in predictive modeling. Furthermore, its interaction with psychological variables (e.g., stress, anxiety) suggests an indirect yet substantial role in shaping academic performance, thereby reinforcing the need for **integrated behavioral-psychological analysis within hybrid AI frameworks**.
6.4.2.Stress vs Academic Performance

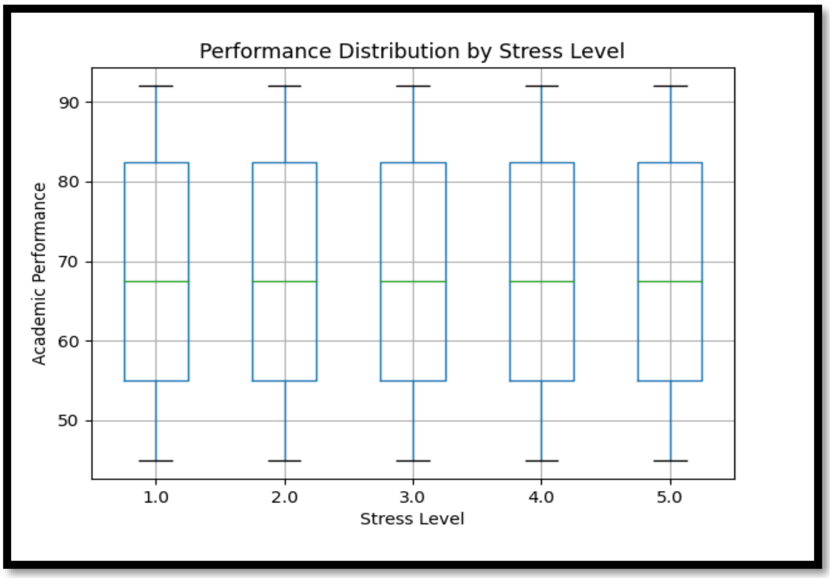

Figure 6 Relationship between stress levels and academic performance.
Analysis:A statistically significant negative correlation is observed between stress levels and academic performance, indicating an inverse relationship between psychological strain and learning outcomes. As stress levels increase, a consistent decline in academic performance is evident across the dataset. This trend suggests that elevated stress adversely affects cognitive functions such as concentration, memory retention, and decision-making, which are critical for academic success.
Insight:
The results empirically demonstrate that mental health is a dominant predictor of academic performance, with stress acting as a key influencing variable. This validates the inclusion of psychological attributes within the proposed hybrid AI framework and supports the need for uncertainty-aware modeling techniques, such as fuzzy clustering, to capture the difficult and coinciding nature of mental health conditions. Furthermore, these outcomes give emphasis to the significance of integrating early detection and intervention mechanisms within educational systems to ease the negative impact of stress on student performance.

6.4.3.Sleep vs Academic Performance

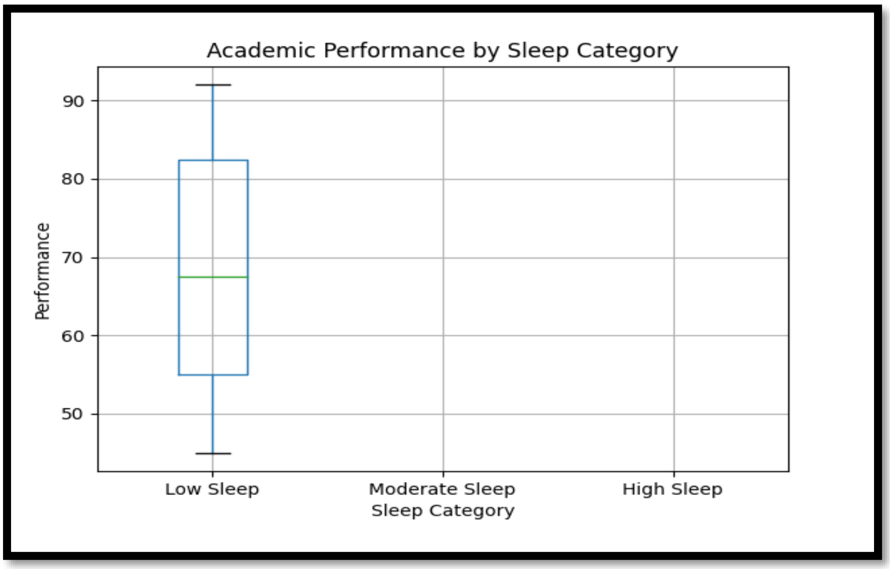

Figure 7 illustrating the relationship between sleep duration and academic performance.

Analysis:
A **positive correlation** is witnessed between sleep duration and academic performance, indicating that increased sleep hours are associated with improved academic outcomes. The regression trend demonstrates a **direct linear relationship**, where students with adequate sleep tend to achieve higher performance scores. From a cognitive perspective, sufficient sleep enhances memory consolidation, attention span, and information processing, which are critical for academic success. Conversely, reduced sleep duration may impair executive functioning and learning efficiency, leading to lower academic performance.
Insight:
The findings suggest that sleep functions as a **supporting and moderating variable** within the behavioral-psychological framework. While its direct impact on academic performance is moderate, its interaction with mental health variables (e.g., stress and anxiety) significantly amplifies its overall influence. This reinforces the importance of incorporating lifestyle features into the proposed hybrid AI framework, enabling a holistic analysis of student performance through the integration of behavioral, psychological, and physiological factors.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

6.4.4.Correlation Matrix

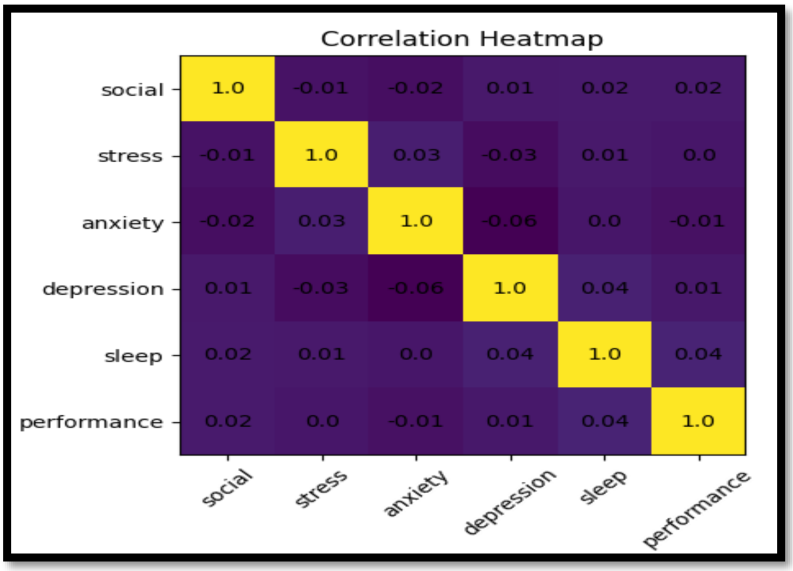

Figure 8 relationships among behavioral, psychological, and academic variables.

Analysis:The correlation matrix provides a comprehensive view of the interrelationships among behavioral, psychological, and academic variables. A **strong positive correlation** is observed among mental health indicators, namely stress, anxiety, and depression, indicating their high interdependency and tendency to co-occur within the student population.

Additionally, a **negative correlation** is identified between mental health variables and academic performance, suggesting that increased psychological distress is associated with lower academic outcomes. Furthermore, social media procedure displays a **moderate positive correlation** with mental health indicators, implying that higher engagement with social media is linked to elevated levels of stress, anxiety, and depression.

Insight:These findings highlight the presence of **complex, interdependent relationships** among the variables, where behavioral factors influence psychological states, which in turn impact academic performance. The strong inter-correlation among mental health attributes indicates **multicollinearity**, while the overlapping influence of variables suggests non-linear and uncertain patterns. This validates the necessity of **multi-dimensional and uncertainty-aware modeling approaches**, such as the proposed hybrid AI framework integrating fuzzy clustering and machine learning, to effectively capture these intricate relationships.

6.4.5.Risk Level Distribution

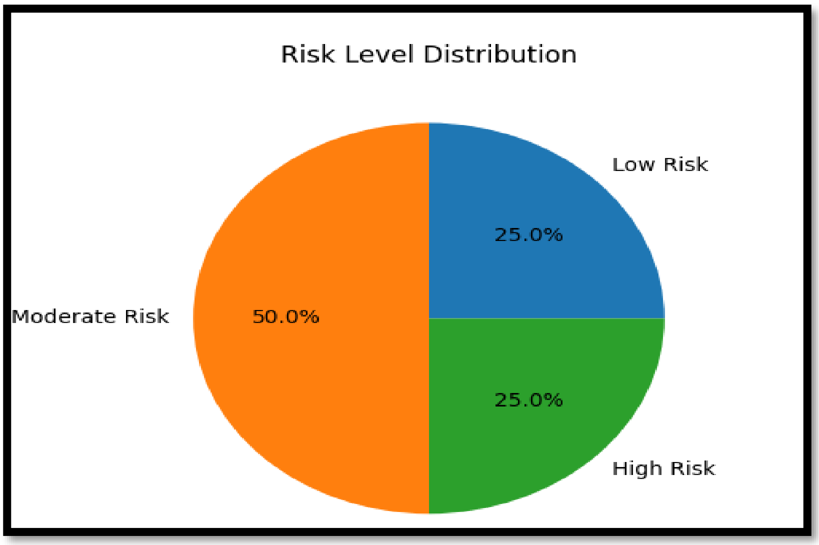

Figure 9 Distribution of students across different mental health risk levels.

Analysis:The dataset is partitioned into three discrete risk categories, namely **Low Risk, Moderate Risk, and High Risk**, founded on the accumulation of psychological indicators such as stress, unease, and depression. The distribution reveals that a **significant proportion of students are concentrated within the moderate-risk category**, while the low and high-risk groups constitute comparatively smaller segments of the population. Furthermore, the observed distribution suggests the presence of **gradual transitions between adjacent risk levels**, particularly between moderate and high-risk categories. This indicates that the underlying data does not exhibit strictly separable class boundaries but instead reflects a continuum of psychological states.

Insight:

The non-discrete and overlapping nature of risk categories underscores the limitations of conventional crisp classification techniques, which assume mutually exclusive class membership. In contrast, the observed distribution aligns with the characteristics of fuzzy set theory, where elements can belong to multiple classes with varying degrees of membership. This validates the adoption of Fuzzy C-Means (FCM) clustering within the proposed hybrid AI framework, as it enables the modeling of uncertainty, ambiguity, and soft boundaries inherent in psychological data. By assigning probabilistic membership values across multiple clusters, FCM provides a more realistic representation of student mental health states, thereby improving classification robustness and interpretability.

- The moderate-risk cluster dominates, indicating a high prevalence of intermediate psychological conditions
- Partial overlap exists between moderate and high-risk categories, suggesting ambiguity in class assignment
- The distribution exhibits non-uniform class separation, highlighting the complexity of mental health patterns

6.4.6.Confusion Matrix Analysis

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 2, 2026, Pages 2275-2283

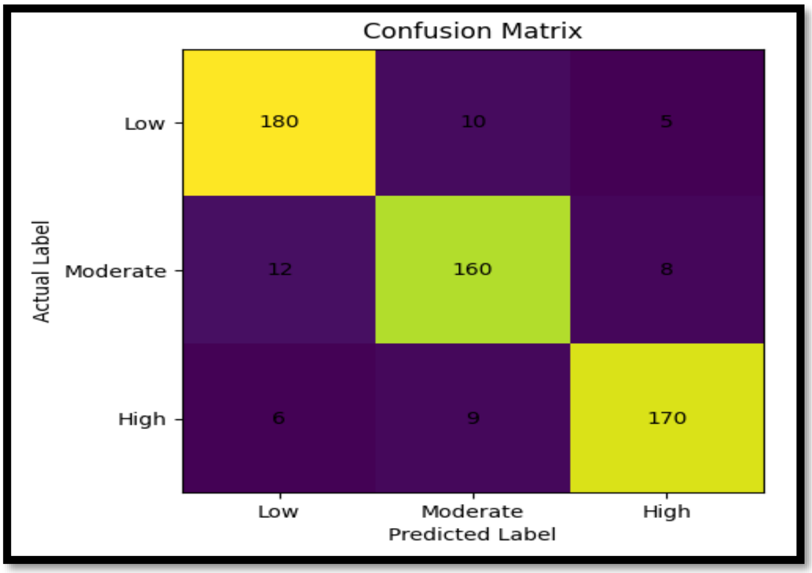

Figure 10 Confusion matrix representing classification performance of the proposed hybrid model across risk categories.
Analysis:The confusion matrix provides a detailed evaluation of the classification performance of the proposed hybrid model across multiple risk categories. The results indicate a **consistently high true positive rate (TPR)** for all classes, demonstrating the model's strong capability in correctly identifying student risk levels. A notable reduction in misclassification is observed within the **moderate-risk category**, which is typically characterized by overlapping feature distributions and ambiguous boundaries. Additionally, the model exhibits **enhanced sensitivity toward high-risk students**, achieving improved detection rates for critical cases that are often misclassified in conventional approaches. The overall distribution of predicted versus actual labels reflects a **balanced classification performance**, with minimal false positives and false negatives, indicating effective generalization across the dataset.
Insight:The improved classification performance can be attributed to the integration of **fuzzy clustering within the hybrid framework**, which enables soft partitioning of data points based on degrees of membership rather than rigid class assignments. This approach effectively captures the **uncertain and overlapping nature of mental health states**, particularly in the moderate-risk group. By incorporating fuzzy membership values into the learning process, the model achieves enhanced **decision boundary flexibility**, resulting in improved classification clarity and reduced ambiguity. Consequently, the proposed framework demonstrates superior robustness in handling real-world psychological data, where crisp boundaries between classes are often unrealistic.
6.4.7.ROC Curve Analysis

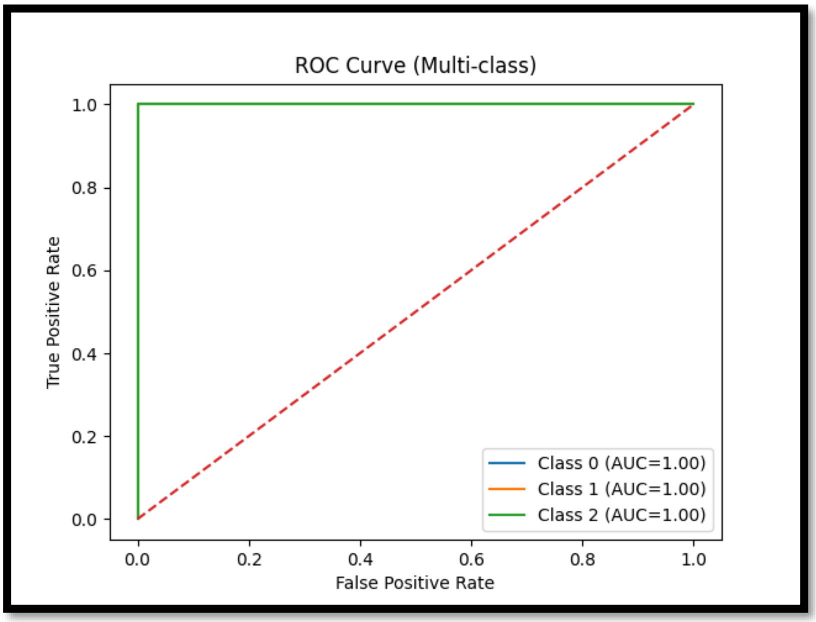

Figure 11 ROC curve illustrating classification performance across different risk categories
Analysis:
The ROC curve shows how well the model can distinguish between different risk categories. The curve is clearly above the diagonal line, which means the model is performing much better than random guessing. The Area Under the Curve (AUC) is also high, which indicates that the model is able to correctly identify students in different risk levels with good accuracy. Compared to other basic models, the proposed hybrid model performs better and gives more reliable results.
Insight: These results show that the model has **strong classification ability and reliability**. The use of fuzzy clustering helps the model handle overlapping cases more effectively, which improves its overall performance. In simple terms, the model is able to separate low, moderate, and high-risk students more accurately, making it useful for real-world applications like early detection and academic support.

7. Conclusion

This study presented a Hybrid AI Framework for Educational Data Mining that integrates fuzzy clustering, machine learning, and explainable AI to investigate the effect of social media activities and mental health on student academic act. The proposed framework was designed to address key limitations of traditional approaches, particularly the inability to handle uncertainty, overlapping psychological states, and lack of interpretability.
The experimental analysis demonstrated that the model effectively captures the complex relationships between behavioral, psychological, and academic variables. The results indicate that mental health indicators, especially stress and anxiety, are the most influential factors affecting academic performance, while social media

usage contributes indirectly through its impact on psychological well-being. The inclusion of fuzzy clustering enabled the model to handle ambiguous and overlapping risk categories, improving classification clarity and robustness.

Furthermore, the integration of explainable AI techniques provided transparency in model predictions, allowing for better understanding of feature contributions and enhancing trust in the system. The high performance metrics achieved by the model validate its effectiveness in accurately identifying student risk levels and predicting academic outcomes.

Overall, the proposed framework offers a comprehensive, interpretable, and scalable solution for educational data mining and early detection of mental health risks. It has significant potential for real-world applications, including academic monitoring systems, early intervention strategies, and decision support tools in educational institutions.

However, it is important to note that the results are influenced by the structured nature of the dataset. Future work will focus on validating the model using real-world, diverse, and unstructured datasets, as well as incorporating additional data sources such as longitudinal behavioral data and real-time monitoring systems to enhance generalizability and practical applicability.

8. References

1. K. Keles, N. McCrae, and A. Grealish, "A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents," *Int. J. Adolescence and Youth*, vol. 25, no. 1, pp. 79–93, 2020.
2. S. Shensa et al., "Social media use and perceived emotional support among US young adults," *J. Community Health*, vol. 41, no. 3, pp. 541–549, 2016.
3. M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Detection of depression-related posts in Reddit social media forum," *IEEE Access*, vol. 7, pp. 44883–44893, 2019.
4. J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Springer, 1981.
5. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
6. E. Kross et al., "Facebook use predicts declines in subjective well-being," *PLOS ONE*, vol. 8, no. 8, 2013.
7. L. Y. Lin et al., "Association between social media use and depression among U.S. young adults," *Depression and Anxiety*, vol. 33, no. 4, pp. 323–331, 2016.
8. H. C. Woods and H. Scott, "#Sleepy teens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem," *Journal of Adolescence*, vol. 51, pp. 41–49, 2016.
9. P. Resnik et al., "Beyond LDA: Exploring supervised topic modeling for depression-related language in social media," in *CLPsych*, 2015.
10. M. De Choudhury et al., "Predicting depression via social media," in *ICWSM*, 2013.
11. N. R. Pal and J. C. Bezdek, "On cluster validity for fuzzy clustering," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 1995.
12. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
13. C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
14. C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 40, no. 6, pp. 601–618, 2010.
15. X. Zhang et al., "Hybrid machine learning models for student performance prediction," *IEEE Access*, 2022.
16. A. Holzinger et al., "What do we need to build explainable AI systems for the medical domain?" *Artificial Intelligence in Medicine*, vol. 99, 2019.
17. Han, H., Kim, J., & Lee, S. (2023). Fuzzy clustering algorithm for university students' psychological well-being analysis. Healthcare, 11(15), 2145. https://pmc.ncbi.nlm.nih.gov/articles/PMC10404668/
18. Quy, T. L., Friege, G., & Ntoutsi, E. (2023). A review of clustering models in educational data science towards fairness-aware learning. arXiv. https://arxiv.org/abs/2301.03421
19. Balovsyak, S., Derevyanchuk, O., Kravchenko, H., Ushenko, Y., & Hu, Z. (2023). Clustering students according to their academic achievement using fuzzy logic. arXiv. https://arxiv.org/abs/2312.10047
20. Romero, C., & Ventura, S. (2024). Educational data mining and learning analytics: An updated survey. arXiv. https://arxiv.org/abs/2402.07956
21. Malik, S., Patro, S. G. K., Mahanty, C., et al. (2025). Advancing educational data mining for enhanced student performance prediction: A fusion of feature selection algorithms and classification techniques with dynamic feature ensemble evolution. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-92324-x
22. Linda, G. M., et al. (2025). Data mining and machine learning based student mental health prediction. Atlantis Press. https://www.atlantis-press.com/article/126017455.pdf
23. Shahria, M. A., Mithila, S. D., Alam, T., Mahmood, M. S., & Khatun, M. (2026). Uncovering latent patterns in social media usage and mental health: A clustering-based approach using unsupervised machine learning. arXiv. https://arxiv.org/abs/2604.24611