2025 IEEE 2nd International Conference on Computing, Applications and Systems (COMPAS) | 979-8-3315-5525-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/COMPAS67506.2025.11381639

Machine Learning-Based Prediction of Nightmares in Children: A Pediatric Psychological Integration

Lamia Islam
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
lamiaislamzy@gmail.com

Md. Basim Al Zabir Shammo
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
basim.210628@pust.ac.bd

Md. Rashed
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
rashedulislam.ice.pust@gmail.com

Md. Imran Hossain
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
imran05ice@pust.ac.bd

Md Akash Miya
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
mdakashmiya0827@gmail.com

Moumita Haque Mohona
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
moumitahaque28@gmail.com

***Abstract*—Identifying fear-related nightmares in children is particularly challenging, as young children often struggle to articulate their experiences or recall dream content. These nightmares, however, can severely disrupt sleep, hinder academic performance, and heighten psychological vulnerability, underscoring the need foraccurate and early prediction for effective intervention. This study explores various ML approaches, including LR, DT, SVM, KNN, NB, and LightGBM, to develop a robust predictive framework. The dataset undergoes comprehensive preprocessing, incorporating categorical encoding, numerical feature scaling, insufficient data handling, and the SMOTE to mitigate class imbalance. Each model is systematically trained and key performance indicators such as precision, recall, and F1 score were used for evaluation, with confusion matrices providing further insight into classification performance. Among the models evaluated, Light-GBM demonstrated the highest predictive performance, achieving an accuracy of 99.77%, a precision of 100%, a recall of 99.53%, and an F1 score of 99.77%, surpassing all other approaches. In this research, LightGBM is introduced as the proposed model, demonstrating superior predictive capability compared to alternative approaches and underscoring its effectiveness in predicting fear-related nightmares with high reliability. The study advances ongoing research efforts focused on machine learning applications in psychological assessment, offering a data-driven approach to early detection and facilitating timely mental health interventions.**

***Index Terms*—Fear-related nightmares in children, Early detection, Machine learning, LightGBM, SMOTE, Mental health assessment.**

I. Introduction

Nightmares are disturbing dreams that evoke intense fear or anxiety, often causing awakenings and disrupting overall sleep quality. Characteristically, they are intense images with powerful emotions, sometimes affecting well-being and even mental health in general [1]. The focus on children's nightmares is important. Research indicates that children experiencing frequent nightmares are at an increased risk for emotional and behavioral issues, including hyperactivity, attention deficits, and diminished academic performance. These recurrent distressing dreams can also lead to insomnia symptoms, such as difficulties in initiating and maintaining sleep [2]. Sleep-related difficulties, including delayed sleep onset, frequent night awakenings, sleepwalking, nightmares, and enuresis,were prevalent among primary school children and were associated with poor academic performance. Suppose a child's nighttime fears begin to interfere with the quality and quantity of their sleep or are accompanied by comorbid sleep or psychiatric disorders. In that case, these fears may have deleterious effects on the child's physical and mental health, academic performance, impulse control, risk-taking behavior, and social functioning [3]. A cross-sectional study involved 27 primary schools in Hannover, Germany, highlighting the significant impact of sleep disturbances on children's daytime functioning and academic success [4]. A study involving more than 6,000 primary school children found that those who frequently experienced nightmares faced an increased risk of developing insomnia, hyperactivity, mood disturbances, and lower academic performance [5]. Additionally, a study by researchers from the University of Tulsa also suggests that nightmare intervention among children can improve their school performance and emotional well-being [6]. These findings strongly affirm the urgent need for proactive identification and intervention in sleep disturbances, given their potential far-reaching consequences on students' mental health and academic performance. The early identification of nightmare-prone children may help prevent adverse outcomes like academic challenges and psychosocial maladjustments [7]. It can be challenging to determine if a child experiences frequent nightmares, especially when the child is too young to verbalize their anxieties or struggles to recall the dream's content [8]. Research suggests that approximately 70% of young children experience nightmares occasionally, with the frequency peaking around the age of ten [9]. Children aged 7 to 9 identified scary dreams as one of their top three sources of intense anxiety and concern [10].

Our study utilizes multiple ML models to predict fear-related nightmares in children based on behavioral and psychological data.Preprocessing of the dataset involves managing missing values, encoding categorical variables, and scaling numerical features. Six ML models, including LR, DT, LightGBM, SVM, KNN, and NB, are trained and evaluated. Model performance is evaluated through accuracy, precision, recall, F1-score, confusion matrices, and ROC curves, resulting in a comprehensive comparison across classifiers. Among the models investigated, LightGBM is proposed as the predictive model due to its superior performance across evaluation metrics. This selection highlights its effectiveness in handling complex feature interactions and imbalanced data.

The key contributions of this study are as follows:

1) This study pioneers the use of traditional machine learning models to predict fear-related nightmares in children, a task not yet addressed in state-of-the-art research.
2) It uniquely incorporates multifactorial data (sleep patterns, parental observations, and life changes) to create a context-sensitive prediction model.
3) Our proposed model achieved a high level of accuracy in predicting nightmares, thereby contributing to improved mental health assessments.

II. Problem Statement

Identifying frequent nightmares in children is not always an easy task, especially when they are too young to articulate their fears or have difficulty recalling the details of their dreams. Young children may struggle to find the words to describe their experiences or feel too frightened or confused to share what they remember. In some cases, they may wake up distressed without being able to explain why. This can make it challenging to determine whether a child's disrupted sleep is due to a bad dream or another underlying issue. Many parents may fail to detect frequent nightmares in children or may not identify them in time due to various factors, such as the child's inability to express their fears or the parent's limited awareness of sleep issues. This is where machine learning (ML) can play a transformative role. Machine learning has been applied to mental health diagnosis for decades, with increasing advancements in recent years. While many studies have focused on diagnosing anxiety, depression, and sleep disorders, the issue of fear-related nightmares in children remains underexplored. This study aims to address that gap by developing machine learning models to predict and detect nightmares, potentially improving early intervention and mental health support for children.

III. Literature Review

The integration of machine learning (ML) in mental health diagnosis dates back to the 1980s. One of the early systems, DTREE, employed decision tree techniques to diagnose mental health conditions based on DSM-IV Axis I disorders [11]. Similarly, INTERNIST/AUTOSCID was developed to conduct structured clinical interviews for diagnosing DSM-IV Axis II personality disorders [12]. In recent years, ML has gained prominence in diagnosing various mental health conditions, particularly among college students. A systematic literature review conducted between 2011 and 2024, based on PRISMA guidelines, selected 30 relevant studies to evaluate the application of deep learning models in diagnosing mental health disorders. Prominent algorithms, such as Convolutional Neural Networks (CNN), Random Forest (RF), Support Vector Machines (SVM), Deep Neural Networks, and Extreme Learning Machines (ELM), were found to show significant potential in detecting mental health conditions, with CNN, RF, and SVM achieving promising accuracy [13].

In predicting anxiety and depression, various ML techniques have been explored. For instance, Qasrawi *et al.* assessed the performance of SVM and RF models, which showed the highest accuracy rates for depression (SVM: 92.5%, RF: 76.4%) and anxiety (SVM: 92.4%, RF: 78.6%) prediction [14]. Another study demonstrated the successful detection of depression in children using RF, XGBoost, and Decision Tree, with RF achieving 95% accuracy and 99% precision [15]. Even though many studies explored psychology in the domain of child psychology, nightmares as a specific issue were not addressed. Sleep disorder detection research has largely focused on conditions like sleep apnea and insomnia, with AI advancements enhancing automated detection accuracy [16]. Despite these advancements, the issue of nightmares in children has been largely overlooked in ML-based studies.

Although sleep disorder detection research has focused on conditions such as sleep apnea and insomnia, the use of ML for detecting nightmares, particularly fear-related ones, remains underexplored. Nightmares, especially fear-based ones, may signify psychological distress, trauma, or anxiety, and early detection could provide valuable insights into a child's emotional state [17]. No existing ML models are dedicated to predicting nightmares, emphasizing a significant gap in the research. Current observational tools like polysomnography lack predictive automation [18]. Our research addresses this critical gap by focusing on the prediction and detection of fear-related nightmares in children using ML. While prior studies have demonstrated the effectiveness of CNN, RF, and SVM in diagnosing anxiety, depression, and sleep disorders, none have explored nightmare detection in children. Our study emphasizes that early detection is crucial for understanding underlying psychological distress, offering predictive automation that overcomes the limitations of current observational tools, such as polysomnography. This novel approach could significantly improve mental health interventions for children.

IV. Proposed Methodology

In order to construct an effective and dependable ML model for the prediction of fear-related nightmares in children, this study introduces a novel methodology that incorporates advanced ML techniques. The prediction of nightmares is a crucial task, as it offers early indicators of a child's mental and emotional well-being, potentially aiding in the detection of underlying psychological distress, such as trauma or anxiety. The proposed methodology is outlined in the following key steps:

A. Dataset Description

This dataset investigates the nightmares experienced by children aged 4 to 10 and their potential psychological effects, particularly in school settings. It includes features such as age, gender, nightmare frequency and content, sleep patterns, recent life changes, and parental observations. The primary target variable is "Psychological Problems at School," which measures the extent of school-related psychological issues. Derived variables such as "Has Recent Life Changes," "Irregular Sleep Pattern," and "Fear-Related Nightmare" are also included to explore additional patterns. It contains psychological and behavioral features that are indicative of nightmare patterns. It was sourced from Kaggle, contributed by Atif Masih under the "Kids Nightmare Patterns and Psychological Impact" repository [19].

B. Data Preprocessing

Effective data preprocessing is essential for ensuring accurate predictions in ML models. The preprocessing steps applied to the dataset are detailed below:

1) **Handling Missing Values:** Missing values in the dataset were addressed effectively to preserve information integrity and prevent data loss. For numerical features, the missing values in the *Age* and *Psychological Problems at School* Problems at School columns were imputed using the median. For categorical features, missing values in *Gender*, *Nightmare Frequency*, *Nightmare Content*, *Sleep Patterns*, *Recent Life Changes*, *Sleep Environment*, and *Parental Observations* were imputed using the mode of each respective column. This approach ensured that the dataset remained complete and reliable for analysis [20].
2) **Encoding Categorical Features:** Categorical features were transformed into a numerical format using Label Encoder, which assigns unique integer labels to each category in columns like *Gender*, *Nightmare Frequency*, *Nightmare Content*, and other similar features. This conversion allows models to process the categorical data effectively.
3) **Feature Scaling:** To standardize the numerical features, StandardScaler was used. This scaler transforms the *Age* and *Psychological Problems at School* columns to have a mean of 0 and a standard deviation of 1. Scaling ensures that all features are on the same scale, which is particularly beneficial for algorithms sensitive to feature magnitude, such as Support Vector Machines (SVM).
4) **Class Imbalance Handling:** The target variable, *Fear-Related Nightmare*, exhibited class imbalance. To address this, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. SMOTE generates synthetic samples for the minority class, ensuring a balanced dataset that prevents biased model predictions [21].
5) **Feature and Target Variable Definition:** The feature set (X) included all columns except the target variable *Fear-Related Nightmare*, while the target variable (y) contained labels indicating the presence or absence of fear-related nightmares.
6) **Data Splitting:** The pre-processed data was split into training and testing sets using an 80:20 ratio. Stratified Sampling ensured that the distribution of the target variable remained consistent across both sets.

Algorithm 1 LightGBM Training Procedure

Input: Training data $\{(x_i, y_i)\}_{i=1}^{N}$, number of trees M, learning rate η
Initialize model $F_0(x) = 0$
for $m = 1$ to M **do**
  Compute gradients $g_i = \frac{\partial l(y_i, \hat{y}_i)}{\partial \hat{y}_i}$ for all i
  Compute Hessians $h_i = \frac{\partial^2 l(y_i, \hat{y}_i)}{\partial \hat{y}_i^2}$ for all i
  Construct histogram-based feature splits
  Train decision tree $h_m(x)$ using leaf-wise growth strategy
  Update model: $F_m(x) = F_{m-1}(x) + \eta h_m(x)$
end for
Output: Final model $F_M(x)$

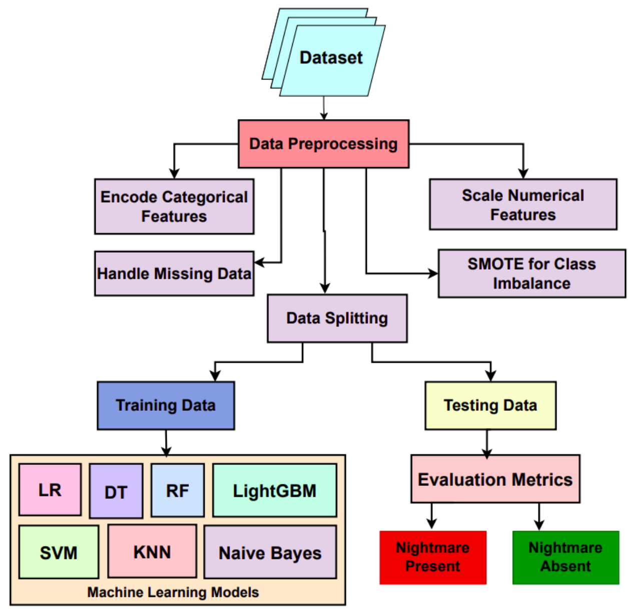

Fig. 1. Proposed Framework Architecture.

C. Proposed Framework

The model development framework in this code involves training multiple ML models to predict fear-related nightmares. A variety of models are employed, including LR, DT, LightGBM, SVM, KNN, and Naive Bayes. We initialized each model with specific hyperparameters to enhance performance. LightGBM, another ensemble learning model, is instantiated with the default parameters but is equipped with random state=42 to ensure reproducibility of results. The SVM is configured with probability=True, which enables the model to compute probability estimates for each class, a necessary step for generating ROC curves and calculating the Area Under the Curve (AUC). KNN is set with seven neighbors (n neighbors), which specifies the number of nearest data points considered when making a classification decision. This choice balances the trade-off between model sensitivity to noise and classification accuracy. Lastly, Naive Bayes uses the default configuration of the GaussianNB class, assuming a normal distribution for the continuous features. Once the models are trained, they are evaluated using accuracy, precision, recall, and F1 scores to assess overall prediction quality and class-specific performance. Confusion matrices are generated to provide a breakdown of true positives, true negatives, false positives, and false negatives, highlighting potential misclassifications and areas for model improvement. Additionally, ROC curves are plotted for all models, with the True Positive Rate (TPR) and False Positive Rate (FPR) calculated across different thresholds. The AUC is computed for each model to quantify its discriminatory power, and all ROC curves are visualized in a single graph to facilitate comprehensive model comparison. This detailed evaluation framework has led us to identify the most suitable model for predicting fear-related nightmares.

V. Results and Discussion

Table 1 compares the effectiveness of six machine learning algorithms in predicting fear-related nightmares. LR has modest performance (60.77% accuracy, 60.85% precision, 60.36% recall, and 60.60% F1-score). DT performs well, achieving 94.01% accuracy, 89.36% precision, 99.88% recall, and 94.35% F1-score. Light-GBM surpasses all models, achieving 99.77% accuracy, 100% precision, 99.53% recall, and 99.77% F1-score.
SVM performs moderately, with accuracy of 68.42%, precision of 65.65%, recall of 72.26%, and F1 score of 70.99%. KNN achieved 84.63% accuracy, 78.93% precision, 94.48% recall, and 86.01% F1-score. Naive Bayes performs well with an accuracy of 62.03%, precision of 61.40%, recall of 64.79%, and F1-score of 63.03%. LightGBM is the highest performance in all metrics.
Fig. 2 visually compares the accuracy of six machine learning algorithms in predicting fear-related nightmares. Light- GBM achieves the highest accuracy (99.77%), followed by Decision Tree (94.01%) and KNN (84.63%). SVM (68.42%) and Naive Bayes (62.03%) show moderate performance, while Logistic Regression achieves the lowest performance at 60.77%. This highlights LightGBM's superior predictive capability. Fig. 3 shows the confusion matrix that represents the performance of our LightGBM model. It shows that out of 2,573 actual No cases, the model correctly predicted all of them (true negatives) with 0 false positives. For actual Yes cases, the model correctly classified 2,561 instances but misclassified 12 as No (false negatives). This indicates a highly accurate model with minimal errors. The ROC curve compares the performance of multiple ML models in classifying fear-related nightmares, with their AUC scores reflecting overall predictive accuracy. LightGBM leads with a perfect AUC of 1.00, indicating flawless classification. Decision Tree follows closely with an AUC of 0.99, showcasing highly effective performance. K-Nearest Neighbors (KNN) has an AUC of 0.93, demonstrating strong predictive power. SVM and Naive Bayes show moderate performance with AUCs of 0.65, indicating a noticeable trade-off between true positives and false positives. Logistic regression has a similar AUC of 0.64, indicating modest classification ability. Overall, LightGBM stands out with a flawless F1-score of 99.77%, perfect precision, and recall.
Decision Tree also performs well with a 94.01% accuracy and 0.99 AUC. KNN shows strong recall but

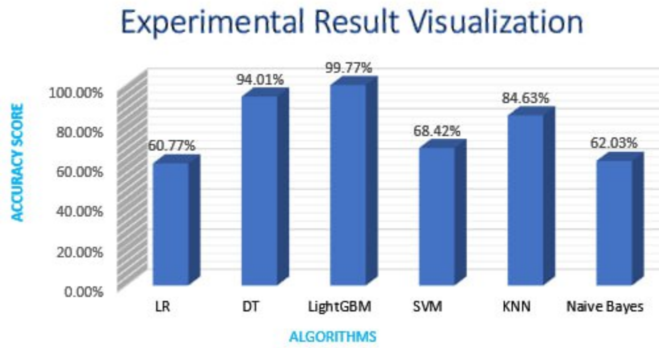

Fig. 2. Performance Visualization of Different Machine Learning Algorithms

TABLE I
PERFORMANCE OF DIFFERENT ML ALGORITHMS

Algorithms	Accuracy	Precision	Recall	F1-score
LR	60.77%	60.85%	60.36%	60.60%
DT	94.01%	89.36%	99.88%	94.35%
LightGBM	99.77%	100%	99.53%	99.77%
SVM	68.42%	65.65%	72.26%	70.99%
KNN	84.63%	78.93%	94.48%	86.01%
Naive Bayes	62.03%	61.40%	64.79%	63.03%

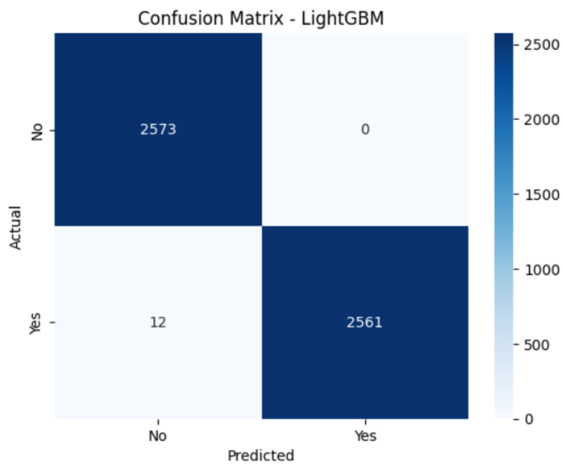

Fig. 3. Confusion Matrix for LightGBM Model.

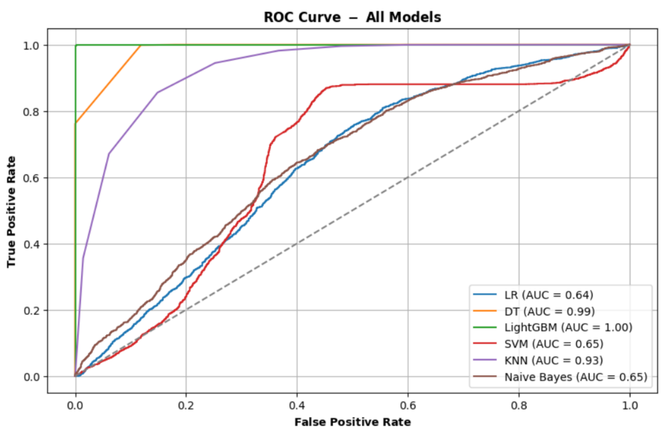

Fig. 4. ROC curve for classifier evaluation.

some false positives, while SVM and Naive Bayes underperform. These results highlight LightGBM as the most accurate and reliable model for this task. Table 2 highlights the differences between existing work and our proposed approach for detecting nightmares in children. Prior studies have largely concentrated on predicting mental health disorders, including depression and anxiety, achieving 92.5% accuracy and 95% using SVM and RF, respectively. However, neither of these studies addressed nightmares or their psychological impact. Our proposed model addresses this gap by detecting fear-related nightmares in children, a condition often linked to anxiety, trauma, and sleep disturbances. Using our proposed LightGBM model, we achieved a significantly higher accuracy of 99.77%, outperforming all other evaluated models. This demonstrates LightGBM's superior capability to handle complex, multifactorial data and its effectiveness in analyzing sleep-related patterns and psychological conditions in pediatric populations. It also highlights the potential of our study for early detection

TABLE II
COMPARISON OF THE PROPOSED MODEL WITH RELATED WORKS

Ref.	Work	Model	Accuracy
[14]	Risk factors for child depression	SVM	92.5%
[15]	Child depression detection	RF	95%
Our	**Our Nightmare detection**	**LightGBM**	**99.77%**

and timely intervention to improve children's mental and emotional well-being.

VI. Conclusion and Future Work

Our proposed LightGBM model, with an accuracy of 99.77patterns, parental observations, and life changes, it offers a holistic and context-sensitive approach to early nightmare detection.Early identification of Children prone to nightmares can help mitigate potential adverse outcomes like academic challenges, social withdrawal, and psychological distress, ultimately enhancing mental health interventions. Our model currently uses a single dataset, limiting its generalizability. Future Improvements should include using multiple datasets from diverse sources and integrating real-time data from wearable trackers

References

[1] A. Zadra and D. Donderi, "Nightmares and bad dreams: Their prevalence and relationship to well-being," *Journal of Abnormal Psychology*, vol. 109, no. 2, pp. 273–281, 2000, doi: 10.1037/0021-843X.109.2.273.

[2] N. Sahu, P. S. Patil, A. -, and I. Longkumer, "Understanding and treating nightmares: A comprehensive review of psychosocial strategies for adults and children," *Cureus*, vol. 16, no. 9, p. e70044, Sep. 2024, doi: 10.7759/cureus.70044. PMID: 39449955; PMCID: PMC11499308.

[3] A. Ivanenko and K. Larson, "24 nighttime distractions: Fears, nightmares, and parasomnias," in *The Oxford Handbook of Infant, Child, and Adolescent Sleep and Behavior*, 2013, pp. 347.

[4] S. Wiechers, A. Schlarb, M. Urschitz, et al., "Sleep problems and poor academic performance in primary school children," *Somnologie*, vol. 15, pp. 243–248, 2011, doi: 10.1007/s11818-011-0535-8.

[5] S. X. Li, M. W. Yu, S. P. Lam, et al., "Frequent nightmares in children: Familial aggregation and associations with parent-reported behavioral and mood problems," *Sleep*, vol. 34, no. 4, pp. 487–493, Apr. 2011, doi: 10.1093/sleep/34.4.487. PMID: 21461327; PMCID: PMC3065259.

[6] University of Tulsa, "University of Tulsa students and faculty study nightmares in children," *University of Tulsa News*, 2019. [Online]. Available: https://utulsa.edu/news/nightmares-in-children/. [Accessed: Dec. 20, 2024].

[7] G. Qu, H. Liu, T. Han, et al., "Association between adverse childhood experiences and sleep quality, emotional and behavioral problems, and academic achievement of children and adolescents," *Eur. Child Adolesc. Psychiatry*, vol. 33, no. 2, pp. 527–538, Feb. 2024, doi: 10.1007/s00787-023-02185-w. PMID: 36869931; PMCID: PMC9985439.

[8] G. Dewar, "Nightmares and night terrors in children: How to identify the problem, and help kids sleep more peacefully," *Parenting Science*, Jan. 16, 2024. [Online]. Available: https://parentingscience.com/night-terrors-in-children/.

[9] A. Gauchat, J. R. Séguin, and A. Zadra, "Prevalence and correlates of disturbed dreaming in children," *Pathologie Biologie*, vol. 62, no. 5, pp. 311–318, 2014.

[10] P. Muris, C. Meesters, H. Merckelbach, and P. Hülsenbeck, "Worry in children is related to perceived parental rearing and attachment," *Behav. Res. Ther.*, vol. 38, no. 5, pp. 487–497, May 2000, doi: 10.1016/s0005-7967(99)00072-8. PMID: 10816907.

[11] B. G. Buchanan and E. H. Shortliffe, *Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley Longman Publishing Co., Inc., 1984.

[12] R. A. Miller, H. E. Pople Jr., and J. D. Myers, "Internist-I, an experimental computer-based diagnostic consultant for general internal medicine," in *Computer-Assisted Medical Decision Making*. New York, NY: Springer, 1985, pp. 139–158.

[13] U. Madububambachu, A. Ukpebor, and U. Ihezue, "Machine learning techniques to predict mental health diagnoses: A systematic literature review," *Clin. Pract. Epidemiol. Ment. Health*, vol. 20, p. e17450179315688, 2024.

[14] R. Qasrawi, S. P. Vicuna Polo, D. Abu Al-Halawa, et al., "Assessment and prediction of depression and anxiety risk factors in schoolchildren: Machine learning techniques performance analysis," *JMIR Form. Res.*, vol. 6, no. 8, p. e32736, Aug. 2022, doi: 10.2196/32736. PMID: 35665695; PMCID: PMC9475423.

[15] U. M. Haque, E. Kabir, and R. Khanam, "Detection of child depression using machine learning methods," *PLoS ONE*, vol. 16, no. 12, p. e0261131, Dec. 2021, doi: 10.1371/journal.pone.0261131.

[16] S. Xu, O. Faust, S. Seoni, et al., "A review of automated sleep disorder detection," *Comput. Biol. Med.*, vol. 150, p. 106100, 2022.

[17] O. Lempen, "Exploring the role of children's dreams in psychoanalytic practice today," *Psychoanal. Study Child*, p. 228, 2007.

[18] J. Rolling, J. Rabot, E. Reynaud, et al., "Nightmares and sleep disturbances in children with PTSD: A polysomnographic and actigraphy approach evaluation," *J. Clin. Med.*, vol. 12, no. 20, p. 6570, Oct. 2023, doi: 10.3390/jcm12206570. PMID: 37892709; PMCID: PMC10607571.

[19] A. Masih, "Kids nightmare patterns and psychological impact," Kaggle, Aug. 24, 2024. [Online]. Available: https://www.kaggle.com/datasets/atifmasih/kids-nightmare-patterns-and-psychological-impact.

[20] S. K. Kwak and J. H. Kim, "Statistical data preparation: Management of missing values and outliers," *Korean J. Anesthesiol.*, vol. 70, no. 4, pp. 407–411, Aug. 2017, doi: 10.4097/kjae.2017.70.4.407. Epub 2017 Jul 27. PMID: 28794835; PMCID: PMC5548942.

[21] N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.