

An Explainable Machine Learning Framework Analyzing Mental Health Influencing Metrics from Screen Time Behavior

Sunaina Kumari¹, Dr. Anshu Bhasin²

¹ Computer Science & Engineering/ I.K. Gujral Punjab Technical University, India

² Computer Science & Engineering/ I.K. Gujral Punjab Technical University, India

Abstract - The growing reliance on digital devices has significantly influenced the behavioural patterns and mental well-being of children and adolescents. This study investigates the impact of screen-time-related factors on children's mental health using machine learning and explainable artificial intelligence techniques. Initially, Random Forest and CatBoost models are compared using evaluation metrics, including accuracy, precision, recall, and F1-score. The experimental results showed that the Random Forest model performed best, achieving an accuracy of 74.06%, and consequently chosen for further analysis.

To enhance model interpretability, Random Forest Feature Importance, SHAP, and Permutation Importance are employed to identify the most influential factors affecting mental health outcomes. The normalized scores obtained from these methods are integrated to propose a novel Mental Health Influence Score (MHIS), which provides a unified ranking of influential behavioural factors. The results revealed that average daily screen time is the most influential factor, followed by the educational to recreational ratio, age, and primary device usage.

Furthermore, LLM-based interpretation is incorporated to generate human-readable explanations of the identified factors, thereby improving the transparency and practical understanding of the proposed framework.

Key Words: Screen Time, Mental Health, Machine Learning, Large Language Model (LLMs), Mental Health Influence Score (MHIS).

1. INTRODUCTION

The increasing use of digital technologies has fundamentally transformed the daily lives of children and adolescents. Smartphones, tablets, laptops, and other internet-enabled devices have become integral to learning, communication, entertainment, and social interaction. While these technologies provide significant educational and social benefits, the rapid increase in screen time has raised substantial concerns regarding its potential impact on the psychological well-being of young individuals. Recent studies have reported that prolonged exposure to digital screens is associated with adverse mental health outcomes, including anxiety, depression, emotional distress, and reduced psychological well-being among children and adolescents.

Furthermore, increasing evidence suggests that excessive screen engagement may disrupt healthy behavioural patterns, including sleep quality, physical activity, and interpersonal interactions, thereby contributing to a higher risk of mental health problems. These findings highlight the growing need to better understand the behavioural factors associated with screen time and their influence on children's mental health [1], [3], [2].

Several studies have demonstrated a significant association between prolonged screen exposure and adverse mental health outcomes among children and adolescents. Excessive screen time has been linked to increased symptoms of anxiety, depression, psychological distress, and reduced emotional well-being. In addition to emotional disturbances, prolonged engagement with digital devices has been associated with sleep disruption, reduced physical activity, and diminished face-to-face social interactions, all of which contribute to poorer psychological health. Longitudinal evidence further indicates that increases in screen use over time are accompanied by corresponding increases in anxiety and depressive symptoms, suggesting that the duration and pattern of digital engagement are important determinants of mental health. Moreover, different forms of screen behaviour, including social media use, internet browsing, and video streaming, may exert varying levels of influence on psychological well-being, emphasizing the need for a comprehensive assessment of screen-related behavioural factors [1], [3], [5], [2].

The increasing availability of large-scale behavioural and demographic data has encouraged the adoption of machine learning (ML) techniques for mental health analysis. Unlike conventional statistical approaches, ML algorithms are capable of identifying complex and non-linear relationships among multiple behavioural factors, thereby improving prediction performance and supporting evidence-based decision-making. Recent studies have successfully employed machine learning models such as Random Forest, CatBoost, Support Vector Machine, and other ensemble learning techniques to predict mental health outcomes using behavioural, demographic, and digital usage characteristics [6], [7], [9]. Although these models achieve high predictive accuracy, their decision-making processes are often difficult to interpret because they operate as complex black-box models. This lack of transparency limits the understanding of

how individual behavioural factors contribute to prediction outcomes, thereby reducing the practical applicability of machine learning models in healthcare and mental health research [10],[12].

Recent advances in Explainable Artificial Intelligence (XAI) have significantly improved the transparency of machine learning models by providing techniques that quantify the contribution of individual features toward prediction outcomes. Among these methods, Random Forest Feature Importance, SHAP (SHapley Additive exPlanations), and Permutation Importance are well known for their ability to interpret feature relevance from different analytical perspectives. However, existing studies generally employ these explainability techniques independently, resulting in feature rankings that may vary depending on the selected interpretation method. Consequently, relying on a single explainability approach may provide only a partial understanding of the behavioural factors influencing children's mental health. Furthermore, very limited research has attempted to integrate multiple explainability techniques into a unified framework capable of producing a comprehensive and reliable assessment of feature influence. This drawback highlights the need for an explainable framework that combines complementary explainability methods to generate robust feature rankings and facilitate transparent interpretation of screen time-related mental health factors [10],[12].

To address the limitations, this study proposes an explainable machine learning framework for identifying the behavioural factors associated with children's mental health based on screen time behaviour. Initially, Random Forest and CatBoost models are comparatively evaluated to determine the most suitable predictive model. The selected model is subsequently interpreted using three complementary explainability techniques, namely Random Forest Feature Importance, SHAP (SHapley Additive exPlanations), and Permutation Importance, each providing a distinct perspective on feature relevance. To obtain a more comprehensive and reliable assessment of influential factors, the normalized outputs of these explainability techniques are integrated into a novel Mental Health Influence Score (MHIS). Furthermore, a Large Language Model (LLM) is employed to generate contextual and human-readable interpretations of the identified influential factors, thereby enhancing the transparency and practical interpretability of the proposed framework.

2. LITERATURE REVIEW

Karim et al. [1] conduct a systematic review of social media use and its relationship with mental health and report associations between excessive or problematic social media use and mental-health concerns, including anxiety, depression, and psychological distress. Similarly, Twenge and Campbell [3] examine screen time among children and adolescents and report that higher levels of screen use are associated with

lower psychological well-being, including lower self-control and emotional stability. Woods and Scott [5] investigate social media use among adolescents and report an association between social media use and poor sleep quality, anxiety, depression, and low self-esteem. Gao et al. [4] examine the relationship between social media exposure and mental-health problems during the COVID-19 outbreak and report that frequent social media exposure is positively associated with anxiety and combined depression and anxiety. However, the study focuses on adults in China during a specific public-health emergency, indicating that its findings should be interpreted within that context.

Pachava et al. [6] investigate the relationship between social media behaviour, demographic characteristics, and mental-health indicators among Indian youth. The study examines factors such as external validation, social media interactions, gender, age, and occupation in relation to mental-health outcomes and applies machine learning techniques to the analysis. The findings highlight the relevance of social media behaviour and demographic factors in understanding mental-health outcomes among Indian youth. Yang and Li [2] apply machine learning and natural language processing techniques to extract psychological and behavioural insights from social media users, demonstrating the potential of computational methods for analysing behavioural information from digital platforms. Xu et al. [8] investigate adolescent subjective well-being through machine learning and examine the role of mental-health and addictive-behaviour factors. Their findings identify several behavioural and psychological factors associated with adolescent subjective well-being, demonstrating the applicability of machine learning for analysing complex behavioural relationships.

Shahzadi et al. [7] investigate the impact of social media on students' academic performance using machine learning methods. The study evaluates K-nearest neighbors (KNN), support vector machine (SVM), and linear regression algorithms to analyse and predict academic performance in relation to social media use. The findings indicate an association between social media use and students' academic performance. Although the study does not directly focus on mental-health outcomes, it demonstrates the application of machine learning to examine behavioural patterns associated with students' digital-media use. This provides a relevant methodological basis for analysing behavioural factors related to screen use in the present study.

Lundberg and Lee [9] introduce SHAP (SHapley Additive exPlanations), a unified framework for interpreting machine learning predictions by assigning feature-importance values to individual predictions. Arrieta et al. [10] provide a comprehensive review of Explainable Artificial Intelligence (XAI) and discuss its concepts, taxonomies, applications, and challenges, highlighting the importance of model transparency and interpretability. Linardatos et al. [11] further review

machine-learning interpretability methods and discuss different approaches for understanding model predictions. These studies establish the importance of explainability when machine learning models are applied to complex and sensitive domains such as mental-health analysis.

Recent studies further demonstrate the importance of explainable machine learning for analysing digital behavioural patterns associated with mental health. Lamba et al. [13] develop an explainable machine learning framework for predicting depression risk from social media behaviour using supervised learning models. The study applies SHAP and LIME to identify influential behavioural factors, with daily screen time, passive scrolling, nighttime usage, and stress-driven engagement emerging as important predictors. The findings demonstrate that combining predictive modelling with explainability can improve the transparency and interpretation of mental-health risk assessment. However, the study primarily focuses on social media behaviour and uses self-reported data, indicating the need for further investigation using structured behavioural datasets.

Although previous studies demonstrate the effectiveness of machine learning in analysing mental-health outcomes, social-media behaviour, screen use, and related behavioural patterns, many studies primarily focus on predictive performance or individual analytical techniques. In addition, the reviewed studies do not provide a unified framework that combines multiple complementary feature-importance and explainability methods specifically for children's screen-time and mental-health analysis. Therefore, there is a need for an integrated approach that evaluates feature influence from multiple perspectives and provides a consolidated interpretation of the factors associated with children's mental-health outcomes. Motivated by this gap, the present study proposes an explainable framework that combines Random Forest Feature Importance, SHAP, and Permutation Importance to develop a unified Mental Health Influence Score (MHIS), followed by LLM-based interpretation to improve the understandability of the obtained results.

3. METHODOLOGY

3.1 Dataset Description

The dataset used in this study is obtained from the Kaggle platform and is stored as Indian Kids Screen Time. The dataset contains information of 9,712 children related to screen time behavior and associated health impacts among children and it includes both numerical and categorical features representing demographic attributes and digital usage patterns. The dataset consists of multiple attributes such as age, gender, average daily screen time, educational -to-recreational usage ratio, primary device used, and other behavioral indicators. Additionally, it includes textual feature Health impacts,

describing psychological conditions associated with screen usage.

3.2 Data Preprocessing

The dataset is preprocessed by converting textual values into a consistent format. Missing values are handled by imputing numerical features using the median strategy and categorical features using the most frequent values. Subsequently, categorical variables were encoded using the OneHot Encoder.

3.3 Model Development

Two ensemble learning algorithms, namely Random Forest and CatBoost, are employed to predict mental health outcomes. Random Forest is used because it combines multiple decision trees and can capture nonlinear relationships between the input features and mental health outcomes. It also provides an intrinsic feature-importance measure that can be used to identify influential variables.

CatBoost is included as a comparative boosting-based algorithm. It builds an ensemble of decision trees sequentially and provides an alternative tree-based learning approach for evaluating the predictive performance of the proposed analysis. Both models are trained using the same pre-processed dataset and are evaluated using accuracy, weighted precision, weighted recall, and weighted F1-score.

To enhance both models' prediction performance, hyperparameter adjustment is carried out. For further explainability investigation, the model with the superior overall classification performance is chosen. Random Forest is chosen as the final model since it outperforms CatBoost in the findings.

3.4 Random Forest Feature Importance

The selected Random Forest model is further determined to identify the relative importance of the input features. Random Forest Feature Importance is obtained from the trained model and represents the contribution of each feature within the tree-based ensemble. The resulting importance scores are used as the first measure of feature influence.

However, relying only on the built-in Random Forest importance measure may provide a limited perspective on feature influence. Therefore, two additional explainability techniques, SHAP and Permutation Importance, are employed to provide complementary evidence.

3.5 SHAP-Based Explainability

SHAP (SHapley Additive exPlanations) is employed to explain the contribution of individual features to the predictions generated by the selected Random Forest model. SHAP assigns a contribution value to each feature for each

observation, indicating how strongly the feature contributes to the model's prediction.

The Tree SHAP explainer is applied because the selected model is tree-based. SHAP values are obtained for the processed observations, and the values corresponding to the Bad mental health class are considered for the analysis. The absolute SHAP values are calculated and averaged across the observations to obtain an overall SHAP importance value for each feature. Thus, features with larger mean absolute SHAP values are considered to have greater influence on the model's predictions.

SHAP is selected because it provides a complementary perspective to the built-in Random Forest importance measure by explaining the contribution of features to individual predictions and their overall influence across the dataset.

3.7 Permutation Importance

Permutation Importance is additionally employed to evaluate feature relevance from a model-performance perspective. In this approach, the values of one feature are randomly shuffled while the remaining features are kept unchanged. The resulting change in model performance indicates the importance of that feature.

In the present study, weighted F1-score is used as the evaluation measure for permutation analysis. The shuffling procedure is repeated 10 times, and the mean importance across the repetitions is used as the final Permutation Importance value for each feature. A larger decrease in model performance after feature shuffling indicates greater dependence of the model on that feature.

Permutation Importance is selected because it provides a different perspective from both Random Forest Feature Importance and SHAP. While Random Forest Feature Importance reflects the importance obtained from the tree ensemble and SHAP measures feature contribution to predictions, Permutation Importance evaluates the effect of removing a feature's information on the model's predictive performance.

3.8 Proposed Mental Health Influence Score (MHIS)

$$MHIS = \frac{RF_{norm} + SHAP_{norm} + PI_{norm}}{3} \quad (1)$$

The proposed Mental Health Influence Score (MHIS) integrates the normalized scores obtained from Random Forest Feature Importance, SHAP, and Permutation Importance. The final score provides a unified ranking of behavioural factors according to their influence on children's mental health outcomes.

Since the three feature-importance approaches produce scores on different numerical scales, the obtained Random Forest,

SHAP, and Permutation Importance scores are independently normalized to a common 0–1 scale using Min-Max normalization. The normalized scores are then combined feature-wise to obtain the proposed Mental Health Influence Score (MHIS). For each feature, MHIS is calculated as the average of its normalized Random Forest, SHAP, and Permutation Importance scores. A higher MHIS indicates greater relative influence of the feature on the model's predictions within the analyzed dataset.

3.9 LLM-Based Interpretation

To enhance the interpretability of the identified influential factors, a Large Language Model (LLM) is used to produce human-readable explanations derived from the Mental health influencing scale ranking. The LLM assessed the impact of each factor and offered contextual insights into how screen-time -related behaviors influence children's mental health. The analysis showed that average daily screen time has the most significant impact on mental health outcomes, with the educational-to-recreational ratio, age, and primary device usage following in order of influence. Incorporating explanations from large language models enhances the transparency and practical comprehension of the proposed framework.

4. RESULTS

The experimental results are presented to evaluate the performance of the developed machine learning models and to identify the factors that contribute most to the prediction of children's mental health outcomes. First, the performance of the Random Forest and CatBoost models is compared using accuracy, precision, recall, and F1-score. The better-performing model is then selected for further feature-importance and explainability analysis. Subsequently, Random Forest Feature Importance, SHAP, and Permutation Importance are used to examine the influence of the input features. The normalized importance scores are further combined to obtain the proposed Mental Health Influence Score (MHIS), providing a unified ranking of the influential factors.

Table no 1 illustrates the performance comparison between the Random Forest and CatBoost models using four evaluation metrics: Accuracy, Precision, Recall, and F1-score. The Random Forest model achieves an accuracy of 74.06%, outperforming CatBoost (72.72%). Similarly, Random Forest achieves higher precision (0.8134), recall (0.7406), and F1-score (0.7114) than CatBoost, which records values of 0.7733, 0.7272, and 0.7018, respectively. These results show that the Random Forest model provides more reliable and consistent classification performance. Therefore, Random Forest is selected as the final prediction model for subsequent explainability analysis, including Feature Importance, SHAP,

Permutation Importance, and the proposed Mental Health Influence Score (MHIS).

Table no 1 Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-score
Random Forest	0.7406	0.8134	0.7406	0.7114
CatBoost	0.7272	0.7733	0.7272	0.7018

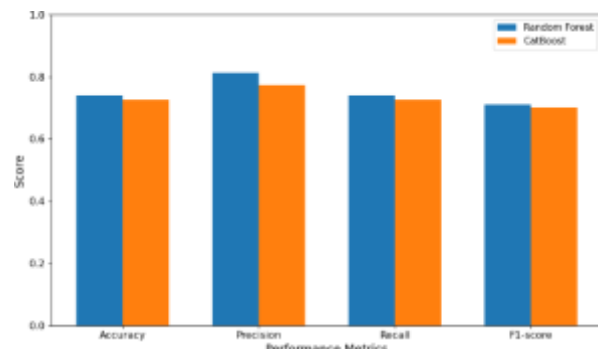


Figure 1: Performance comparison of Random Forest and CatBoost models

Figure 1 further illustrates the performance difference between the two models. Random Forest shows higher values for accuracy, precision, recall, and F1-score than CatBoost, indicating better overall classification performance on the given dataset. Based on these results, Random Forest is selected for the subsequent feature-importance and explainability analysis.

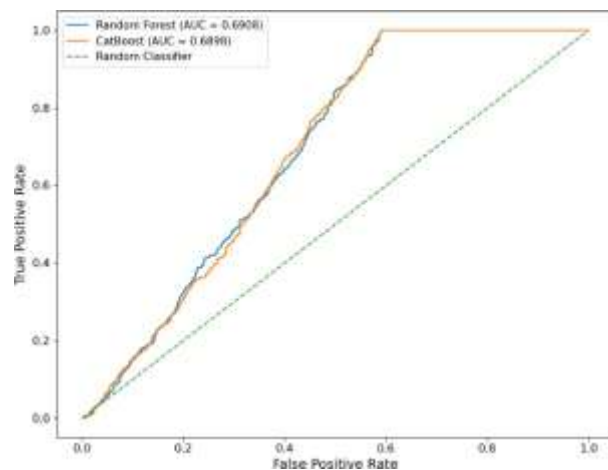


Figure 2: ROC curve (Random Forest vs CatBoost)

Figure 2 Presents the Receiver Operating Characteristic (ROC) curves of the Random Forest and CatBoost models for mental health prediction. The Random Forest model achieved an AUC value of **0.691**, while CatBoost acquired an AUC value of **0.690**. The ROC curves of both models remain above the random-classifier reference line, indicating that both

models provide discriminatory capability beyond random classification. Random Forest achieves a slightly higher AUC than CatBoost, which is consistent with its comparatively better overall performance reported in Table 2. However, the AUC values also indicate that the models provide moderate class-discrimination capability rather than strong separation between the two outcome classes. Therefore, the ROC analysis further supports the selection of Random Forest as the primary model for subsequent explainability analysis.

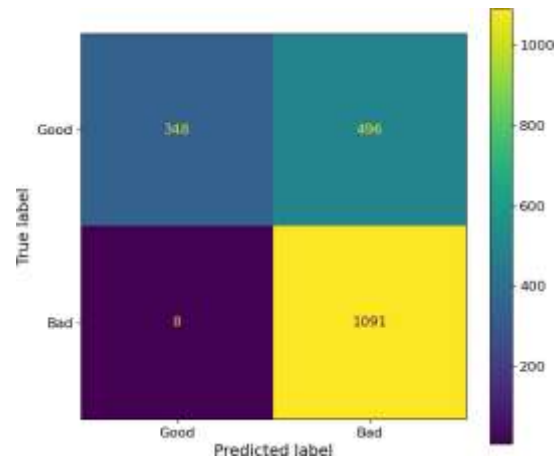


Figure 3: Confusion Matrix (Random Forest)

Figure 3 shows the confusion matrix of the Random Forest model for mental health classification. The model correctly classified 348 instances belonging to the Good mental health category and 1091 instances belonging to the Bad Category. However, 496 Good instances were incorrectly classified as Bad, whereas only 8 Bad instances are misclassified as Good. These results indicate that the Random Forest model is highly effective in identifying children with poor mental health outcomes while producing very few false-negative predictions. The confusion matrix further supports the selection of Random Forest for the subsequent explainability analysis and MHIS computation due to its superior classification performance.

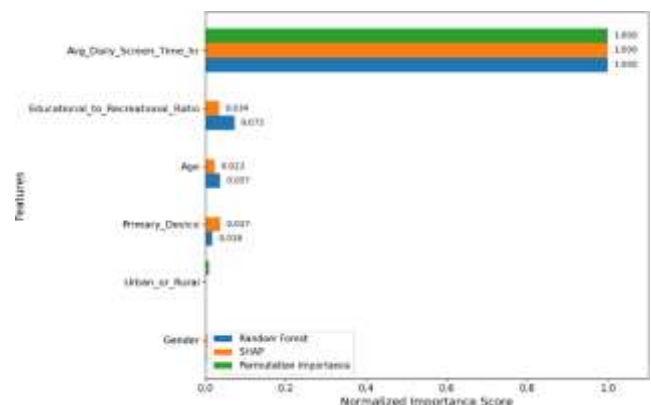


Figure 4. Comparison of Explainability Techniques and Proposed MHIS

Figure 4 presents a comparison of the normalized feature-importance scores obtained from Random Forest Feature Importance, SHAP, and Permutation Importance. The scores are normalized to a common 0–1 scale to enable comparison across the three explainability methods and to provide a consistent basis for calculating the proposed Mental Health Influence Score (MHIS). Average daily screen time obtains the maximum normalized score of 1.000 across all three methods, indicating consistent agreement regarding its dominant influence on the model predictions. The educational-to-recreational ratio obtains normalized scores of approximately 0.072 using Random Forest Feature Importance and 0.034 using SHAP, while its Permutation Importance remains close to zero. Similarly, age obtains normalized scores of approximately 0.037 and 0.023 using Random Forest Feature Importance and SHAP, respectively, with a very small contribution from Permutation Importance. Primary device shows a normalized SHAP score of approximately 0.037 and a Random Forest Feature Importance score of approximately 0.018. Urban or rural status shows a small contribution, particularly under Permutation Importance, whereas gender demonstrates comparatively negligible influence.

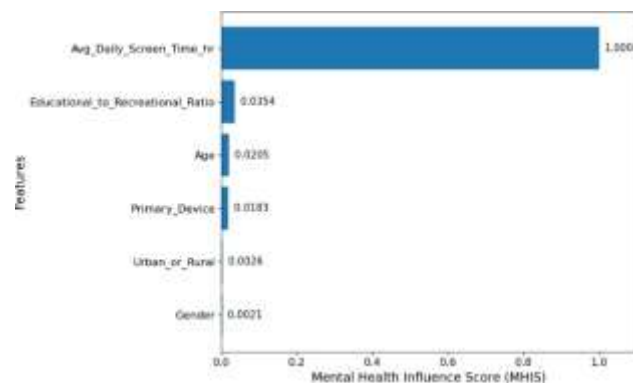


Figure 5: Mental Health Influence Score

Figure 5 displays the final ranking of influential factors based on proposed Mental Health Influence Score (MHIS) is calculated by combining the normalized Random Forest Feature Importance, SHAP importance, and Permutation Importance scores. As shown in Fig. 4, average daily screen time obtains the highest MHIS value of 1.0000, indicating that it is the most influential feature among the examined variables. The educational-to-recreational screen-time ratio follows with an MHIS of 0.0354, while age and primary device obtain scores of 0.0205 and 0.0183, respectively. Urban or rural status and gender show considerably lower influence, with scores of 0.0026 and 0.0021, respectively.

The substantially higher MHIS value of average daily screen time indicates that total daily screen exposure contributes more strongly to the model's predictions than the other examined variables. The educational-to-recreational ratio also shows a comparatively higher influence than demographic

characteristics, suggesting that the purpose and distribution of screen use provide additional information for the prediction task. Age and primary device demonstrate smaller but measurable contributions, whereas urban or rural status and gender have relatively limited influence. These results indicate that screen-use behaviour is more influential than the examined demographic factors within the present dataset

5. DISCUSSION

The Results of this study shows that the proposed framework provides an interpretable analysis of screen-time-related factors associated with mental health outcomes. Random Forest and CatBoost are initially compared as predictive models, with Random Forest achieving better overall classification performance and therefore being selected for further analysis. To strengthen the interpretation of the selected model, Random Forest Feature Importance, SHAP, and Permutation Importance are subsequently employed. These methods provide complementary perspectives on feature influence: Random Forest Feature Importance reflects the importance obtained from

the trained model, SHAP explains the contribution of features to model predictions, and Permutation Importance evaluates the effect of disrupting a feature on model performance. The agreement among these approaches provides stronger support for the findings obtained from the initial Random Forest analysis. Their normalized importance scores are therefore combined to obtain the proposed Mental Health Influence Score (MHIS), providing a unified ranking of the factors influencing the model's predictions.

The MHIS results identify average daily screen time as the most influential factor, followed by the educational-to-recreational screen-time ratio, age, and primary device. The original Random Forest analysis's conclusion is reinforced by the average daily screen time's persistent dominance across the three importance metrics, which suggests that overall daily screen exposure is especially significant within the current dataset. The goal shouldn't always be to eradicate educational screen use because youngsters use screens for both recreational and educational purposes. Instead, a good balance between digital and offline activities should be maintained in order to manage excessive or protracted overall exposure. The educational-to-recreational ratio's comparatively high rating further implies that screen usage type and purpose should be taken into account in addition to overall screen time. While gender and urban/rural status have relatively less predictive significance, age and primary device also contribute to the model.

By integrating prediction, explainability, feature ranking, and LLM-based interpretation, the suggested paradigm expands the study beyond straightforward model comparison. While MHIS presents a combined ranking of the aspects under

investigation, Random Forest Feature Importance, SHAP, and Permutation Importance offer complementary viewpoints on feature influence. The numerical ranking is further transformed into a more comprehensible form by the LLM-based interpretation. Pragmatically, the results indicate that screen-time management should concentrate on minimizing needless or extended exposure while preserving a suitable balance between offline, recreational, and instructional activities. The suggested framework should be viewed as an analytical and decision-support method rather than a clinical diagnostic system.

6. CONCLUSIONS

This study explores the impact of screen-time-related behavioural factors on children's mental health using machine learning and explainable artificial intelligence techniques. Initially, Random Forest and CatBoost models are compared using multiple evaluation metrics, including accuracy, precision, recall, and F1-score. The experimental results show that the Random Forest model achieves superior performance and is therefore selected for further explainability analysis.

To improve the interpretability of the prediction model, three explainability methods Random Forest Feature Importance, SHAP, and Permutation Importance are used. The normalized scores from these methods are combined to introduce the Mental Health Influence Score (MHIS), offering a consistent and dependable ranking of factors that impact children's mental health. The analysis finds that average daily screen time is the most impactful factor, with the educational-to-recreational ratio, age, and primary device usage ranking next underscoring the critical role of behavioral factors in mental health evaluation.

The proposed MHIS framework enhances the transparency and reliability of feature importance analysis by integrating several explainability techniques rather than relying on a one method. Furthermore, the integration of LLM-based interpretation improves the practical usability of the framework by converting numerical importance scores into human-readable explanations. This enables researchers, educators, parents, and healthcare professionals to better understand the relationship between screen-time behaviours and mental health outcomes, thereby supporting more informed decision-making.

In future work, the proposed framework can be extended by incorporating larger and more diverse datasets, advanced deep-learning models, and real-time mental health monitoring systems. Additional behavioural, environmental, and psychological factors may also be combined to further improve prediction performance and interpretability. Moreover, the proposed MHIS framework can be validated across different age groups and populations to evaluate its

generalizability and applicability in broader mental health assessment scenarios.

REFERENCES

1. F. Karim, A. Oyewande, L. F. Abdalla, R. C. Ehsanullah, and S. Khan, "Social Media Use and Its Connection to Mental Health: A Systematic Review," *Cureus*, vol. 12, no. 6, Art. no. e8627, 2020, doi: 10.7759/cureus.8627.
2. X. Yang and G. Li, "Psychological and Behavioural Insights From Social Media Users: Natural Language Processing-Based Study," *JMIR Formative Research*, 2025.
3. J. M. Twenge and W. K. Campbell, "Associations Between Screen Time and Lower Psychological Well-Being Among Children and Adolescents: Evidence From a Population-Based Study," *Preventive Medicine Reports*, vol. 12, pp. 271–283, 2018, doi: 10.1016/j.pmedr.2018.10.003.
4. J. Gao et al., "Mental Health Problems and Social Media Exposure During COVID-19 Outbreak," *PLOS ONE*, vol. 15, no. 4, Art. no. e0231924, 2020, doi: 10.1371/journal.pone.0231924.
5. H. C. Woods and H. Scott, "#Sleepyteens: 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.
6. Pachava, O. A. Lasekan, S. K. Golla, and S. Gosikonda, "Machine Learning Analysis of Social Media's Impact on Mental Health of Indian Youth," *International Research Journal of Multidisciplinary Scope*, vol. 5, no. 2, pp. 623–635, 2024, doi: 10.47857/irjms.2024.v05i02.0592.
7. K. Shahzadi, M. S. Sarfraz, M. Saqib, and H. M. Shah, "Machine Learning Methods for Analysing the Impact of Social Media on Students' Academic Performance," *Journal of Applied Artificial Intelligence*, vol. 1, no. 1, pp. 13–22, 2020, doi: 10.48185/jaai.v1i1.96.
8. Y. Xu, L. Xie, M. Bao, X. Yang, S. Chen, and Z. Gao, "Exploring Subjective Well-Being in Adolescents: The Role of Mental Health and Addictive Behaviours Through Machine Learning," *International Journal of Mental Health Promotion*, vol. 27, no. 5, pp. 667–682, 2025, doi: 10.32604/ijmhp.2025.062808.
9. S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information A.*
10. Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI," *Information Fusion*, vol. 58, pp. 82–115, 2020, doi: 10.1016/j.inffus.2019.12.012.
11. G. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable AI: A Review of Machine Learning

Interpretability Methods,” Entropy, vol. 23, no. 1, Art. no. 18, 2021, doi: 10.3390/e23010018.

12. S. S. Kar, R. Dube, B. K. M. Goud, Q. S. Gibrata, A. A. El-Balbissi, T. A. Al Salim, and R. N. M. A. K. Fatayerji, “Impact of Screen Time on Development of Children,” Children, vol. 12, no. 10, Art. no. 1297, 2025, doi: 10.3390/children12101297.

13. K. Lamba, S. Rani, and M. Shabaz, “Explainable machine learning for mental health prediction from social media behavior: A nested cross-validation study with SHAP and LIME interpretability,” Discover Mental Health, vol. 6, Art. no. 25, 2026, doi: 10.1007/s44192-026-00373-z.