2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) | 979-8-3503-7024-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCCNT61001.2024.10724267

Analyzing the Impact of Social Media Usage on Mental Health: A Machine Learning Approach

Shardha Purohit
School of Journalism & Mass Communication
Noida International University
Gautam Budh Nagar, India
shradhatpurohit20@gmail.com

Rusha Mudgal
Amity School of Communication
Amity University
Gurugram, Manesar, India
mudgal88@outlook.com

Satvik Vats
Department of Computer Science & Engineering
Graphic Era Hill University
Dehradun, India
svats@gehu.ac.in

Pooja Rana
Amity School of Communication
Amity University
Gurugram, Manesar, India
prana1@ggn.amity.edu

Owais Ahmad Shah
Department of Electronics and Communication Engineering
Dayanand Sagar University
Harohalli, India
mail_owais@yahoo.in

Aditya Verma
Department of Computer Science & Engineering
Graphic Era Deemed to be University
Dehradun, India
adityaverma.gehu@gmail.com

***Abstract*—This study investigates the correlation between social media usage and its impact on mental health, utilizing a dataset of 481 individuals to explore how varying degrees of social media engagement affect symptoms of ADHD, Anxiety, Depression, and Self Esteem. Through the application of statistical analysis and machine learning models, significant findings indicate that increased time spent on social media correlates positively with heightened mental health issues, particularly among younger demographics. The Pearson correlation matrix and machine learning algorithms, including Logistic Regression and Gaussian Naive Bayes, were utilized to analyze the data, achieving predictive accuracies of 99.31% and 94.44%, respectively. These models suggest the potential for using machine learning to identify and mitigate risks associated with excessive social media use. The study highlights the need for targeted mental health interventions tailored to specific demographic groups to alleviate the negative impacts of social media. This research underscores the importance of developing strategies to manage and prevent mental health issues in the digital age, offering valuable insights for policymakers, educators, and healthcare providers.**

Keywords— Social media, Mental health, Machine learning, Predictive modelling, Demographic analysis

I. Introduction

The profound integration of social media into the fabric of daily life, described as transforming global communication and embedding itself deeply into the lives of billions, sets the stage for significant inquiries into its psychological impacts.

In recent years, the relationship between social media usage and mental health has become a pivotal area of research, particularly as digital communication platforms grow increasingly prevalent. Studies are delving into various aspects of this relationship, uncovering both detrimental and potentially positive impacts. [1] conducted a study on the influence of social media intensity on mental health among middle school students in Indonesia, revealing a negative correlation between mental health and the intensity of social media use. This suggests that high usage can disrupt mental well-being, emphasizing the need for awareness and moderation in social media consumption among young populations. Similarly, [2] explored the impact of social networks on mental health from a risk factor perspective, investigating the pathological associations that might arise from prolonged use of such platforms. Although the specifics of their findings are not detailed, the title suggests a comprehensive analysis of how social networks could exacerbate mental health issues. [3] offers an understanding of the connection between social media use and mental health problems, likely examining the mechanisms through which social media impacts psychological well-being, although specific results are not outlined in the abstract. This research could be crucial for developing interventions to mitigate negative effects. Finally, [4] discusses the impact of social media on human flourishing and mental health from a personalist perspective, exploring how personal values and ethical considerations interact with social media use, impacting overall mental health and personal. These studies contribute to a nuanced understanding of how social media influences mental health, revealing a complex interplay of factors that could potentially harm or enhance individual well-being.

The study, leveraging a Kaggle dataset derived from surveys on social media usage, aligns with findings such as those by [3], who investigated the mechanisms linking social media use to mental health issues. Similar to the analysis by [3], current approach using machine learning techniques to identify patterns in social media behaviour that correlate with mental health risks is innovative and vital for developing preventive measures.

Moreover, the development of a predictive model relates closely to the proactive measures discussed by [1], where the focus on risk factors and pathological associations from social media use is examined. Several studies have investigated the impact of social media usage on mental health [5]-[9]. The study model, aimed at early screening for psychological distress triggered by social media habits, offers a practical application that could guide both users and policymakers in fostering a healthier digital environment. The increasing complexity of VLSI design necessitates an understanding of how social media usage impacts mental health, as engineers strive to create technologies that enhance user well-being [10], [11].

II. Dataset Description

The dataset employed in this study comprises a structured survey data collection aimed at exploring the association between social media usage and mental health outcomes. It encompasses a sample size of 481 participants, each represented across 21 distinct variables (see Table 1). These variables not only capture demographic information such as age, gender, relationship status, occupation status, and affiliations with organizations but also detail the variety of social media platforms used by the respondents and the extent of their engagement, quantified in hours spent on these platforms. The frequency distribution histogram plot of all the variables in the dataset is presented in Fig 1. Furthermore, the pair plot between the variables are depicted in Fig. 2.

A significant aspect of the dataset is the use of a series of Likert scale questions designed to measure the frequency or intensity of mental health symptoms that may correlate with social media usage. These questions are carefully crafted to assess various psychological traits and conditions:

- **ADHD-related Symptoms**: Questions address the purposeless use of social media, the ease with which one is distracted by social media, and difficulty in concentrating. These factors are indicative of potential attention deficit issues exacerbated by digital interactions.
- **Anxiety-related Symptoms**: The dataset includes measures such as restlessness when not using social media and being overly bothered by worries, which may signal underlying anxiety disorders.
- **Self-Esteem Issues**: It investigates how individuals compare themselves to their peers on social media, their feelings about these comparisons, and the extent to which they seek validation from their online interactions.
- **Depression-related Symptoms**: The survey probes feelings of depression, fluctuations in interest, and sleep issues, providing insights into the depressive states that may be influenced by social media habits.

TABLE I. Dataset statistical analysis.

Statistics	Age	ADHD Score	Anxiety Score	Self Esteem Score	Depression Score	Total Score	Outcome
Count	480	480	480	480	480	480	480
Mean	26.14375	13.46042	6.147917	6.266667	9.639583	35.51458	0.370833
Std	9.923621	3.898302	2.08709	2.759635	3.104528	9.274507	0.483532
Min	13	4	2	2	3	14	0
25%	21	11	5	4	7.75	29	0
50%	22	14	6	6	10	36	0
75%	26	16	8	8	12	42	1
Max	91	20	10	14	15	58	1

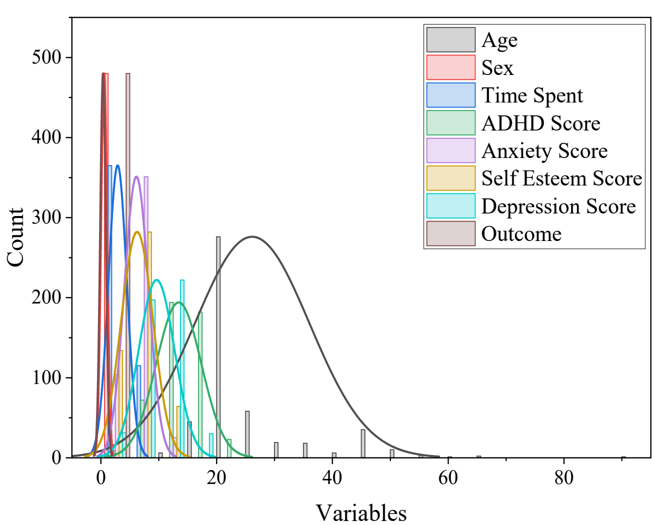

Fig. 1. Frequency distribution histogram plots of the features in the dataset.

Fig. 2. Pair plot of all the variables in the dataset.

The Pearson correlation matrix [12] depicted in Fig. 3 reveals several critical findings regarding the impact of social media on mental health. Importantly, there is a strong positive correlation between the time spent on social media and increased symptoms of ADHD, Anxiety, and Depression. This suggests that higher social media usage could exacerbate these mental health conditions. Additionally, the strong intercorrelations between ADHD, Anxiety, and Depression scores indicate that these conditions might influence each other, compounding the mental health challenges faced by individuals. The 'Outcome' variable, which is highly correlated with these mental health scores, likely reflects an overall mental health impact, underscoring the potential severity of these symptoms.

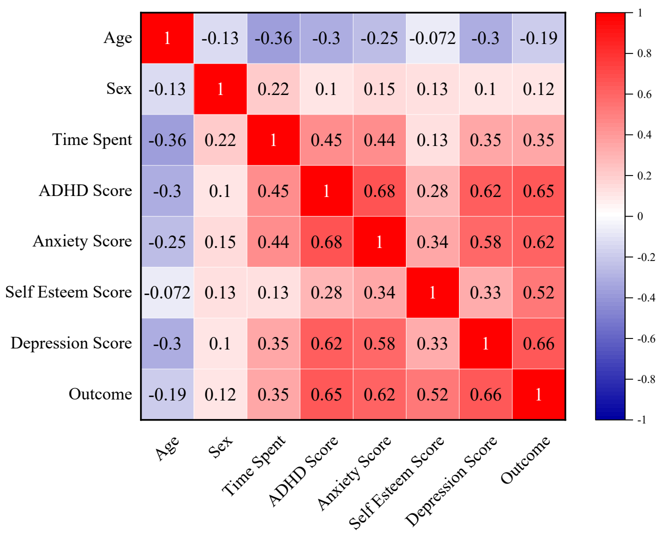

Fig. 3. Pearson correlation matrix.

III. Methodology

The methodology employed in this study was meticulously designed to analyze the impact of social media usage on mental health. Utilizing a dataset sourced from Kaggle, the research was conducted using Python within a Jupyter Notebook environment, leveraging powerful libraries such as Pandas for data manipulation and Seaborn along with Matplotlib for visualization purposes. The complete methodology flowchart is presented in Fig. 4.

The dataset consisted of responses from 481 individuals, covering 21 variables that included demographic details, social media usage, and responses to mental health symptom measurements via Likert scale questions. The initial phase of data preprocessing involved thorough cleaning steps, such as handling missing values and encoding categorical data for analytical readiness. This also included normalizing the data to ensure a consistent scale across different measures, facilitating more accurate comparisons and analyses.

A foundational aspect of the analysis was the construction of a Pearson correlation matrix. This matrix was crucial in identifying and visualizing the relationships between various variables, particularly focusing on how time spent on social media correlated with different mental health outcomes namely ADHD, anxiety, and depression symptoms.

Advancing from statistical insights, the study employed machine learning algorithms to develop predictive models aimed at estimating mental health outcomes based on patterns of social media usage. Selection of appropriate models was based on a rigorous evaluation of performance metrics such as accuracy, precision, and recall, all of which were derived from extensive cross-validation procedures to enhance the models' reliability and robustness.

To assess the efficacy of the predictive models, a hold-out validation set was used. This step was essential to evaluate the practical application of the models, ensuring they could accurately forecast mental health risks from given data inputs under real-world conditions.

The study adhered strictly to ethical standards concerning the handling and processing of data, with particular attention to maintaining the anonymity and confidentiality of the participants' information. The methodology was crafted to comply with all pertinent ethical guidelines related to research involving human subjects.

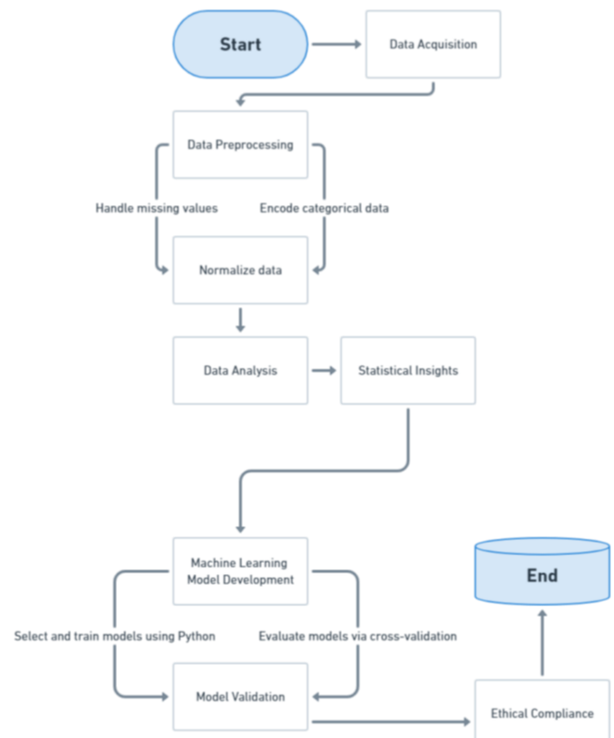

Fig. 4. Flowchart of the study with ML model approach.

IV. Results and Discussions

The analysis of the dataset, encompassing 480 participants, revealed significant correlations between social media usage and various mental health issues. Key findings include a notable positive correlation between time spent on social media and increased symptoms of ADHD (0.45), Anxiety (0.44), and

Depression (0.35), suggesting a detrimental impact on mental health with higher social media engagement.

Age displayed negative correlations with mental health scores, indicating younger individuals are particularly susceptible to these effects. The outcome variable, potentially representing overall mental health risk, showed strong correlations with ADHD (0.65) and Anxiety (0.62) scores.

Machine learning models were employed to predict mental health outcomes based on social media usage patterns. The Logistic Regression model demonstrated an impressive accuracy of 99.31%, while the Gaussian Naive Bayes model achieved 94.44% accuracy as can be seen in Figs. 5 & 6. These results highlight the potential of using predictive models to identify and mitigate risks associated with social media usage on mental health.

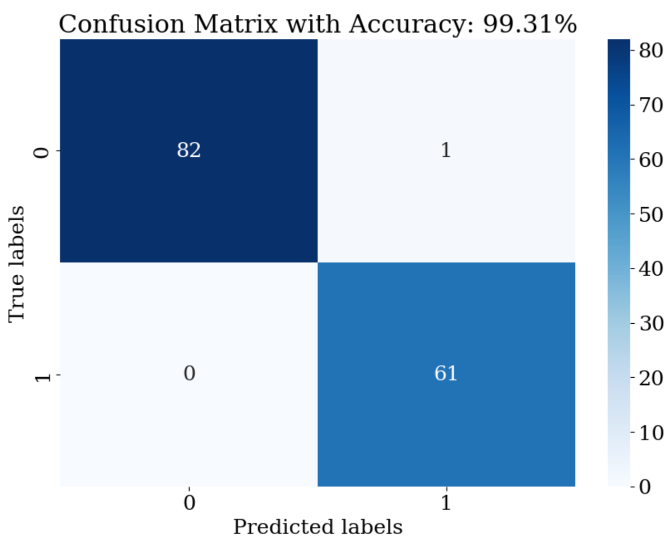

Fig. 5. Confusion matrix of logistic regression classification model

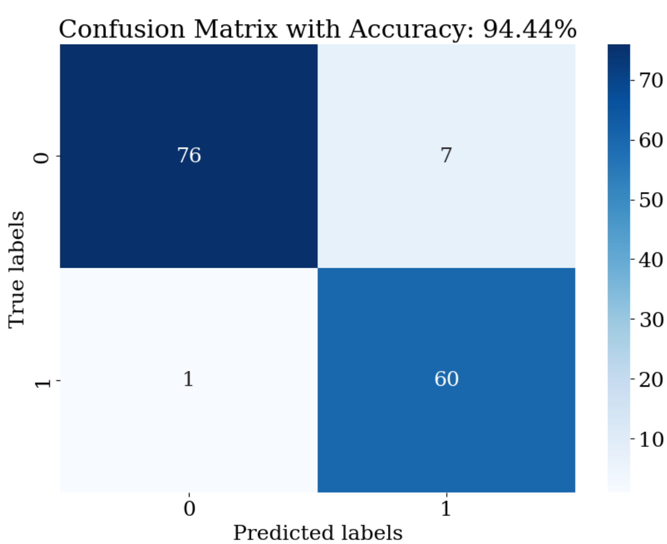

Fig. 6. Confusion matrix of GaussianNB classification model

The models indicated that key variables such as the time spent on social media, and the scores for ADHD, Anxiety, and Depression were predictive of the overall mental health outcome. This suggests that machine learning can effectively be used to gauge the severity of mental health deterioration due to social media usage. It also implies that interventions can be tailored based on individual usage patterns, offering a more personalized approach to mental health management.

The findings align with existing literature in digital psychology and mental health, which consistently reports negative impacts of prolonged social media use on mental health, particularly among younger populations. For instance, studies such as those by [13], [14] have documented similar associations between screen time and mental health issues like depression and anxiety. The predictive capacity of our models adds to this body of research by not only confirming these associations but also quantifying the risk levels, which could be instrumental in clinical settings.

Moreover, study results contribute to the emerging discussion on the applicability of machine learning in mental health assessment. Recent scholarly work suggests that data-driven approaches can revolutionize mental health diagnostics and intervention strategies [15]. By effectively identifying patterns that precede mental health declines, these models could facilitate earlier and more effective interventions, potentially curbing the adverse effects of social media on mental health.

Fig 7 illustrates the distribution of mental health scores (Anxiety, Self Esteem, and Depression) across different age groups. The plot shows fluctuating levels of Anxiety, Self Esteem, and Depression scores as age increases, with a notable increase in all scores as age approaches 90. This dramatic rise could indicate specific age-related factors influencing mental health or potential outliers in the dataset. Generally, the scores seem to stabilize or decrease slightly in middle age before this rise.

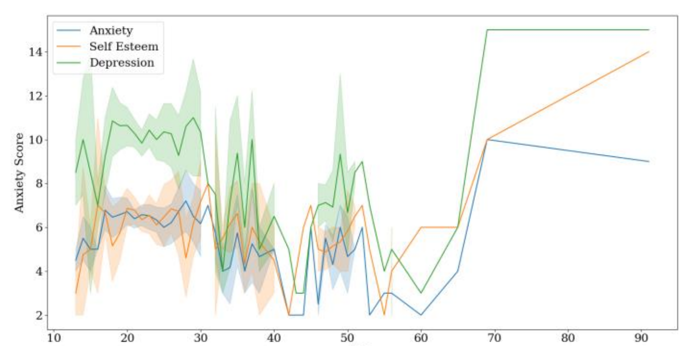

Fig. 7. Mental health scores by age

Fig. 8 depicts the relationship between the amount of time spent on social media and mental health scores. It shows a clear positive trend in Anxiety and Depression scores as the time spent increases, suggesting that higher social media usage correlates with worse mental health outcomes. The Self

Esteem score remains relatively stable initially but begins to decrease slightly as time spent increases beyond 4 hours, indicating potential negative impacts on self-esteem from excessive social media use.

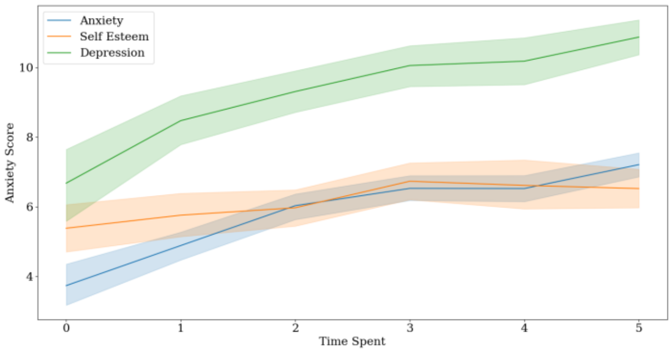

Fig. 8. Mental health scores by time spent

Finally, Fig. 9 examines mental health scores by sex, presenting Anxiety, Self Esteem, and Depression scores across different genders (presumably coded as 0, 1, and possibly non-binary as 2). Anxiety scores are highest for gender 0 and decrease with other genders, while Depression shows a steady decrease. Self Esteem scores, however, increase slightly for gender 1. This suggests varying impacts of social media on mental health depending on gender, with gender 0 possibly experiencing higher adverse effects.

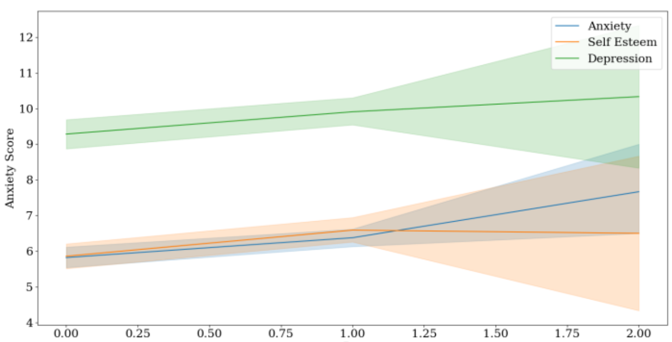

Fig. 9. Mental health scores by sex

V. Conclusions

This study has elucidated the intricate relationships between social media usage and various mental health outcomes, leveraging statistical analysis and predictive modeling to draw insightful conclusions. Key findings from the research demonstrate a clear association between increased time spent on social media and heightened symptoms of Anxiety, Depression, and deteriorations in Self Esteem. These outcomes are particularly pronounced in younger demographics, suggesting that age significantly modulates the psychological impact of social media.

The predictive models employed, specifically Logistic Regression and Gaussian Naive Bayes, have proven highly effective in forecasting mental health risks based on social media usage patterns. With accuracies of 99.31% and 94.44% respectively, these models underscore the potential of machine learning as a tool for early identification and intervention in mental health issues related to digital habits.

Furthermore, the study has shown distinct variations in mental health impacts across different genders, indicating that social media affects diverse populations in unique ways. This finding calls for tailored approaches in mental health interventions to address the specific needs of different demographic groups.

References

[1] L. O. M. N. A. R. Syah, S. Yusran, and L. O. A. I. Ahmad, “Pengaruh intensitas penggunaan media sosial terhadap kesehatan mental di smp n 3 parigi kabupaten muna tahun 2023,” *Jurnal Wawasan Promosi Kesehatan*, vol. 4, no. 3, pp. 203–208, Dec. 2023, Accessed: May 07, 2024. [Online]. Available: https://ojs.uho.ac.id/index.php/winsjournal/article/view/46503

[2] J. F. López-Gil *et al.*, “Global Proportion of Disordered Eating in Children and Adolescents: A Systematic Review and Meta-analysis,” *JAMA Pediatr*, vol. 177, no. 4, pp. 363–372, Apr. 2023, doi: 10.1001/JAMAPEDIATRICS.2022.5848.

[3] M. Mohammed and Y. M. Mohammed Yunus, “Understanding the Link Between Social Media Use and Mental Health Issues,” *Journal of Clinical Research and Applied Medicine*, vol. 3, no. 3, pp. 31–36, Nov. 2023, doi: 10.5530/JCRAM.3.3.7.

[4] V. Ajluni, “The Impact of Social Media on Human Flourishing and Mental Health: A Personalist Perspective,” *https://doi.org/10.1177/00243639231209972*, Oct. 2023, doi: 10.1177/00243639231209972.

[5] T. Saha, S. M. Reddy, S. Saha and P. Bhattacharyya, "Mental Health Disorder Identification From Motivational Conversations," in *IEEE Transactions on Computational Social Systems*, vol. 10, no. 3, pp. 1130-1139, June 2023, doi: 10.1109/TCSS.2022.3143763.

[6] A. Pourkeyvan, R. Safa and A. Sorourkhah, "Harnessing the Power of Hugging Face Transformers for Predicting Mental Health Disorders in Social Networks," in *IEEE Access*, vol. 12, pp. 28025-28035, 2024, doi: 10.1109/ACCESS.2024.3366653.

[7] B. Saha, T. Nguyen, D. Phung and S. Venkatesh, "A Framework for Classifying Online Mental Health-Related Communities With an Interest in Depression," in *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 4, pp. 1008-1015, July 2016, doi: 10.1109/JBHI.2016.2543741.

[8] M. E. Villa-Pérez, L. A. Trejo, M. B. Moin and E. Stroulia, "Extracting Mental Health Indicators From English and Spanish Social Media: A Machine Learning Approach," in *IEEE Access*, vol. 11, pp. 128135-128152, 2023, doi: 10.1109/ACCESS.2023.3332289.

[9] M. E. Aragón, A. P. López-Monroy, L. C. González-Gurrola and M. Montes-y-Gómez, "Detecting Mental Disorders in Social Media Through Emotional Patterns - The Case of Anorexia and Depression," in *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 211-222, 1 Jan.-March 2023, doi: 10.1109/TAFFC.2021.3075638.

[10] Rakesh, O. A. Shah and S. Vats, "Floorplanning and Comparative Analysis of 16-bit Synchronous Up/Down Counter in Different CMOS Technology," *2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*, Greater Noida, India, 2023, pp. 867-871, doi: 10.1109/CISES58720.2023.10183608.

[11] O. A. Shah, K. Sahu, Rakesh, Z. Ishrat, K. B. Ali and S. Vats, "Performance Evaluation of Pulse Triggered Flip-Flops in 32 nm CMOS Regime," *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, Salem, India, 2023, pp. 1477-1482, doi: 10.1109/ICAAIC56838.2023.10141264.

[12] K. Pearson, “Notes on the History of Correlation,” *Biometrika*, vol. 13, no. 1, p. 45, Oct. 1920, doi: 10.2307/2331722.

[13] 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," *Prev Med Rep*, vol. 12, pp. 271–283, 2018.
[14] H. C. Woods and H. Scott, "# Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem," *J Adolesc*, vol. 51, pp. 41–49, 2016.
[15] P. J. Uhlhaas *et al.*, "Towards a youth mental health paradigm: a perspective and roadmap," *Mol Psychiatry*, vol. 28, no. 8, pp. 3171–3181, 2023.