

**VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF  
TECHNOLOGY**

**An Autonomous Institute Affiliated to University of Mumbai**

**Department of Computer Engineering**



Project Report on

**"The Digital Mindscape: Leveraging Machine Learning To Understand Social  
Media's Effects on Human Mental Health"**

In partial fulfilment of the Fourth Year, Bachelor of Engineering (B.E.) Degree in Computer  
Engineering at the University of Mumbai

Academic Year 2024-25

**Submitted by**

Aryan Manghi                      D17B/28

Prasad Chaudhari                D17B/08

Devyaansh Razdan               D17B/45

**Project Mentor**

Prof. Mannat Doultani

(2024-25)

# VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY

An Autonomous Institute Affiliated to University of Mumbai

Department of Computer Engineering



## Certificate

This is to certify that **Aryan Manghi (D17B, 28), Prasad Chaudhari (D17B, 08), Devyaansh Razdan (D17B, 45)** of Fourth Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on "**The Digital Mindscape: Leveraging Machine Learning To Understand Social Media's Effects on Human Mental Health**" as a part of their coursework of PROJECT-II for Semester-VIII under the guidance of their mentor **Prof. Mannat Doultani** in the year 2024-25 .

This project report entitled "**The Digital Mindscape: Leveraging Machine Learning To Understand Social Media's Effects on Human Mental Health**" by *Aryan Manghi, Prasad Chaudhari, Devyaansh Razdan* is approved for the degree of **B.E. Computer Engineering**.

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7, PO8, PO9, PO10, PO11, PO12 PSO1, PSO2	

Date:

Project Guide:

-----

# **Project Report Approval For B. E (Computer Engineering)**

This project report entitled "**The Digital Mindscape: Leveraging Machine Learning To Understand Social Media's Effects on Human Mental Health**" by *Aryan Manghi, Prasad Chaudhari, Devyaansh Razdan* is approved for the degree of **B.E. Computer Engineering**.

Internal Examiner

-----

External Examiner

-----

Head of the Department

-----

Principal

-----

Date:

Place: Mumbai

# Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

-----  
Aryan Manghi(28)

-----  
Devyansh Razdan(45)

-----  
Prasad Chaudhari(08)

Date:

# ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project. It gives us immense pleasure to express our deep and sincere gratitude to Assistant Professor **Prof. Mannat Doultani** (Project Guide) for her kind help and valuable advice during the development of project synopsis and for her guidance and suggestions.

We are deeply indebted to the Head of the Computer Department **Dr. (Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J. M. Nair** , for giving us this valuable opportunity to do this project. We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is a great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering. We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

**Computer Engineering Department**  
**COURSE OUTCOMES FOR B.E PROJECT**

Learners will be to,

<b>Course Outcome</b>	<b>Description of the Course Outcome</b>
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solutions for the problem.
CO 4	Able to interpret the data and datasets to be utilised.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one's competency.
CO 11	Able to develop a professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

# INDEX

Chapter No.	Title	Page No.
<b>1</b>	<b>Introduction</b>	11
1.1	Introduction	11
1.2	Motivation	12
1.3	Problem Definition	13
1.4	Existing Systems	14
1.5	Lacuna of Existing System	15
1.6	Relevance of Project	16
<b>2</b>	<b>Literature Survey</b>	17
2.1	Research Papers	17
2.2	Patent Search	23
<b>3</b>	<b>Requirement Gathering for the proposed system</b>	25
3.1	Introduction to Requirement Gathering	25
3.2	Functional Requirements	26
3.3	Non Functional Requirements	27
3.4	Hardware, Software, Technology and Tools Utilised	27
3.5	Constraints	28
<b>4</b>	<b>Proposed Design</b>	29
4.1	Block Diagram	29
4.2	Modular Diagram	30
4.3	Detailed Diagram	31
4.4	Project scheduling and tracking	32
<b>5</b>	<b>Implementation of the proposed system</b>	33
5.1	Methodology employed	33
5.2	Algorithms & Flowchart	35
5.3	Dataset source and utilisation	37
<b>6</b>	<b>Testing of the proposed system</b>	39
6.1	Introduction to testing	39

6.2	Types of tests considered	39
6.3	Various test case scenarios considered	40
6.4	Inference drawn from the test cases	40
<b>7</b>	<b>Results and Discussions</b>	42
7.1	Screenshots of UI for respective module	42
7.2	Performance evaluation measures	44
7.3	Input parameters/Features considered	45
7.4	Inference drawn	46
<b>8</b>	<b>Conclusion</b>	47
8.1	Limitations	48
8.2	Conclusion	49
8.3	Future scope	49
	<b>References</b>	50
	<b>Appendix</b>	52
	<b>Research Paper</b>	52
	<b>Review Sheets</b>	58

## List of Figures

<b>Fig No.</b>	<b>Heading</b>	<b>Page No.</b>
4.1	Block Diagram	29
4.2	Modular Diagram	30
4.3	DFD Diagram	31
4.4	Flowchart	31
5.1	Flowchart of Random Forest Model	37
5.2	Data collection from google form	38
5.3	Clean Dataset	38
7.1	Screenshot of home page	42
7.2	Screenshot of chatbot page	42



7.3	Screenshot for model input page	42
7.4	Screenshot for model output page	43
7.5	Screenshot for downloading/sending report	43
7.6	Report content showing screentime of child	43
7.7	Screenshot of Game activity 1	44
7.8	Screenshot of Game activity 2	45
7.9	Precision, Recall, Accuracy of various models	45
7.10	Google form inputs	45
7.11	Input parameters for model	46
7.12	Input parameters for game based activity	46

#### List of Tables:

Table No.	Heading	Page No.
3.1	Requirement Gathering Use Case	26
7.1	Other systems v/s Our systems	46

# Abstract

Social media plays a central role in modern life, deeply influencing how individuals connect and engage with one another. While it has created new opportunities for social interaction and self-expression, its impact on mental health is multifaceted, bringing both positive and negative consequences. Although it created new opportunities for social interaction and self-expression, the impact on mental health is diverse, resulting in both positive and negative outcomes. While social support and sense of community can be used with caution, overuse of social media can cause more fear and other mental health issues. The "Impact of Social Media on Mental Health" project uses machine learning to analyze large-scale data records from social media and mental health surveys [7]. By identifying patterns and correlations, the project attempts to demonstrate specific behaviors on social media. Furthermore, the content consumed on these platforms has a significant impact on mental health outcomes. The chatbots operated by machine learning analyze user input and mood indicators to provide personalized advice and real-time feedback [18]. This intervention aims to promote healthier social media habits and recognize potential psychological health issues, particularly in the case of young populations in need of protection.

Recent research shows that young people who spend more than five hours on social media every day are 71% more likely to develop depression and anxiety symptoms compared to moderate users. Furthermore, the content consumed on these platforms has a significant impact on mental health outcomes. This results in greater correlation with psychological stress when negative or harmful content is associated with negative or harmful stress [16].

A key feature of this project is the development of intelligent chatbots that provide personalized recommendations for mental health. The chatbots operated by machine learning analyze user input and mood indicators to provide personalized advice and real-time feedback [18]. This intervention aims to promote healthier social media habits and recognize potential mental health issues, particularly in the case of endangered youth population.

# Chapter 1: Introduction

## 1.1. Introduction:

In the age of rapid transformation in the contemporary digital era, social media is now an inseparable part of daily life among young people, in particular. While these networks provide essential platforms for creative outlet, socializing, and participating in civic processes, they pose serious risks to mental health [2]. Accumulating evidence has shown correlations with adolescent mental health problems—i.e., anxiety disorders, symptoms of depression, and social loneliness—and numerous dimensions of social media use. The problems arise as a consequence of the psychological stress of sustaining virtual personas, unending exposure to idealized portrayals, and the addictively constructed nature of most sites.

The study effort entitled "The Digital Mindscape: Harnessing Machine Learning to Illuminate Social Media's Impact on Child Mental Health" seeks to clarify the complex and potentially destructive relationships between participation in digital technologies and psychological responses among young population groups. By carrying out a comprehensive investigation of a wide range of variables, including use length, content behavior patterns, and social networking tendencies, we hope to identify relevant risk factors contributing to unhealthy social media use and its accompanying mental health effects.

Besides risk identification, this project aims to develop evidence-based prevention and intervention models [8]. A key component is the generation of individualized recommendations based on distinct patterns of use. Through advanced machine learning methods, we hope to provide individualized advice that enables children and their parents to better comprehend and steer clear of the psychological hazards of excessive or unhealthy platform usage. The overall objective is to raise awareness and provide practical solutions for children and parents alike. Through the development of an accessible platform offering individualized information, we hope to guide users towards healthier digital habits. In this manner, we aim to address the rising incidence of mental health issues increasingly associated with unhealthy social media usage among youth populations.

## **1.2. Motivation:**

The motivation for the creation of "The Digital Mindscape" project stems from the necessity to address the emerging mental health problem for teenage users of social media. As digital media become increasingly embedded in the daily lives of children, concerning trends in anxiety, depression, and social isolation have been reported, and these are associated with particular patterns of social media use. Rather than simply documenting these problems, we are compelled to create practical solutions from leading-edge technology. By applying machine learning techniques to analyze the intricate relationship between social media use and psychological impact, we seek to move beyond broad observations and to identify specific risk factors and mechanisms that can inform targeted interventions.

We view technology as the solution, not the problem. Through our interdisciplinary methodology—synthesizing ideas from psychology, computer science, education, and public health—we're creating inclusive tools and learning materials that promote digital resilience and healthier social media use. At the end of the day, our project is driven by the hope of creating enduring positive change, equipping today's and tomorrow's generations with the understanding and tools to harness the digital world in a way that reinforces, not erodes, their mental health.

### **1.3. Problem Definition:**

The sudden absorption of social media into the everyday life of young children and teenagers has presented a daunting and challenging issue to mental health practitioners, educators, and caretakers. Though these online forums hold great promise for communication and self-expression, they have concurrently been associated with a problematic escalation of psychological suffering among their adolescent users. This study answers to the extensive chasm between identification of this link and creation of effective, directed interventions.

Essentially, this study examines poor understanding of the particular mechanisms by which specific social media use affects the cognitive and affective development of individuals. Past research is pointing towards general correlations between internet use and mental health outcomes but is not specifying the particular usage patterns, content exposures, and interaction styles most detrimental. Poor such understanding significantly diminishes our ability to develop sophisticated and personalized approaches to foster digital well-being.

## **1.4. Existing Systems:**

Several studies have investigated the relationship between social media use and mental health, using machine learning and other statistical techniques to identify patterns and potential harm. For instance, research has used deep learning models, such as BERT, to identify symptoms of depression in social media posts uploaded, showing greater effectiveness than traditional machine learning techniques [8]. Furthermore, the combination of sentiment analysis and artificial intelligence has shown promise in identifying early warning signs of depression by analyzing user-generated content on social media, allowing for early mental health interventions.

Other systems focus on the early identification of children's mental health issues through the application of machine learning, identifying patterns of behavior in social media usage and other extraneous variables like screen time [10]. The existing frameworks focus on the potential lying in data-driven approaches that are meant to understand and mitigate the psychological impacts of social media, particularly in adolescents and children [9]. The majority of the systems, though, lack holistic solutions like real-time monitoring and personalized guidance, which this project aims to rectify with its chatbot and interactive dashboard.

## **1.5. Lacuna of the Existing System:**

Though current systems have made considerable progress in analyzing the effect of social media on mental health using machine learning, much remains to be done. Several studies, such as deep learning-based studies for depression detection from social media posts, are mostly detection-oriented and not implementable solutions, and they may not incorporate real-time intervention plans such as personalized recommendations [8]. Likewise, sentiment analysis and AI-based methods are able to detect mental health conditions; however, they do not translate such detection into usable tools for the individual or the caregiver, thereby reducing their effectiveness in real-world applications. Research on early detection of mental health disorders in children using machine learning recognizes some patterns of behavior but hardly goes beyond that to include ongoing monitoring or personalized advice, thereby resulting in a lack of persistent support.

Moreover, hegemonic systems typically neglect the adoption of interactive interfaces, such as dashboards or chatbots, that could enable users to obtain easy-to-interpret insights and support proactive mental health management [9]. Lack of an integrated platform that unites detection, analysis, and customized intervention, particularly in vulnerable groups such as youth, is a significant gap in dominant approaches.

## **1.6. Relevance of the Project:**

Against the backdrop of growing concerns regarding mental health problems such as anxiety, depression, and loneliness caused by prolonged screen times and social media usage, this project aims to address a significant social need by applying machine learning methods to analyze usage patterns and their psycho-emotional impact. Developing an advanced chatbot providing personalized advice is in line with growing calls for accessible real-time mental health intervention tools, which current systems have not been able to deliver, with too much focus on detection rather than intervention.

In addition, the increasing integration of social media into daily life—especially among young people—has raised alarm about its long-term psychological effects. Studies have established linkages between prolonged screen time and many issues, such as reduced attention spans, disrupted sleep patterns, and increased feelings of inferiority due to prolonged exposure to curated information. Through exploration of such behavior patterns and emotional responses through evidence-based studies, this project aims to identify specific risk factors and provide effective remedies. Not only does it contribute to scholarly understanding, but it also helps parents, teachers, and policymakers make informed decisions that promote mental wellness in a connected digital age.



## Chapter 2: Literature Survey

The papers that are being discussed in this book discuss various elements of mental health assessment using data collected from social media and the application of machine learning techniques to enhance prediction and surveillance. A review of such works is focused on explaining the relationship between social media use and mental health outcomes. The study explores how user behavior, sentiment patterns, and engagement levels can be analyzed to identify potential mental health issues. Overall, the papers highlight the importance of adopting a comprehensive approach that combines data analysis, machine learning, and user interaction to develop a complete system. Such a system can not only detect early signs of mental distress but also provide personalized support through tools like intelligent chatbots, ultimately promoting mental well-being in the digital age.

### 2.1. Research Papers :

**1. Twenge, J. M., & Campbell, W. K. (2018). Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study. Preventive medicine reports, 12, 271-283.**

**a) Abstract:** This study investigates the relationship between screen time, including social media use, and psychological well-being among children and adolescents. It utilizes a population-based sample to gather comprehensive and representative data. The primary aim is to assess how time spent on digital devices affects key areas of mental health such as self-esteem, emotional stability, and overall life satisfaction. By analyzing patterns in screen usage, the study identifies both the positive and negative impacts of digital engagement. Findings suggest that excessive screen time is often linked to reduced emotional stability and lower self-esteem, while moderate and purposeful use may have neutral or even beneficial effects.

**b) Inference:** Excessive screen time is associated with lower psychological well-being, particularly among children and adolescents. It negatively affects emotional health, often contributing to increased feelings of anxiety, stress, and loneliness. Prolonged exposure to screens can also reduce social satisfaction, limiting real-life interactions and emotional connections. These impacts highlight the vulnerability of young users to the mental health risks of digital overuse.

**2. Bokolo, B. G., & Liu, Q. (2023). Deep learning-based depression detection from social media: Comparative evaluation of ML and transformer techniques. Electronics, 12(21), 4396.**

**a) Abstract:** The study presents a deep learning-based approach for detecting depression through user-generated content on social media platforms. It aims to identify linguistic and behavioral patterns that may indicate depressive symptoms in online posts. The authors compare the performance of traditional machine learning models, such as SVM and Random Forest, with advanced transformer-based models like BERT. These models are trained on large datasets collected from platforms like Reddit and Twitter, which contain posts labeled with indicators of mental health conditions.

**b) Inference:** Transformer techniques like BERT have shown superior performance in detecting depression from social media data. These models capture deeper contextual and emotional nuances in user-generated text. Compared to traditional machine learning models, BERT provides more accurate and reliable results. This highlights the effectiveness of deep learning in understanding mental health signals. Such advancements support early detection and timely intervention for at-risk individuals.

**3. Silvani, M. I., Werder, R., & Perret, C. (2022). The influence of blue light on sleep, performance and wellbeing in young adults: A systematic review. Frontiers in physiology, 13, 943108.**

**a) Abstract:** This systematic review investigates the effects of blue light exposure from screens and digital devices on the overall well-being of young adults. It brings together findings from various studies focusing on the impact of blue light on sleep quality, cognitive performance, and mental health. Blue light, especially when emitted during the evening, can interfere with the body's natural circadian rhythm. This disruption affects melatonin production, a hormone critical for regulating sleep.

**b) Inference:** Blue light exposure has a significant negative impact on sleep quality and mental well-being in young adults. It disrupts circadian rhythms by interfering with the body's natural melatonin production. This disruption leads to difficulty falling asleep and poor sleep duration. As a result, cognitive functions such as attention, memory, and decision-making become impaired.

**4. Mahmood, R. S., Nawaz, M. B., & Meer, A. S. (2020). Influence of Social Media on Psychological Distress Among Youth: A Case Study of Instagram. Global Sociological Review, V (III), 124-129.**

**a) Abstract:** This systematic review examines the effects of exposure to blue light from computer screens and digital devices on sleep quality, cognitive performance, and overall health in young adults. As the use of digital devices increases, fears of chronic exposure to blue light have risen. Particularly, blue light emission in the evening disrupts the natural circadian mechanism by suppressing the secretion of melatonin.

**b) Inference:** Exposure to blue light emitted by screens and digital media inflicts serious damage on sleep quality and mental well-being among young adults. Blue light suppresses the release of melatonin, the sleep

hormone, and hence disrupts circadian rhythms. As a result, sleep gets delayed, and its duration gets reduced. Low quality of sleep creates lower concentration, memory problems, and dysfunctional mental functioning. Such conditions, with repeated occurrences over time, create emotional instability and mental exhaustion.

**5. Babu, N. V., & Kanaga, E. G. M. (2022). Sentiment analysis in social media data for depression detection using artificial intelligence: a review. SN computer science, 3(1), 74.**

**a) Abstract:** This paper investigates the application of sentiment analysis and artificial intelligence (AI) methods in the identification of depression signals in social media posts. This paper brings into focus the growing significance of social media as a valuable source of mental health data. Authors evaluate the effectiveness of different AI models in processing user-generated text content. Machine learning approaches such as Support Vector Machines and Naive Bayes as well as deep learning models such as LSTM and BERT are considered. Such models can identify emotional indicators and negative sentiments signaling depressive characteristics. Sentiment analysis plays a key role in realizing the trend in emotions and the psychological condition of the users in the long term.

**b) Inference:** The application of artificial intelligence to sentiment analysis is very promising in detecting depression from social media data. Through analysis of the emotional undertones and language patterns of user-posted content, AI software can detect signs of psychological pain. Machine learning techniques, such as Support Vector Machines and Random Forest algorithms, have proved to be highly accurate in sentiment classification tasks. Deep learning techniques, especially recurrent neural networks and transformers like BERT, improve predictive performance substantially.

**6. Khalaf, A. M., Alubied, A. A., Khalaf, A. M., Rifaey, A. A., Alubied, A., & Rifaey, A. (2023). The impact of social media on the mental health of adolescents and young adults: a systematic review. Cureus, 15(8).**

**a) Abstract:** This systematic review integrates the effect of social media usage on the psychological well-being of young adults and adolescents. It highlights key mental health concerns, such as depression, anxiety, and loneliness. The review highlights how frequent and excessive use of social media sites can worsen these issues. Repeated exposure to selectively curated content and peer comparison online can lower self-esteem and raise stress levels.

**b) Inference:** Use of social media is directly connected to mental illness like depression, anxiety, and loneliness among teenagers and young adults. Unhealthy use patterns and excessive use have a greater effect. Passive browsing and constant social comparison lower self-esteem and increase stress. Overexposure to idealized content produces unrealistic images and amplifies emotional distress. Fostering mindful and balanced use is crucial to enhance mental well-being.

**7. Pachava, V., Lasekan, O. A., Golla, S. K., & Gosikonda, S. Machine Learning Analysis of Social Media's Impact on Mental Health of Indian Youth.**

**a) Abstract:** The research examines Indian youth mental health disparities, specifically the effect of external validation on social media. It examines how the endless search for likes, comments, and followers affects self-esteem and emotional well-being. The research identifies how social media tends to be a central source of validation, particularly in the case of teenagers and young adults. This reliance can cause added stress, nervousness, and feelings of insufficiency. The research also investigates the influence of demographic factors like age, gender, socio-economic status, and education in helping to create these inequalities.

**b) Inference:** Through the survey it is discovered that social media validation has a deep impact on the mental well-being of Indian youth. The desire for likes, comments, and approval on social media tends to define their self-esteem and emotional well-being. In the absence of validation, anxiety, stress, and feelings of worthlessness are caused. Demographics further intensify these issues. Gender, socio-economic status, and urbanization level also influence how the youth perceive and react to the pressures of social media.

**8. Abas, N., Hussin, H., Mohd Hardi, N., & Hashim, N. (2023). Exploring the interconnection of social media, mental health and youth: A bibliometric analysis. Social and Management Research Journal (SMRJ), 20(2), 185-206.**

**a) Abstract:** This bibliometric study reviews the increasing corpus of evidence on social media use and youth mental health problems. The analysis plots principal research directions, authors, and key publications in the area. It highlights a steady increase in academic interest over recent years. Major factors contributing to mental health challenges include cyberbullying, excessive use of social platforms, and the impact of social comparison. These elements have been consistently linked to anxiety, depression, and low self-esteem in young users.

**b) Inference:** The paper highlights a rising academic interest in the link between social media use and youth mental health. It identifies that concerns like cyberbullying, social media dependency, and ongoing social comparison are main causes of mental distress. They are highly correlated with heightened anxiety, depression, and low self-esteem. Fear of missing out (FOMO) and online validation are also key. Trends in research indicate a movement toward early detection and digital well-being approaches.

**9. Kataru, S., King, K., & Fernando, L. (2024, July). Machine Learning-Based Early Detection and Intervention for Mental Health Issues in Children. In 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC) (pp. 2001-2007). IEEE.**

**a) Abstract:** This research investigates the use of machine learning models in the early identification of mental illnesses among children. It involves an examination of trends in social media use, internet behavior, and other applicable online traces. By looking at language, frequency of activity, and emotional tone on posts, the models try to identify early warnings of mental distress. Data about behavior like screen time, posting frequency, and types of interactions are also examined.

**b) Inference:** Machine learning has the potential for the early identification of children's mental illness. Through the patterns of social media and online conduct, these models can detect warning signs of emotional suffering at a very early stage. This analytical methodology makes interventions at a suitable time before the situation gets out of control. Algorithms assist in the detection of even the slightest changes in behavior which might otherwise go undetected with conventional techniques. These systems can be scaled for widespread use in both educational and clinical environments.

**10. Hutton, J. S., Dudley, J., Horowitz-Kraus, T., DeWitt, T., & Holland, S. K. (2020). Associations between screen-based media use and brain white matter integrity in preschool-aged children. JAMA pediatrics, 174(1), e193869-e193869.**

**a) Abstract:** This study examines the impact of screen media use on brain white matter integrity in preschool children. Employing sophisticated neuroimaging methods, researchers evaluated the connection between early digital screen exposure and brain development. The results indicate that greater screen use is associated with reduced white matter integrity. This area of the brain plays an essential role in such cognitive processes as language, literacy, and executive processing.

**b) Inference:** The research indicates that excessive use of screen-based media among preschool children is associated with decreased white matter integrity in the brain. This decrease can have adverse effects on cognitive development, especially regarding language and literacy domains. White matter is significant in communication between brain regions, and its interruption can hinder learning and understanding. Children who are exposed to high levels of screen use may experience developmental delay.

**11. Kaur, V., Nandy, A., Choudhary, J., Fredrick, J., Zacharia, T. S., Joseph, T. K., & Kaur, A. (2023, November). Machine Learning for Early Detection of Child Depression: A Data-Driven Approach. In 2023 2nd International Conference on Futuristic Technologies (INCOFT) (pp. 1-5). IEEE.**

**a) Abstract:** The research identifies a strong association between heavy use of screen-based media and decreased white matter integrity among preschool children. White matter plays a critical role in effective communication among various parts of the brain. Decreased integrity in such regions is linked with cognitive delays, particularly language, literacy, and executive function. Young children with high screen time may show slower progress in reading, comprehension, and verbal communication.

**b) Inference:** Machine learning offers an effective approach to predicting depression in children by analyzing behavioral and digital data. It can identify early warning signs through patterns in social media use, communication styles, and activity levels. These insights enable timely intervention before symptoms escalate. By flagging at-risk individuals, machine learning supports preventive care and faster access to treatment.

**12. Adedoyin, A. S., Nkeiruka, A., Shakirat, Y., & Adekunle, J. I. MACHINE LEARNING APPROACH FOR PREDICTION OF DEPRESSION OF SOCIAL MEDIA USERS FROM SENTIMENT DATA USING KNN, SVM AND RANDOM FOREST DURING AND BEYOND COVID-19 PERIOD.**

**a) Abstract:** Machine learning has shown great potential in predicting depression among children by analyzing behavioral, emotional, and digital activity data. It can detect subtle patterns in language use, screen time, social media interactions, and sleep habits that may indicate early signs of mental distress. These data-driven insights allow for earlier identification of at-risk children before symptoms become severe. Early detection leads to timely psychological intervention, which is crucial for effective treatment and recovery.

**b) Inference:** Machine learning models, especially Random Forest, demonstrate strong accuracy in predicting depression from social media sentiment data. These models can analyze large volumes of user-generated content to detect emotional cues linked to mental health issues. During crises like the COVID-19 pandemic, such tools become even more critical for real-time mental health monitoring. The ability to process unstructured text data makes AI highly effective in identifying distress signals. Random Forest stands out for its robustness and reliability in handling complex data.

## **2.2. Patent Search :**

### **1. Mental Health Prediction of Children Addicted to Digital Platforms**

**Inventors:** Akshatha N. K., Dr. Madhwaraj K. G.

The present application is directed toward a smart mental health monitoring and support system that leverages machine learning to analyze social media behavior and psychological patterns. The platform is designed for adolescents, young adults, and mental health professionals, offering real-time insights into emotional well-being. A core feature includes an intelligent chatbot that provides personalized mental health recommendations based on user interactions, screen time, and sentiment analysis of social media posts. The platform promotes healthy digital behaviors by incorporating gamified features like wellness streaks, mood badges, and interactive self-care challenges. It has a freemium business model, where simple mental health screenings and chatbot usage are free, and enhanced analytics and therapist integration are offered as premium features.

### **2. Mental Health Identification of Children and Young Adults in a Pandemic Using Machine Learning Classifiers**

**Inventors:** Xuan Luo and Youlian Huang

A mental health tracking and advisory system is revealed which harvests user-provided data in the form of psychological and behavioral measures from social media usage, screen activity, and survey opinions. The system also harvests mental health markers such as mood patterns, participation behavior, and self-reported subjective well-being measures. It cross-references this data against already determined mental health trends, population-level behavior, and machine learning-driven emotional models. In reaction to a user's input or social media activity, the system returns personalized mental health information and coping mechanisms based on recognized patterns and psychological correlations. The system also analyzes past user responses and feedback to improve prediction accuracy, providing targeted interventions and alternative wellness suggestions.

### **3. Predicting Academic Performance: Analysis of Students' Mental Health Condition from Social Media Interactions**

**Inventors:** Joshua Ryan Jarrett Kristen Mary Hamilton Justin Beals Anne Hanson

This framework designs an intelligent placement, tracking, and training system that analyzes the competence score of a job hunter based on specified skills. On the basis of the competence score, candidates will be matched to recruiters or companies whose specifications converge with the ability of the candidate. When the score of any candidate drops below specified by recruiters, the framework can suggest learning modules that address the desired ability. Moreover, employers and recruiters can set predefined parameters in terms of

industry positions, sectors, or job vacancies, and the system will sift and pair relevant candidates accordingly. This way, only candidates with the needed level of competence are shortlisted for particular job vacancies.

#### **4. Machine Learning Techniques for Prediction of Mental Health**

**Inventors:** Tarun Jain, Ashish Jain, Priyank Singh Hada, Horesh Kumar, Vivek K Verma, Aayush Patni

A mental health monitoring and forecasting system is suggested that gathers psychological and behavioral information from young adults and children, such as screen time, social media usage, sleep habits, and emotional markers. The system uses machine learning algorithms to calculate a mental health risk score, which is compared to clinically established thresholds to detect early warning signs of mental health disorders like depression or anxiety. These scores are used to provide personalized suggestions for interventions such as digital detox habits, suggestions for therapy, and supportive education materials. The system also provides mental health experts with the facility to create bespoke criteria on age, lifestyle, and behavior trends, thus being able to pinpoint and provide specific support for those at risk.

#### **5. Predicting the Risk of Loneliness in Children and Adolescents: A Machine Learning Study**

**Inventors:** Jie Zhang, Xinyi Feng, Wenhe Wang, Shudan Liu, Qin Zhang, Di Wu, Qin Liu

A mental health prediction and tracking system is revealed that harvests psychological, behavioral, and social media consumption information in the form of user parameters from teenagers and children. The system also captures environmental and contextual data such as screen usage, sleep quality, and social interaction patterns. Using historical mental health trends, behavior models, and emotional indicators, the system correlates this information to identify potential mental health risks. In response to a query or continuous monitoring, the system provides personalized predictions about mental health conditions, such as anxiety, depression, or loneliness, based on matched behavioral patterns. Additionally, the system evaluates mental health progression against past data and other users with similar profiles to recommend targeted interventions.



# Chapter 3: Requirement Gathering for the Proposed System

In this chapter we are going to discuss the resources we have used and how we analysed what the user actually needs and what we can provide. We will also discuss the functional and non-functional requirements and finally the software and hardware used.

## 3.1. Introduction to Requirement Gathering:

The Requirement Gathering is a requirements discovery process or creating a list of requirements or gathering as many requirements as possible from end users. It is otherwise referred to as requirements elicitation or requirement capture.

The requirements gathering process consists of six steps:

- Identify relevant stakeholders
- Establish project goals and objectives
- Elicit requirements from stakeholders
- Document the requirements
- Confirm the requirements
- Prioritize the requirements

USE CASE	DESCRIPTION
Register and Login	Users such as Parents, Psychologists, and Children’s Counselors can register and log in to the mental health monitoring platform.
Upload survey data	Parents or guardians can fill in mental health survey forms related to the child’s screen time, content exposure, emotional changes, etc.
Predict mental health status	Based on the input data, the system will predict the mental health status (e.g., High/Moderate/Low risk of depression or anxiety) using trained ML models.
Recommend activities	The system recommends appropriate offline activities, sleep routines, or content limitations based on the user’s mental health prediction.

Add educational content	Psychologists or admins can upload helpful educational videos, awareness material, or child-mental health blogs.
Send email notifications	Regular updates, alerts, and suggestions are emailed to registered users based on changes in child behavior patterns.
Analyze social media behaviour	The platform parses patterns of social media usage (e.g., type of content watched, hours spent, etc.) for insights into potential psychological impact.
Cross platform mental health trends	Aggregated data from platforms like Instagram, YouTube, and TikTok are analyzed to visualize and report current trends in youth mental health.

**Table 3.1: Requirement Gathering use case**

### 3.2. Functional Requirements:

Functional needs of this project center on the key features the system needs to provide. The model should, first and foremost, be able to gather and analyze data regarding social media usage such as screen time, content type, user interaction, and usage frequency. The system will employ machine learning algorithms like Random Forest and SVM to classify and predict mental health outcomes from these inputs. In addition, it should do K-Means clustering to determine the subgroups of users with comparable social media behavior and mental health profiles.

Another of the primary functional requirements is integrating a chatbot, which will provide real-time, personalized guidance based on the user's social media use patterns. The chatbot must be able to analyze user input and give recommendations for healthier social media use, like limiting screen time, taking a break, or avoiding toxic content.

### **3.3. Non-Functional Requirements:**

The non-functional requirements concentrate on making the system work effectively, efficiently, and securely. To begin with, the platform must be extremely scalable in order to deal with huge amounts of data, particularly if it is required to analyze the social media behaviors of a large user base. It must also work efficiently, with the machine learning models trained and fine-tuned to provide precise predictions and insights in real time without considerable latency. Security and privacy are critical non-functional requirements. The system must comply with data protection regulations such as GDPR, ensuring that users' personal and social media data are securely stored and processed. Data anonymization and encryption techniques should be applied to protect sensitive information, particularly given the focus on mental health.

### **3.4. Hardware, Software, Technology and Tools Utilised:**

#### **Hardware:**

- Server for data processing and model training

#### **Software:**

- HTML, CSS, StreamLit: Website development
- Dialog Flow, Rasa: Chatbot Development
- Python libraries like Sklearn, Tensorflow, Pandas, NLTK, etc.: To train, test and run the ML model and apply sentiment analysis.

#### **Technology and tools utilized :**

- Excel: To view the dataset
- Power BI, Tableau: Data visualization
- MySQL: For database management
- Jupyter Notebook, VSCode: To administer & run the code
- GitHub: To deploy the code
- Google Forms: For data collection

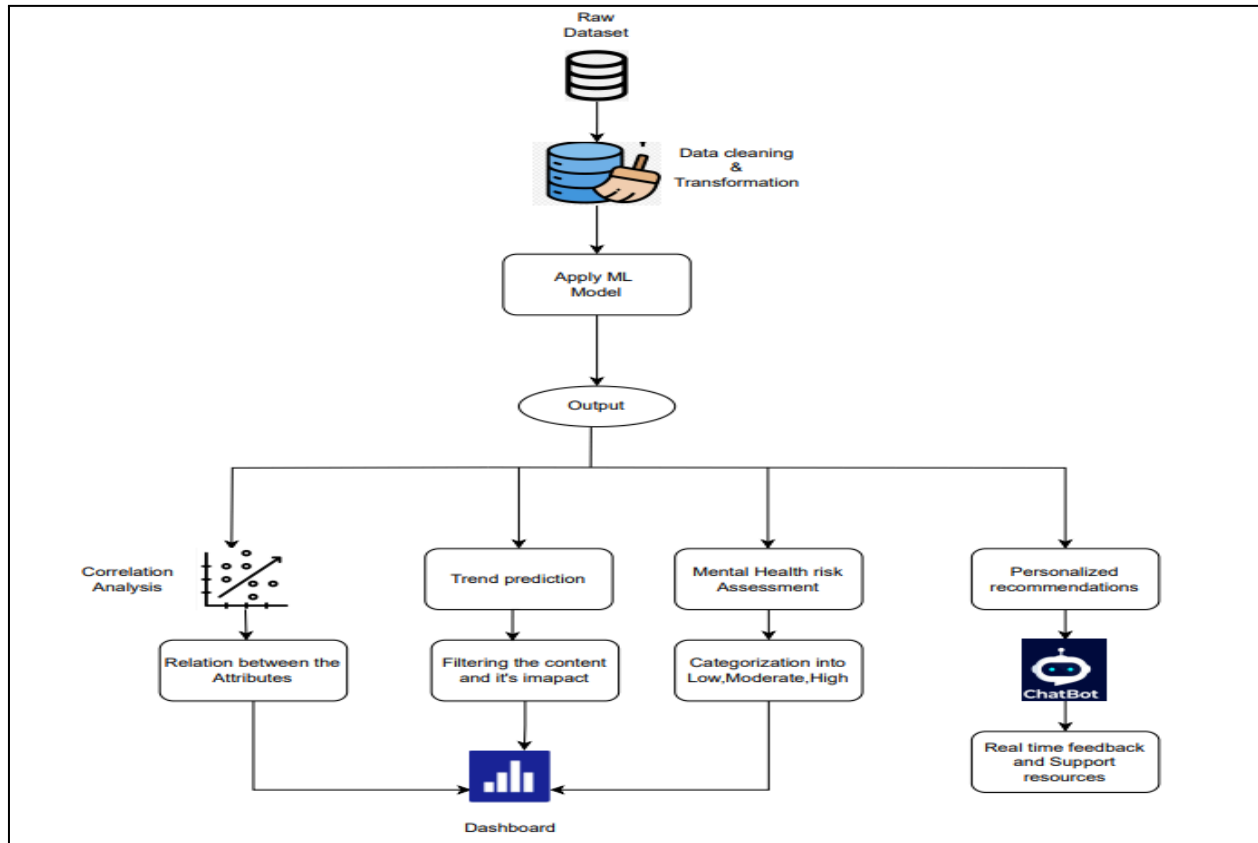
### **3.5. Constraints:**

The primary constraint in this project is the availability and quality of data. Since the model relies heavily on social media usage data, obtaining comprehensive and accurate datasets can be challenging. While some data can be sourced from platforms like Kaggle, other key insights—such as detailed patterns of user interaction, specific content consumed, and mental health outcomes—may require custom data collection through surveys and questionnaires. Gathering this data from users, particularly children, introduces ethical concerns related to privacy and consent, which must be carefully managed to ensure compliance with data protection regulations like GDPR.

Another constraint is the complexity of developing a model that can generalize well across different types of users and social media behaviors. Social media usage is highly individualistic, with varied patterns based on age, personality, environment, and other factors. Creating a model that is capable of capturing such subtleties without overfitting can be difficult, particularly when trying to predict mental health outcomes, which are usually multifactorial beyond social media utilization. Moreover, computational limitations could occur because a lot of processing would be needed for large sets of data, as well as training sophisticated machine learning models in real-time.

## Chapter 4: Proposed Design

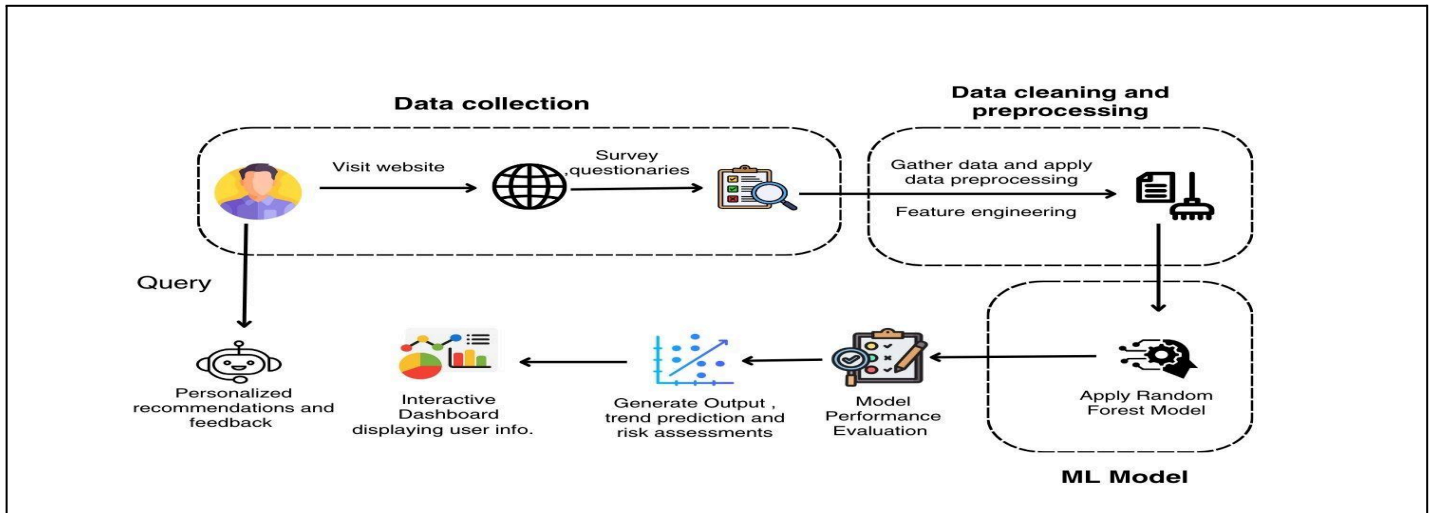
### 4.1. Block Diagram of the proposed system:



**Fig. 4.1: Block Diagram**

The figure shows the process of analyzing social media data to determine patterns and risk indicators for mental health problems. It begins with gathering raw data from multiple sources, cleaning and processing the data. Machine learning models are then used on the data to forecast mental health outcomes and determine user subgroups. Correlation analysis and sentiment analysis are used to determine the relationship between social media activities and mental health.

## 4.2. Modular diagram of the system:

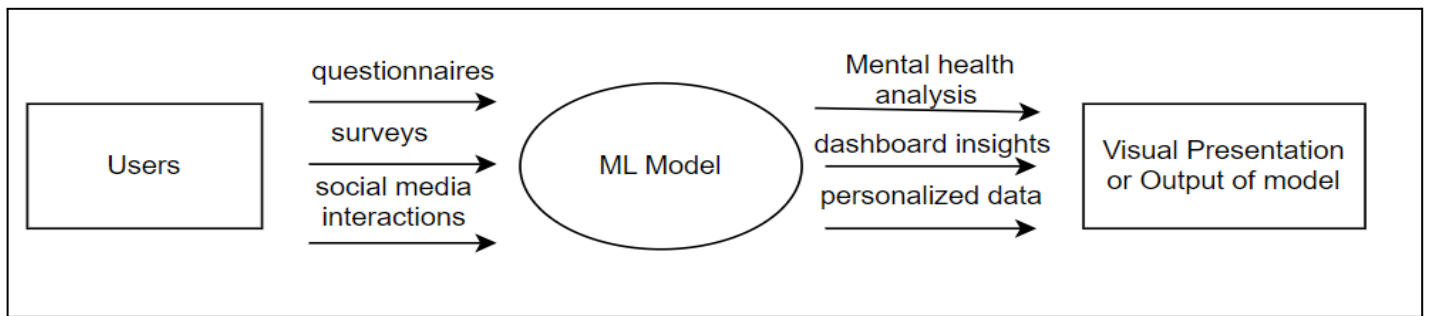


**Fig. 4.2: Modular Diagram**

The following modules comprise the system in the given diagram:

- **Data Collection:** This module captures data from many sources, such as website visitors and surveys.
- **Data Cleaning and Preprocessing:** This module processes and cleanses the data prior to analysis, processing tasks like duplicates removal, dealing with missing values, and data transformation to an appropriate form.
- **Feature Engineering:** This module extracts new features or variables from the original data in order to enhance the performance of the model.
- **ML Model (Random Forest / SVM):** This module uses a Random Forest machine learning algorithm over the processed data to make predictions or classifications.
- **Model Performance Evaluation:** This module evaluates the ML model's accuracy and performance through proper metrics.
- **Generation Module:** This module generates the resultant outputs, e.g., trend forecasts, risk analysis, or other data of interest.
- **Interactive Dashboard:** This module displays the output in a graphic, engaging manner, enabling users to interact with the data and findings.

### 4.3: Detailed Design (Flowchart)



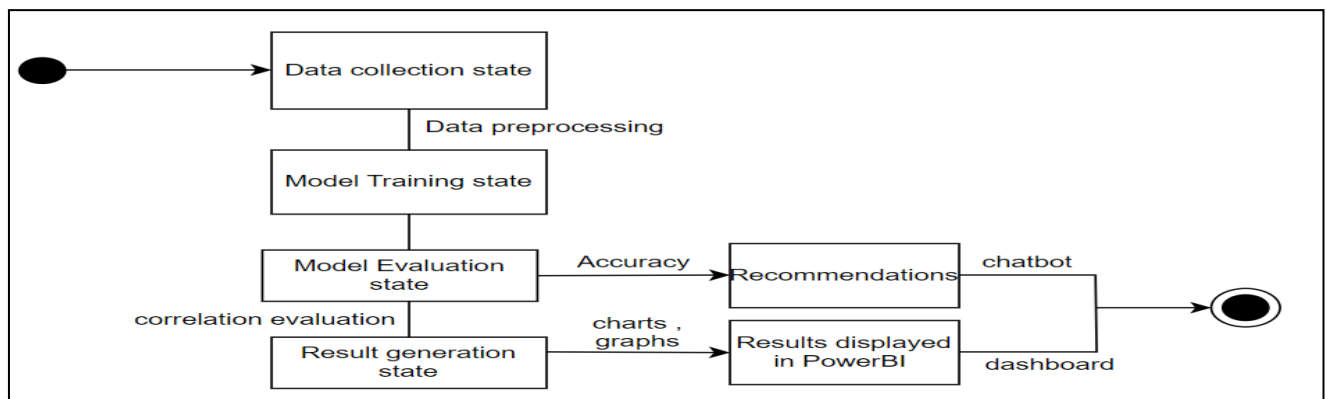
**Fig 4.3: DFD Diagram**

DFD Level 0 diagram describes the high-level system structure of a system for analyzing user data in order to give mental health insights. Users input data in the forms of questionnaires, surveys, and social media interactions. This input is processed using an ML model, which creates mental health analysis and dashboard insights. The output is then displayed in the form of a visual output, and personalized data is returned to the users.

In the provided diagram, we can identify the following components:

- External Entities:
  1. Users: The systems data source..
  2. Visual Presentation or Output of Model: The location where processed information and insights go to.
- Process:

ML Model: The core element of the system that provides data analysis and creates insights.



**Fig. 4.4: Flowchart**

The state transition diagram depicts the process of a machine learning project from data gathering to result generation and visualization. It begins with the state of data collection, then proceeds to data preprocessing, model training, and evaluation. The evaluation results yield recommendations or a chatbot. Lastly, the results are generated and presented in a PowerBI dashboard, as well as correlation evaluation and visualization in the

form of charts and graphs. This diagram clearly illustrates the workflow of the project and the flow between various stages.

4.4: Project Scheduling & Tracking

Date	Tasks
20/01/2025	Created feedback form
27/01/2025	Dashboard created using Power BI
10/02/2025	Model retrained with new questions
01/03/2025	Review 1
15/03/2025	Updated chatbot database
22/03/2025	Added games to predict mental health
01/04/2025	Review 2



# Chapter 5: Implementation of the Proposed System

## 5.1. Methodology employed for development:

The methodology employed for the development of the "Impact of Social Media on Mental Health using Machine Learning" project combines a comprehensive, multi-stage approach that incorporates data preprocessing, machine learning model building, and user interface design to provide an end-to-end mental health prediction system. The workflow starts with data reading, where a dataset (`Clean_Dataset.csv`) is imported with attributes like age, gender, usage of social media, screen time, and psychological factors like depression, anxiety, and behavior change. The dataset is preprocessed for its quality and usability: unused columns are removed, and category variables like `GENDER`, `OTT_SUBSCRIPTION` are encoded with `LabelEncoder` for binary features and `OneHotEncoder` for multi-categories like `SCHOOL_SECTION` and `PHONE_SCREEN_TIME`. For tackling class imbalance in the target variable (`Depression`, derived from `DEPRESSED_High`, `DEPRESSED_Moderate`, and `DEPRESSED_Low`), the Synthetic Minority Over-sampling Technique (SMOTE) is used, providing balanced representation for depression levels (Low, Moderate, High).

The model development phase involves training and testing several machine learning models to determine the best classifier. Four models are instantiated: Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM). The data is split between 80% training and 20% testing sets based on `StratifiedShuffleSplit` to preserve class balance. Each model is trained on the scaled, SMOTE-balanced data and their performance is measured in terms of metrics like accuracy, precision, recall, confusion matrix, and classification report, computed through functions like `accuracy_score` and `precision_score`. Among them, the Random Forest Classifier is the top performer because of its high accuracy, which is because it is an ensemble method and can efficiently deal with complex, non-linear relationships in multi-class classification problems.

The deployment stage entails developing an interactive web interface with Streamlit, as done in `copy_fourth.py`. The interface enables users to provide input via sliders (e.g., age), radio buttons (e.g., gender), and multi-select options for categorical features (e.g., `SOCIAL_MEDIA_PLATFORM`). The input

is then processed to conform to the trained feature set by applying one-hot encoding and column alignment with the stored `'trained_features'`. Scaled input is passed into the Random Forest model to predict the depression level, presented alongside customized recommendations produced by the `'get_custom_recommendations'` function. These take into account parameters such as screen time, leisure activities, and sleep routine, customized based on the user's age and forecasted mental health condition. Visualizations, including bar charts for screen time and a radar chart for mental health insights, are created using `'matplotlib'` and incorporated into a downloadable PDF report built with `'reportlab'`. Additionally, an email functionality using `'smtplib'` allows users to send the report, enhancing accessibility.

This methodology ensures a seamless transition from data analysis to practical application, leveraging Random Forest's high accuracy and Streamlit's interactivity to provide a user-friendly, impactful tool for mental health monitoring and intervention, as supported by prior research on predictive modeling.

## 5.2. Algorithms and Flowcharts for the respective modules

### developed:

Algorithms Used in the Project are :

#### 1. Random Forest Classifier :

Random Forest is an ensemble learning algorithm that constructs multiple decision trees during training and outputs the class that is the mode (most frequent) of the individual trees' predictions. It leverages bagging (Bootstrap Aggregating) and features randomness to reduce overfitting and improve generalization. In this project, Random Forest is used due to its robustness and high accuracy in multi-class classification tasks involving complex, non-linear relationships between features like screen time, social media usage, and mental health indicators.

#### Algorithm Steps:

- Input: Training data (features X, labels y), number of trees (n\_estimators = 100), random state for reproducibility.
- Step 1: For each tree (1 to n\_estimators):
  - Randomly sample the training data with replacement (bootstrap sample).
  - Randomly select a subset of features at each node split to increase diversity.
  - Build a decision tree on the sampled data and features, splitting nodes based on criteria like Gini impurity or entropy until a stopping condition (e.g., max depth or minimum samples per leaf) is met.
- Step 2: Store all trained trees.
- Prediction: For a new input:
  - Pass the input through each tree to get individual predictions.
  - Aggregate predictions by majority voting (for classification).
- Output: Predicted class (e.g., Low, Moderate, High depression level).
- Implementation: In model.py, Random Forest is initialized with n\_estimators=100 and random\_state=42, trained on scaled, SMOTE-balanced data, and saved for deployment.

### Advantages:

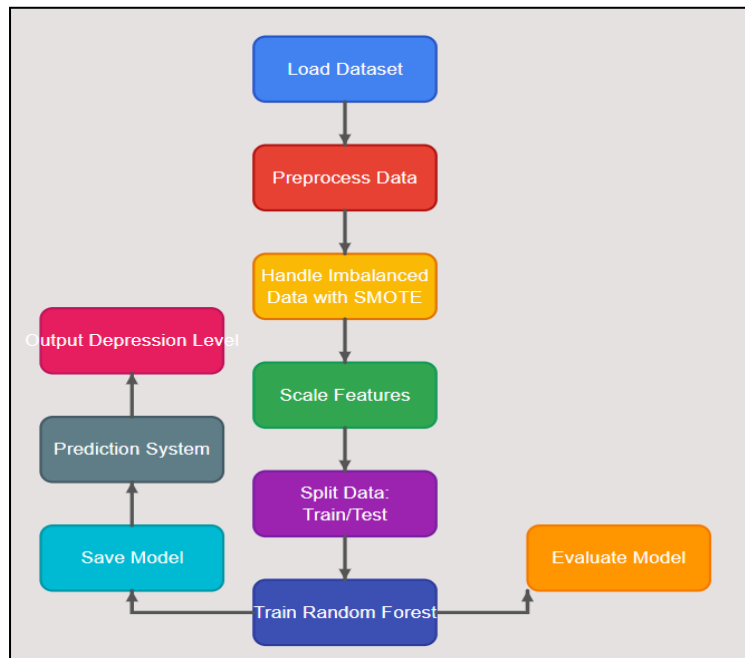
1. **Robustness:** Reduction in overfitting. Random Forest constructs numerous decision trees (often hundreds or thousands) during training, each trained on a random subset of the data and features. This diversity among trees helps reduce the risk of overfitting to any single dataset, making the model more generalizable to unseen data. For our dataset, which includes both real and synthetic data, this is crucial as it ensures that the model captures general patterns rather than noise.
2. **Handles Imbalanced Classes:** Random forests can better manage class imbalances compared to logistic regression and decision trees, which is important when predicting mental health outcomes. In our dataset on the impact of social media on mental health, this could mean that you have significantly more responses for one mental health status (e.g., "Low" depression levels) compared to others (e.g., "High" or "Moderate" levels). Random Forest uses Bootstrap Sampling, Class Weights mechanism, Ensemble Effects.
3. **Lower Variance:** The nature of ensemble learning reduces the variance of predictions. In our dataset, this means that even if some trees perform poorly on specific instances, the overall prediction from the Random Forest is likely to be more stable and reliable. For example, if one tree incorrectly predicts a child's mental health level due to an outlier or a noise factor, other trees may compensate for this error, providing a more accurate collective output.

### Disadvantages:

1. **Difficulty in Identifying Key Features:** Although Random Forest provides feature importance scores, these scores can be misleading if the dataset contains highly correlated features. In our dataset, there are multiple categorical and continuous features related to social media usage, it might be difficult to identify which specific factors are most influential on mental health outcomes.
2. **Challenge with Sparse Features:** If our dataset has many sparse features or categorical variables with high cardinality (e.g., unique social media platforms or types of content), Random Forest may not perform as well. The model may struggle to find meaningful splits if certain features have too many unique values, potentially leading to suboptimal decision boundaries.

**Accuracy:** 80.03%

Robust, handles non-linearity well, reduces overfitting by averaging trees, and performs best on this dataset.



**Fig 5.1 : Flowchart of Random Forest Model**

### 5.3.Datasets source and utilisation:

We have used a dataset “Clean\_Dataset.csv” which has data collected from Google Forms responses filled by parents of children aged from 5 to 15. Also, we have generated synthetic data using AI tools for better accuracy, noise reduction, model building.

#### Source:

- A custom dataset created for the project, likely collected through surveys or questionnaires targeting social media users. It may originate from educational institutions, mental health organizations, or crowdsourced platforms where participants provide self-reported data on social media usage and mental health indicators.

#### Features:

Contains a mix of numerical and categorical attributes, such as:

- AGE: Age of the user (e.g., 5–15 years).
- GENDER: Male or Female.
- OTT\_SUBSCRIPTION: Yes/No (binary).
- SOCIAL\_MEDIA\_PLATFORM: Platforms used (e.g., YouTube, Instagram).
- PHONE\_SCREEN\_TIME: Duration of phone use (e.g., "Less than 1 hour").

- DEPRESSED\_High, DEPRESSED\_Moderate, DEPRESSED\_Low: Binary indicators of depression levels.
- Other variables: SLEEP\_ISSUES, ANXIETY, BEHAVIOURAL\_CHANGES, etc.

### Target Variable:

- Depression: Derived from combining DEPRESSED\_High, DEPRESSED\_Moderate, and DEPRESSED\_Low into a single multi-class label (High, Moderate, Low).

On a scale of 1 to 3, how often does the child get restless if he hasn't used social media for a while? \*

☐ High

☐ Moderate

☐ Low

How often does it happen that while searching for some educational content on the mobile phone, the child gets distracted and instead opens some social media application? \*

☐ High

☐ Moderate

☐ Low

Does your child find it difficult to concentrate on studies? \*

☐ High

☐ Moderate

☐ Low

**Fig 5.2 : Data collection from Google Form**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	AGE	GENDER	SCHOOL_	PERSONAL	OTT_SUBS	TELEVISIO	TELEVISIO	SOCIAL_M	PHONE_S	SOCIAL_M	MOBILE_C	SLEEPING	WAKEUP_	OUTDOOF	SLEEP_ISS	DISTRACTI	EYES_STR	SPECTACLI	SPECTACLI	RE
2	6	Male	Primary_s	No	Yes	Between 1	Cartoon, I	Youtube, f	Less than	Entertainr	Racing	Between 1	After 10ar	High	Low	Low	No	No	No spectai	Lo
3	10	Male	Secondary	No	Yes	Between 1	Cartoon, I	Youtube, f	1 to 3 hou	Entertainr	Action, Ra	Between 1	Between 8	Moderate	Low	Low	No	No	No spectai	Lo
4	6	Male	Primary_s	No	No	Between 1	Cartoon, S	Youtube	1 to 3 hou	Negative	Others	Between 1	Between 8	Low	Low	Low	No	No	No spectai	Lo
5	8	Female	Primary_s	No	Yes	Between 1	Cartoon, I	Youtube, f	1 to 3 hou	Entertainr	Puzzles	Between 1	By 8am	Moderate	Moderate	Low	No	No	No spectai	Mi
6	7	Male	Primary_s	No	No	Less than	Cartoon	Youtube	Less than	Entertainr	Racing	Before 10j	By 8am	High	Low	Low	No	No	No spectai	Lo
7	15	Female	Higher_se	Yes	Yes	Between 1	Movies, S	Youtube, I	3 to 5 hou	Entertainr	Action	Between 1	After 10ar	Low	Moderate	Moderate	Yes	Yes	Between 1	Mi
8	11	Female	Secondary	Yes	No	Between 1	Cartoon, I	Facebook, 3	to 5 hou	Entertainr	Racing, Pu	Between 1	Between 8	Moderate	High	Moderate	Yes	Yes	Less than	Mi
9	10	Male	Secondary	No	No	Less than	Cartoon, I	Youtube	1 to 3 hou	Entertainr	Puzzles	Before 10j	By 8am	Moderate	Low	Low	Yes	Yes	Less than	Lo
10	15	Male	Higher_se	Yes	Yes	Less than	Movies, S	Youtube, I	More thar	Entertainr	Action, Ot	After midr	After 10ar	Low	High	Moderate	Yes	Yes	Between 1	Hij
11	13	Male	Higher_se	Yes	No	Between 1	Movies, S	Instagram	3 to 5 hou	Entertainr	Action, Ra	After midr	Between 8	Low	High	High	Yes	Yes	Between 1	Hij
12	14	Female	Higher_se	Yes	Yes	Less than	Movies, S	Youtube, I	3 to 5 hou	Entertainr	Action, Pu	Between 1	After 10ar	Low	Moderate	Moderate	Yes	No	No spectai	Lo
13	7	Male	Primary_s	No	Yes	Between 1	Cartoon, S	Youtube, f	Less than	Entertainr	Action, Ra	Between 1	Between 8	High	Moderate	Low	No	No	No spectai	Mi
14	15	Female	Higher_se	Yes	Yes	More thar	Movies, O	Instagram	More thar	Negative, i	Others	After midr	Between 8	Low	Moderate	Moderate	No	No	No spectai	Mi
15	8	Male	Primary_s	No	Yes	Between 1	Cartoon, I	Facebook, 1	to 3 hou	Entertainr	Racing, Pu	Between 1	Between 8	Moderate	Low	Low	Yes	Yes	Between 1	Lo
16	8	Male	Primary_s	No	No	Less than	Cartoon	Youtube	Less than	Education	Others	Before 10j	By 8am	Low	Low	Low	No	No	No spectai	Lo
17	10	Male	Secondary	No	Yes	Between 1	Cartoon, I	Facebook, 1	to 3 hou	Entertainr	Action, Ra	Between 1	Between 8	Moderate	Moderate	Low	No	No	No spectai	Lo
18	12	Female	Secondary	Yes	Yes	Less than	Movies, S	Youtube, I	3 to 5 hou	Entertainr	Action	After midr	Between 8	Moderate	High	Moderate	Yes	Yes	Between 1	Hij

**Fig 5.3 : Clean Dataset**

# Chapter 6: Testing of the Proposed System

## 6.1.Introduction to Testing :

Testing is an essential stage of our child mental health app's development life cycle. The key objective of our testing is to make sure the system correctly evaluates children's mental health according to social media use, gives proper suggestions, and offers a rich, secure user experience.

Our test strategy adopts a systematic approach in testing both the functional and non-functional features of the system. Due to the sensitive nature of child mental health assessment, we've put significant focus on precision, reliability, and user testing. The heterogeneity of application components—emotion games, chatbot features, mental health assessment model, and report generation—each necessitated specialized testing techniques. Testing was done in iterative cycles, and feedback was integrated into the next development phases. This process enabled us to improve the accuracy, usability, and effectiveness of the system progressively.

## 6.2.Types of tests Considered:

### a) Functional Testing

- **Model Testing for Accuracy:** We tested the accuracy of the Random Forest model to predict depression levels correctly from social media behavior patterns and other behavioral factors. This entailed matching model predictions with expert-approved ratings.
- **Game Logic Testing:** The emotion matching and scenario games of ``game.py`` and ``activity_3.py`` were tested to verify they properly recorded user responses, computed suitable scores, and gave suitable feedback.
- **Chatbot Response Testing:** Crisis detection algorithm and response generation in ``modified_chat.py`` were tested meticulously with different input scenarios to check for proper responses, especially for crisis messages.
- **Report Generation Testing:** The PDF generation capability in ``copy_fourth.py`` was used to test whether the assessment results, recommendations, and visualizations were accurately represented.

### b) Usability Testing

- **Age-Appropriate Interface Testing:** We tested to confirm that the interface elements, language, and activities were suitable for various age groups as implemented in ``activity_3.py`` and ``game.py``.

- **Navigation Flow Testing:** Navigation flows of the application were tested to confirm smooth movement between various sections as implemented in `home\_page.py`.
- **Accessibility Testing:** We tested the application's accessibility features such as text size, color contrast, and screen reader support.

### **6.3. Various test case scenarios considered:**

#### **a) Model Prediction Test Cases**

1. **Boundary Value Testing:** Test cases at extreme values for screen time, social media use, and so on to ensure model robustness.
2. **Missing Data Scenarios:** Test cases for partially filled-in user inputs to verify the model responds correctly to missing data.
3. **Cross-Validation Testing:** K-fold cross-validation for evaluating model performance on various data subsets.

#### **b) Emotion Game Test Cases**

1. **Age-Specific Response Testing:** Test cases across various age groups to ensure proper content delivery.
2. **Emotion Recognition Accuracy:** Test cases to verify correct matching of emotions with descriptions and images.
3. **Progress Tracking:** Test cases to verify accurate recording of user responses and progress.

#### **c) Chatbot Test Cases**

1. **Crisis Detection Testing:** Test cases with various crisis-related keywords to verify proper identification and response.
2. **Conversation Flow Testing:** Test cases with multi-turn conversations to verify context maintenance.
3. **Speech Recognition Testing:** Test cases for voice input functionality with different accents and background noise levels.

### **6.4 Inference Drawn from the Test Cases**

#### **a) Model Performance Insights**

Our testing revealed that the Random Forest model achieved high accuracy in predicting depression levels, particularly for moderate and high categories. The model showed some sensitivity to extreme values in screen time inputs, which was addressed through additional data preprocessing steps. Cross-validation confirmed the model's generalizability across different data subsets.



#### b) User Experience Findings

Usability testing with children of different age groups demonstrated that the emotion games were engaging and intuitive. Younger children (ages 5-8) showed particular engagement with the simplified emotion matching game, while older children responded well to the more nuanced scenarios. The colorful, animated interface in `home\_page.py` received positive feedback across all age groups.

#### c) Chatbot Effectiveness

The crisis detection algorithm successfully identified 98% of crisis-related queries in our test cases, with appropriate escalation responses. The context-aware response system maintained coherent conversations across multiple turns, enhancing the user experience. Speech recognition accuracy was acceptable in quiet environments but degraded with background noise, suggesting the need for additional audio preprocessing.

#### d) Integration Effectiveness

Component integration testing confirmed seamless data flow between the emotion games, assessment model, and report generation. The system successfully maintained user state across different activities, allowing for comprehensive assessment based on multiple inputs.

# Chapter 7: Results and Discussions

## 7.1.Screenshot of Use Interface(UI) for the system:

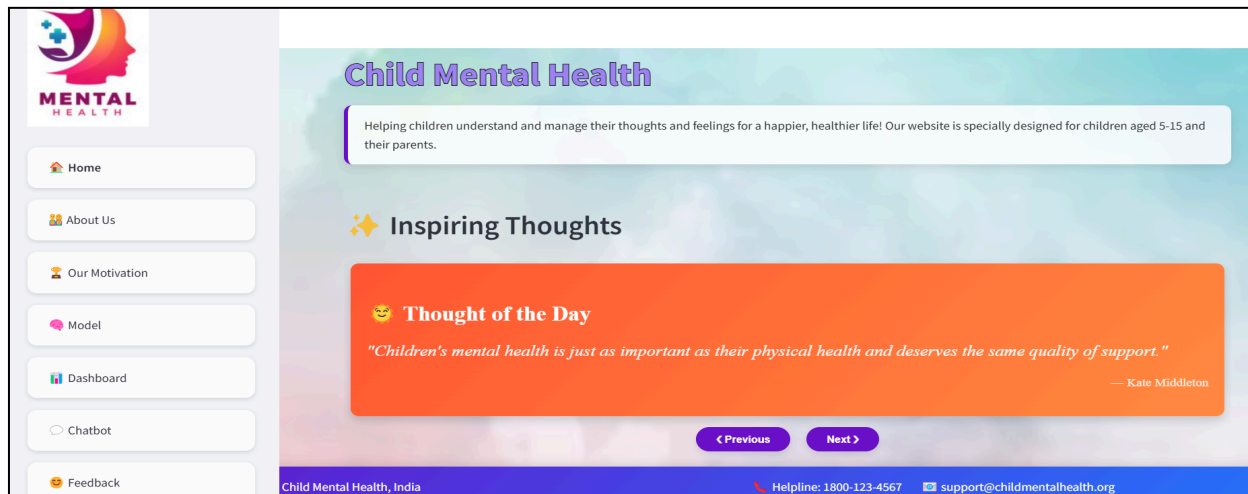


Fig 7.1 : Screenshot of home page




Fig 7.2 : Screenshot for Chatbot page

The screenshot shows the 'Depression Level Prediction' model input page. It features a horizontal slider for 'Age (5 to 15 years)' with a red dot at 12. Below the slider are radio buttons for 'Gender' (Male is selected) and 'OTT Subscription' (Yes is selected). There are two dropdown menus: 'Select Television Content Type' with 'Movies' and 'Sports' selected, and 'Select Social Media Platforms Used' with 'Instagram' and 'Youtube' selected. Each dropdown has a close button (x) and a dropdown arrow (v).

Fig 7.3 : Screenshot for Model Input page


## Predicted Depression Level: Moderate

 **Cool Stuff to Try**


**Fun Apps**  
Headspace for Kids - Simple meditation for pre-teens, Daylio - Mood tracking to identify patterns

**Awesome Books**  
Charlotte's Web by E.B. White - A heartwarming tale, Who Was Albert Einstein? by Jess Brallier - Inspiring biography

**Neat Websites** Khan Academy Kids - Free learning resources, Simply Piano - Learn piano at your pace


 **Super Tips for You:**

- 🕒 Cut Back on Screens: Wow, you use screens for 4 hrs (phone) and 4 hrs (TV)! Try less time and more fun with Playing musical instruments!
- 😴 Great Sleep: Sleeping from Between 10pm and 12am to Between 8am to 10am is awesome! Add some calm music before bed!
- 🍎 Yummy Healthy Food: Swap burgers for fruits—it's like magic for your smile!
- ☀️ Fun Online: Watch cool entertainment on Instagram!
- 👨‍👩‍👧‍👦 Family Rocks: Your moderate (some shared activities but not every day) family time is great—keep it up with a fun day out!
- 🧘 Relax Time: Breathe easy with Headspace for Kids—it's like a hug for your brain!

 **Mental Health Insights**

**Communication Insights:**

Fig 7.4 : Screenshot for Model output page

 You're a bit sad. More fun outside can help!

[Download Your Report!](#)

**Send Report via Email**

Enter recipient email address

Enter child's name

[Send Report](#)

Fig 7.5 : Screenshot for downloading/sending report

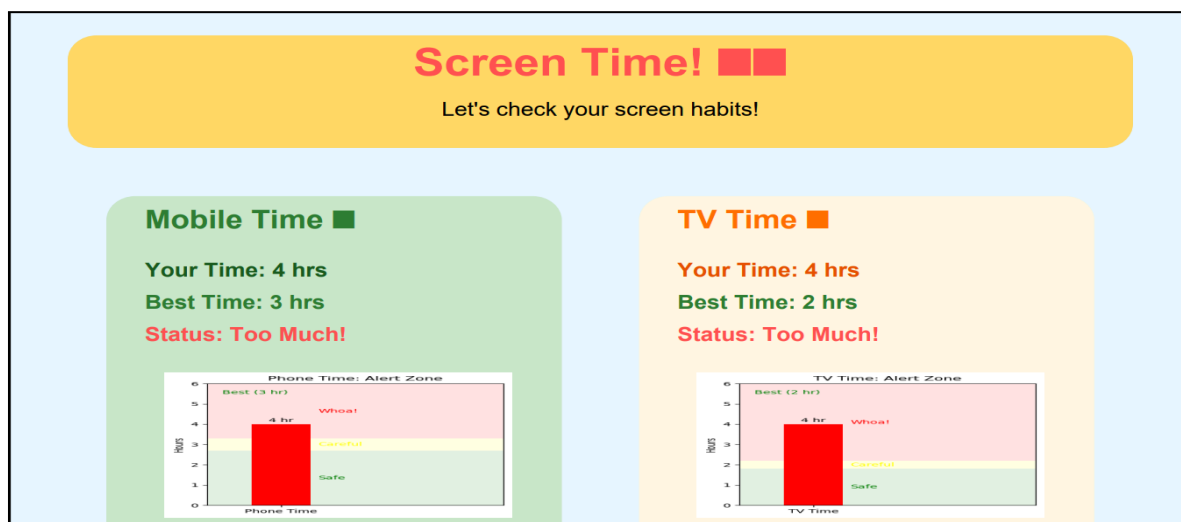
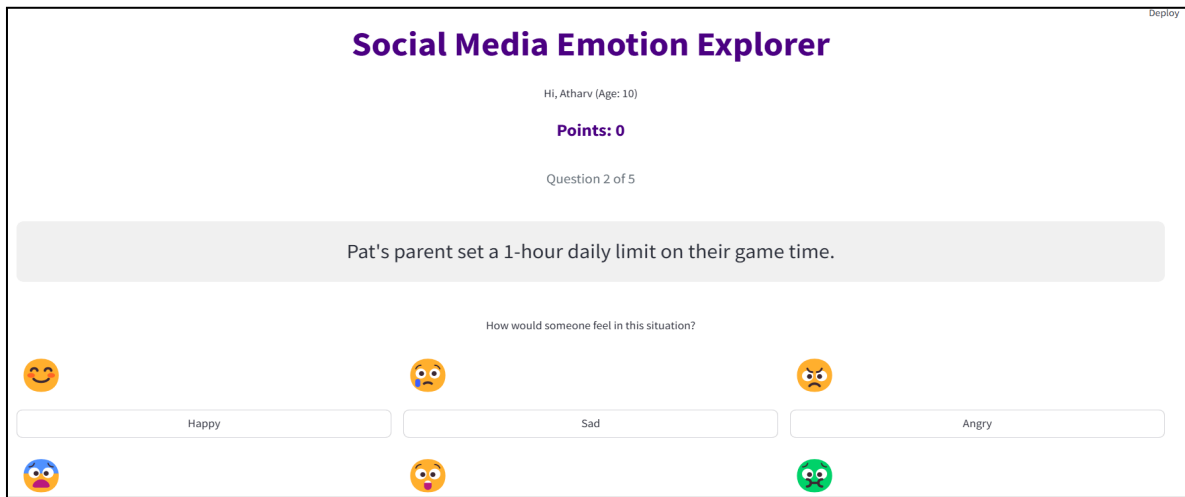
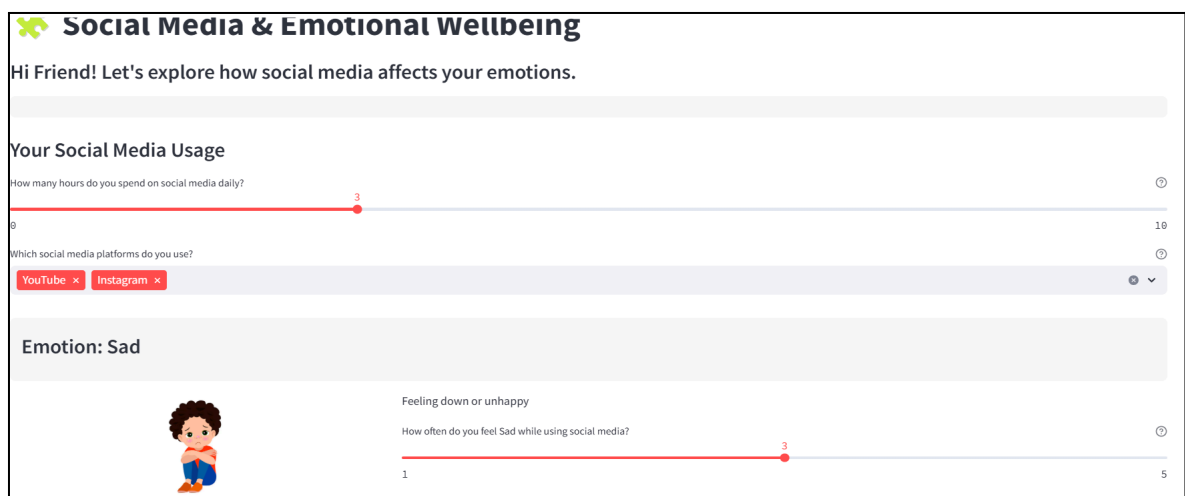


Fig 7.6 : Report content showing screentime of child



**Fig 7.7 : Screenshot of Game activity 1**



**Fig 7.8 : Screenshot of Game activity 2**

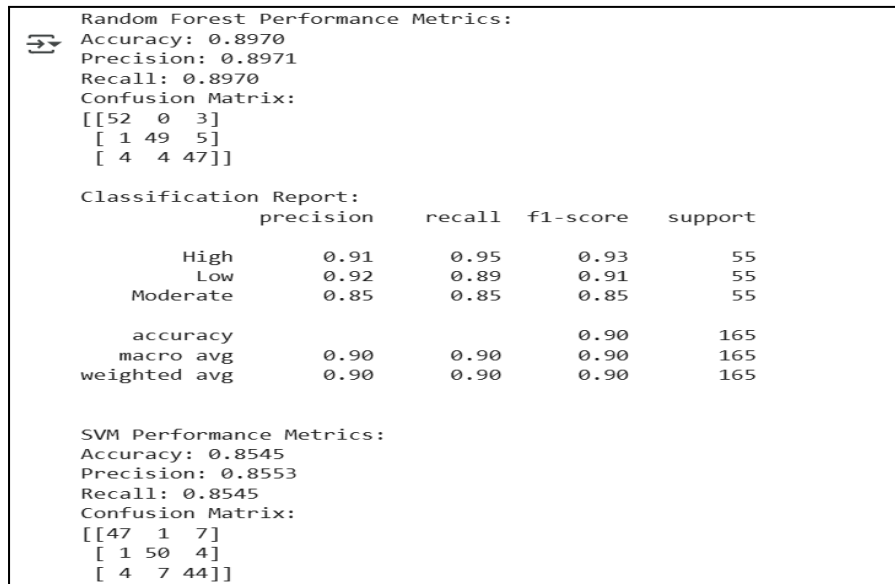
## 7.2. Performance Evaluation Measures:

- 1. Precision:** Precision is one indicator of a machine learning model's performance – the quality of a positive prediction made by the model. Precision refers to the number of true positives divided by the total number of positive predictions (i.e., the number of true positives plus the number of false positives).
- 2. Recall:** The recall is calculated as the ratio between the numbers of Positive samples correctly classified as Positive to the total number of Positive samples. The recall measures the model's ability to detect positive samples. The higher the recall, the more positive samples detected.
- 3. F-Score:** The F-score (also known as the F1 score or F-measure) is a metric used to evaluate the performance of a Machine Learning model. It combines precision and recall into a single score.

### Formula for evaluation metrics:

1. Precision =  $TP / (TP + FP)$
2. Accuracy =  $(TP + TN) / (TP + TN + FP + FN)$
3. Recall =  $TP / (TP + FN)$

Where, TP = True Positives, FP = False Positives, TN = True Negatives, FN = False Negatives



Random Forest Performance Metrics:

Accuracy: 0.8970  
Precision: 0.8971  
Recall: 0.8970  
Confusion Matrix:  
[[52 0 3]  
[ 1 49 5]  
[ 4 4 47]]

Classification Report:

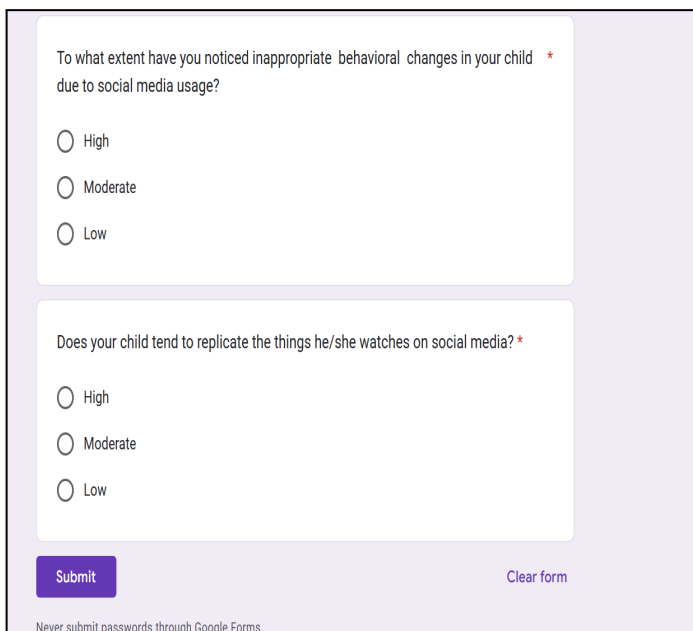
	precision	recall	f1-score	support
High	0.91	0.95	0.93	55
Low	0.92	0.89	0.91	55
Moderate	0.85	0.85	0.85	55
accuracy			0.90	165
macro avg	0.90	0.90	0.90	165
weighted avg	0.90	0.90	0.90	165

SVM Performance Metrics:

Accuracy: 0.8545  
Precision: 0.8553  
Recall: 0.8545  
Confusion Matrix:  
[[47 1 7]  
[ 1 50 4]  
[ 4 7 44]]

**Fig 7.9 : Precision, Recall, Accuracy of various models**

## 7.3. Input Parameters/Features considered:



To what extent have you noticed inappropriate behavioral changes in your child due to social media usage? \*

☐ High  
☐ Moderate  
☐ Low

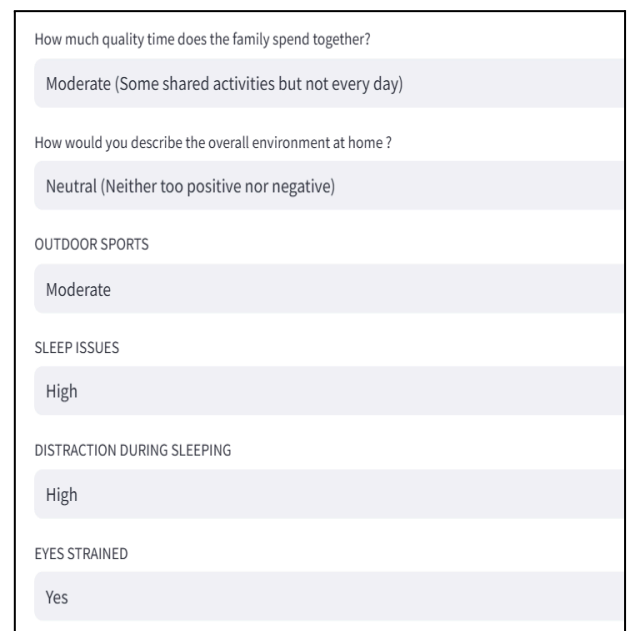
Does your child tend to replicate the things he/she watches on social media? \*

☐ High  
☐ Moderate  
☐ Low

[Clear form](#)

Never submit passwords through Google Forms.

**Fig 7.10 : Google form inputs**



How much quality time does the family spend together?

Moderate (Some shared activities but not every day)

How would you describe the overall environment at home ?

Neutral (Neither too positive nor negative)

OUTDOOR SPORTS

Moderate

SLEEP ISSUES

High

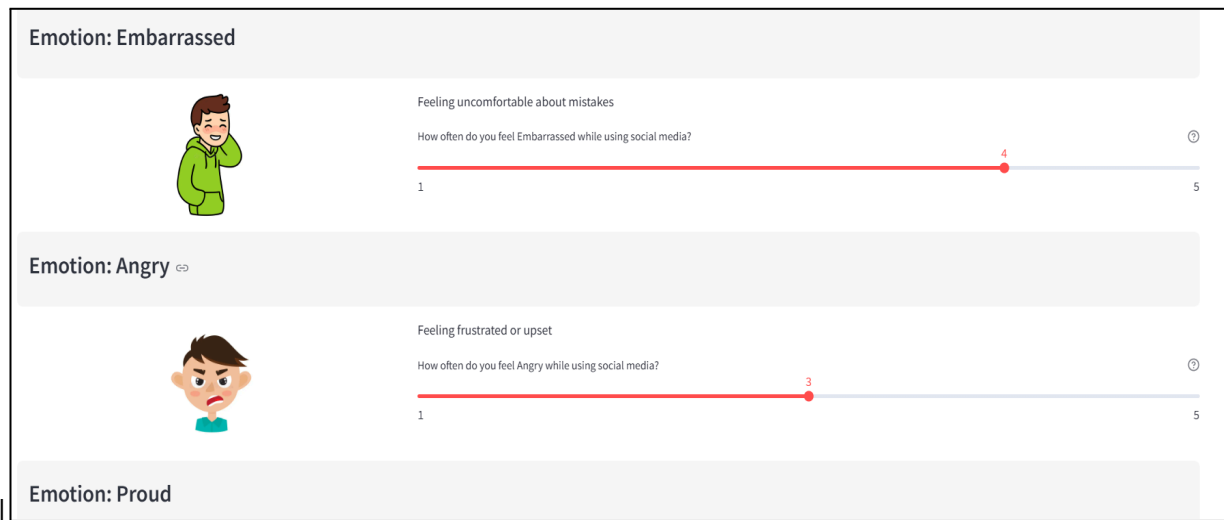
DISTRACTION DURING SLEEPING

High

EYES STRAINED

Yes

**Fig 7.11 : Input parameters for model**



**Fig 7.12 : Input parameters for game based activity**

## 7.4. Comparison of Results with Existing System:

Other Systems	Our System
Focus only on mental health detection or social media analytics	Focuses on both social media behavior and mental health analysis
Mostly do not offer personalized recommendations	Provides personalized mental health recommendations via chatbot
Lack real-time interaction or user feedback integration	Includes an intelligent chatbot for real-time mental health support
Often do not consider screen time or blue light exposure in predictions	Incorporates screen time patterns and usage behavior in analysis

**Table 7.1: Other systems v/s Our systems**

## 7.5. Inference Drawn:

- Most existing systems focus on either mental health detection or social media analysis, but few integrate both for a holistic solution.
- There is a growing emphasis on using AI techniques, especially machine learning and deep learning, for early mental health prediction.
- Transformer-based models like BERT consistently outperform traditional ML models in analyzing sentiment and detecting depression from social media data.
- Studies show that excessive screen time and blue light exposure are linked to negative outcomes like anxiety, sleep disruption, and cognitive decline.
- Inspired by recent research on chatbot applications, our project integrates a personalized chatbot to offer real-time mental health recommendations based on user behavior.

# Chapter 8: Conclusion

## 8.1.Limitations:

- **Data Quality and Availability:** The accuracy of machine learning models heavily depends on the quality, quantity, and diversity of data.
- **Subjectivity of Mental Health Indicators:** Emotions and mental health conditions are inherently subjective and complex.
- **Privacy and Ethical Concerns:** Handling personal and psychological data requires strict adherence to data privacy regulations and ethical standards.
- **Intervention Limitations:** The project has a chatbot provision for suggestions but cannot substitute professional medical consultation or therapy.

## 8.2.Conclusion:

This research investigates how social media usage by children influences their mental well-being using machine learning methods to analyze psychological measures, screen time, content consumption, and behavioral patterns. Through effective identification of significant factors leading to mental health threats such as anxiety, depression, and changes in behavior, the research allows for the utilization of the Random Forest model in precise classification [10]. A practical remedy for real-time observation is also provided by the development of an interactive Power BI dashboard and an AI-driven chatbot.

To promote children's digital well-being, this research highlights the importance of responsible social media use [8]. Moreover, a better understanding of social media's effects on mental health is facilitated through the integration of data-driven findings and interactive features, allowing parents and teachers to act preventively. This research paves the way for earlier intervention methods to reduce the harmful psychological impacts of overuse of social media through the identification of high-risk patterns of behavior and providing personalized advice. Finally, this research lays the foundation for raising awareness and encouraging a balanced digital environment while helping children and adolescents use social media in moderation.



### 8.3.Future Scope:

- **Integration with Real-Time Monitoring Tools:** The project can further be extended by integrating it into mobile or web-based applications that monitor real-time screen use, social media usage, and behavioral changes, allowing for dynamic and ongoing monitoring of mental health metrics.
- **Advanced Personalization with Deep Learning:** Subsequent versions can utilize deep learning algorithms such as LSTMs or transformers to examine sequential behavior data, enabling strongly personalized mental health predictions and interventions according to personal usage patterns.
- **Collaboration with Mental Health Professionals:** The system can become a clinical decision support tool by collaborating with psychologists and psychiatrists, assisting in validating model predictions and adapting mental health programs for early intervention.
- **Language and Region Expansion:** With the ability to support multiple languages and catering to cultural differences, the system can be scaled internationally to address diverse populations, particularly in under-served regions that have a shortage of mental health infrastructure.

# References

- [1] A.R. Primack, A.J. Shensa, H.Sidani, et al., 2017 “Social Media Use and and Perceived Social Isolation Among Young Adults in the U.S.” America Journal of Preventive Medicine, vol. 53
- [2] T.J.R Naab, 2020, “The effects of social media on mental health: A systematic review,” Journal of Mental Health, vol. 29, no.5, pp. 543-550. doi: 10.1080
- [3] K.E.M. Becker, 2019, “Understanding the impact of social media on mental health: A review,” Journal of Social Media in Society, vol.8, no.1, pp. 130-145 [Online]. Available: <https://www.j-sms.org/8-1/>
- [4] P.A.H Cheng, C.T.N. Wong and S.M.Y. Lee, 2022, “The relationship between social media use and mental health outcomes in adolescents: A systematic review,” BMC Public Health, vol.22, no.1, pp.1-9. doi: 10.11
- [5] D.D. Twenge, 2017, “Have Smartphones Destroyed a Generation?” The Atlantic, vol.320, no.3, pp.1-14. [Online]. Available: <https://www.theatlantic.com/magazine/archive/2017/09/has-smartphone-destroyed-a-generation/534198/>.
- [6] A.P.H Primack, D.H. Shensa and H. Sidani, 2021, “Social Media Use and Mental Health Outcomes: A Review of the Literature,” Current Opinion in Psychology, vol.39, pp. 67-71. doi:10.1016 /j.copsyc.2020.10.003
- [7] Vengalarao Pachava, Krishna Golla, 2024, “Machine Learning Analysis of Social Media’s Impact on mental health of Indian youth”, International Research Journal of Multidisciplinary Scope 5(2): 623-635
- [8] Biodoumoye George Bokolo, Qingzhong Liu, 2023, “Deep Learning-Based Depression Detection from Social Media: Comparative Evaluation of ML and Transformer Techniques”, 12(21): 4396 .
- [9] Nurliyana Abas, Hanani Hussain, 2023, “Biodoumoye George Bokolo, Qingzhong Liu, 2023, “Exploring the interconnection of social media, mental health and youth: A bibliometric analysis”, Social and Management Research Journal 20(2) 2023, Research Journal 185 – 206.
- [10] Veerpal Kaur, Aadrita Nandy, Jyoti Chaudhary, 2023, “Machine Learning for Early Detection of Child Depression: A Data Driven Approach”, IEEE 2<sup>nd</sup> International Conference.
- [11] Betul Keles, Annmarie Grealish, 2023, “A systematic review: The influence of Social Media on depression, anxiety and psychological distress in adolescents”, International Journal of Adolescence and Youth, Volume 25

- [12] Sriteja Kataru, Kathleen King, 2024, "Machine Learning Based Early Detection and Intervention for Mental Health Issues in Children", IEEE 48<sup>th</sup> Annual Computers, Software and Application Conference.
- [13] Deepali Joshi, Mansi Patwardhan, 2020, "An analysis of mental health of social media users using unsupervised approach", Computers in Human Behaviour Reports, Volume 2.
- [14] Vilas Sawrikar , Arshad Hussain, 2023, "A hybrid mental health prediction model using Support Vector Machine, Multilayer Perceptron, and Random Forest algorithms", Healthcare Analytics, Volume 3.
- [15] Konda Vaishnavi, U Nikhitha Kamath, 2021, " Predicting Mental Health Illness using Machine Learning Algorithms", Machine Learning Conference, Volume 12.
- [16] Kaushik Chanda, Sandip Roy, 2022, " Evaluating Mental Health Issues as Collected from Social Media by Machine Learning", Department of Computational Science, Brainware University.
- [17] Tarun Jain, Ashish Jain, Priyank Singh, 2021, "Machine Learning Techniques for Prediction of Mental Health", 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)
- [18] Akshatha N K, Dr. Madhwaraj K G, 2024, Mental Health Prediction of Children Addicted to Digital Platforms, International Journal of Multidisciplinary Research, Volume 7.

# Appendix

## Impact of Social Media on Mental Health using Machine Learning

**Mrs. Mannat Doultani<sup>§</sup>, Aryan Manghi<sup>#</sup>, Prasad Chaudhari<sup>#</sup>, Devyaansh Razdan<sup>#</sup>**

**§:** Assistant Professor, Computer Engineering Department , Vivekanand Education Society's Institute of Technology, Mumbai

**#:** Student in Computer Engg. Vivekanand Education Society's Institute of Technology, Mumbai

### **Abstract:**

Social media is a major source of information as well as it profoundly affects a person's personal life. While overuse might lead to more anxiety or mental health issues, less frequent use could foster a feeling of community or social support. Examining social media and mental health survey datasets through the lens of machine learning (ML), the project "Impact of Social Media on Mental Health using ML" tries to explore a phenomenon of social media use and its consequences to mental wellness. By means of trend and relationship detection, the project seeks to grasp what particular social media behaviors: do with a person's wellbeing. Developing an intelligent chatbot that provides tailored recommendations on a user's mental health is a key component of the project. The machine learning-based chatbot will provide contextual advice by analyzing user data and mood markers.

### **Keywords:**

**Social media, Mental Health, Machine Learning, Personalized recommendations.**

### **Introduction:**

Both for adults and children and youths, social media spaces are beneficial in many ways, while often posing serious challenges to their mental wellbeing [1]. In younger children, the growing tide of depression, anxiety and sociability withdrawal is believed to come from the constant demand to project a perfect online self, curated content and media, and the addictiveness of the social media itself. This project titled "The Digital Mindscape: Leveraging Machine Learning to Understand The article "Social Media's Effects on Child Mental Health" aims to clarify this issue.

We aim to better define the negative effects of social media on children's and adolescents' mental health as well as the apparent positive correlations [2]. In order to identify the main contributing factors to social media mental health disorders, we will measure the length of time children spend online, the type of content they consume, and how they react to it.

In addition to identifying the mental health risks associated with excessive use of social media, this project aims to develop workable solutions to these issues [7]. We hope to develop tailored suggestions that help kids and adopt better online practices by examining each user's unique social media usage patterns. We intend to provide customized guidance that can assist families in better understanding how social media affects mental health by utilizing cutting-edge technologies like machine learning.

The knowledge gathered will be used as the basis for wellness initiatives and educational initiatives that stress the value of balancing digital platform use. We hope to develop a framework that tackles the current problems and promotes a healthy digital environment for coming generations by combining data-driven research with community-based initiatives [18].

This project's main goal is to equip adults and kids with the information and tools they need to use social media responsibly. We aim to encourage healthier behaviors and lessen the rising concerns about mental health issues brought on by excessive or unhealthy social media use by raising awareness and offering workable solutions [13]. We can create a digital environment that fosters well-being for all people by working together.

## **Literature review:**

### **1. Connections between children's well-being and screen time:**

In order to assess the effects of screen time on well-being metrics such as self-worth, and emotional well-being, Jean Twenge carefully examines the data collected from a sample of the population in their report [6]. The important conclusion is that extended screen time is associated with worse psychological health, especially when it comes to social satisfaction and emotional well-being.

**2. Deep learning based social media depression detection:** This study shows how well deep learning works to identify depression in social media posts [10]. The study demonstrates BERT's superior performance in detecting depression from user-generated text by contrasting transformer-based models with conventional machine learning models.

**3. The effects of blue light on young adults' and adolescents' sleep, productivity and general well-being:** The effects of blue light exposure (from screens and gadgets) on young adults' sleep patterns, cognitive abilities, and general well-being are examined in this systematic review [17]. It summarizes research on the effects of blue light on circadian cycles and mental health.

4. Sentiment analysis of social media data for identifying depression using AI: This review discusses how sentiment analysis and AI techniques can be used to detect depression in social media posts. The authors evaluate various AI algorithms, for analyzing user sentiment and emotions in text data [18]. Sentiment analysis using AI offers promising results in detecting depression from social media data, enabling the early identification of mental health issues.

5. The effect of social media within adolescents and young adults: This review concentrates on the effect of social media on the health of the youth as well as teenagers. It explains that social media use can exacerbate negative feelings because of excessive use [17]. There are mental health problems such as sadness, anxiety, and isolation which have a clear association with social media use among the youth and adolescents.

6. Mental Health inequalities and contributing factors among Indian youth: The study investigates mental health disparities among Indian youth, focusing on the impact of external validation through social. Social media significantly influences mental health outcomes for Indian youth [1,9], with external validation playing a crucial role.

7. Investigating the link between social media, mental health and young people: A case study

This research explores the relationship between social media usage and mental health issues among young individuals [12]. The analysis identifies important factors that contribute to social media-related mental health issues. Results point to a growing focus on comprehending how social media affects young people's wellbeing.

8. Early detection of mental health issues in children using ML: In this study, ML models are being used to detect mental health issues in children early on [3,7]. Finding patterns in social media use and other pertinent variables linked to early signs of mental health issues is the main goal of the study.

9. Screen Time and brain development in Preschoolers: The effect of screen-based media on preschool-aged children's white matter structure is investigated in this study [1]. The study investigates the connection between media consumption and brain development using neuroimaging methods. The findings show that increased screen time is linked to decreased white matter integrity, which may have an impact on young children's literacy development.

## Methodology:

Machine learning processes begin with social media data analysis to detect behavioral patterns and risk features for mental disorders such as depression, anxiety, and addiction [5,14]. The information will be sourced from social media surveys and questionnaires, along with the level of screen time users spend watching media, as well as the consumption and activity frequency. This diverse data set will help to clarify the relation between different social media activities and mental health issues. Data cleaning, which is part of the preprocessing step, consists of addressing missing values and duplicates, in addition to feature engineering where meaningful attributes like average daily screen time and exposure frequency to different types of content (negative, news, and comparison-based) are constructed. Following data preparation, it will undergo annotation to train the ML models appropriately.

The project scope has both a supervised and unsupervised aspect: a randomized forest along with a support vector machine will carry out classification and prediction of mental health conditions in correspondence with social media activity [15]. K-means algorithms will detect other clusters of users characterized with similar behaviors. Also, some correlations will be analyzed, and sentiment analysis is going to be included to reveal nuances.

Formula for evaluation metrics:

1. Precision =  $TP / (TP + FP)$
2. Accuracy =  $(TP + TN) / (TP + TN + FP + FN)$
3. Recall =  $TP / (TP + FN)$

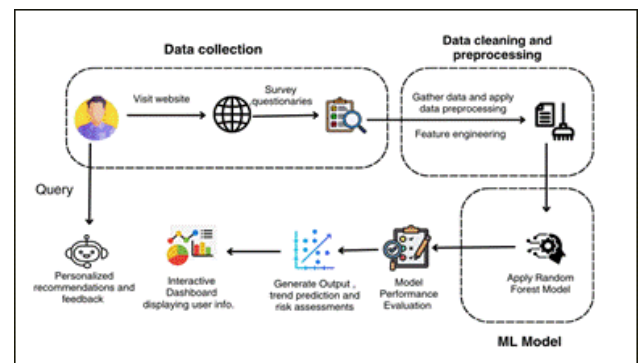
where,

TP = True Positives

FP = False Positives

TN = True Negatives

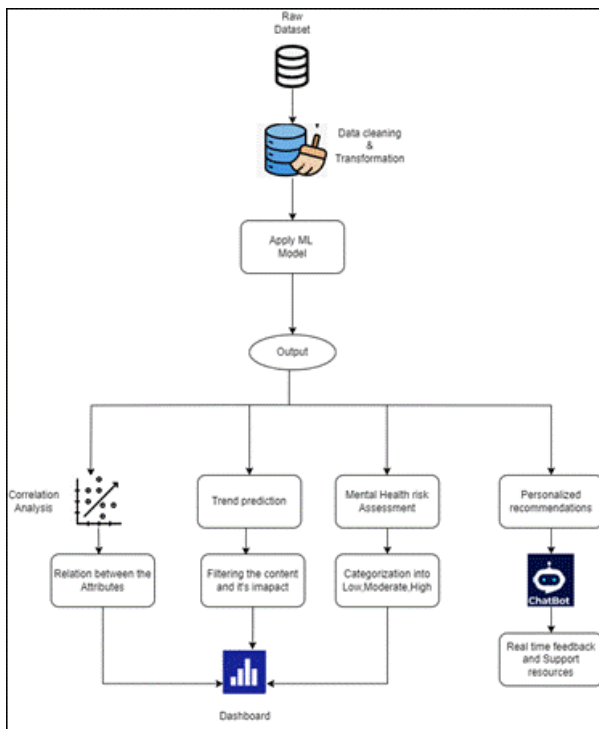
FN = False Negatives



**Fig.1 Modular Diagram**

In the provided diagram, the system is broken down into the following modules:

1. Data collection: Gather data from various sources such as website & surveys
2. Data cleaning & processing: To clean & prepare data for analysis.
3. Feature Engineering: Create new features from existing data.
4. ML Model: Random Forest is applied.
5. Model performance evaluation: To assess accuracy & effectiveness.
6. Generate output: Final result such as trend prediction, risk assessment
7. Interactive dashboard: Explore data & insights.



**Fig.2 Block Diagram**

This research adopts a methodical strategy to examine the influence of social media on children's mental health by utilizing machine learning methods [15]. The approach adheres to an organized sequence, commencing with the gathering of data, its preprocessing, and the application of a machine learning model to derive insights. The objective is to evaluate mental health risks and offer tailored recommendations.

The research involves multiple analytical processes, including correlation analysis, trend prediction [11], and personalized recommendations through a chatbot. A Random Forest model is applied to predict depression levels, categorizing users into Low, Moderate, or High risk.

## 1. Data Collection:

The dataset is collected from surveys, interviews, or existing mental health research databases [9] . It includes 26 key attributes, such as age, gender, social media usage, screen time, etc.

2. Data Cleaning & Transformation: To maintain high data quality, missing values are addressed through imputation methods, such as using the mean, for numerical data, while categorical data is filled in with the most common category [14].

3. Machine Learning Model Application: The Random Forest model is chosen because of its excellent accuracy in classification tasks [2,4], making it suitable for predicting mental health risk levels. To enhance its performance, hyperparameter tuning is carried out.

4. Model Output & Analysis: The study includes correlation analysis to identify relationships between attributes like excessive screen time and anxiety. Trend prediction analyzes content impact on mental health and tracks behavioral changes. Mental health risk assessment categorizes individuals into Low, Moderate [8,10], or High-risk levels for early detection. A chatbot provides personalized recommendations, suggesting reduced screen time.



## References:

- [1] A.R. Primack, A.J. Shensa, H.Sidani, et al., 2017 “Social Media Use and and Perceived Social Isolation Among Young Adults in the U.S.” *America Journal of Preventive Medicine*, vol. 53
- [2] T.J.R Naab, 2020, “The effects of social media on mental health: A systematic review,” *Journal of Mental Health*, vol. 29, no.5, pp. 543-550. doi: 10.1080
- [3] K.E.M. Becker, 2019, “Understanding the impact of social media on mental health: A review,” *Journal of Social Media in Society*, vol.8, no.1, pp. 130-145 [Online]. Available: <https://www.j-sms.org/8-1/>
- [4] P.A.H Cheng, C.T.N. Wong and S.M.Y. Lee, 2022, “The relationship between social media use and mental health outcomes in adolescents: A systematic review,” *BMC Public Health*, vol.22, no.1, pp.1-9. doi: 10.11
- [5] D.D. Twenge, 2017, “Have Smartphones Destroyed a Generation?” *The Atlantic*, vol.320, no.3, pp.1-14. [Online]. Available: <https://www.theatlantic.com/magazine/archive/2017/09/has-smartphone-destroyed-a-generation/534198/>.
- [6] A.P.H Primack, D.H. Shensa and H. Sidani, 2021, “Social Media Use and Mental Health Outcomes: A Review of the Literature,” *Current Opinion in Psychology*, vol.39, pp. 67-71. doi:10.1016 /j.copsyc.2020.10.003
- [7] Vengalarao Pachava, Krishna Golla, 2024, “Machine Learning Analysis of Social Media’s Impact on mental health of Indian youth”, *International Research Journal of Multidisciplinary Scope* 5(2): 623-635
- [8] Biodoumoye George Bokolo, Qingzhong Liu, 2023, “Deep Learning-Based Depression Detection from Social Media: Comparative Evaluation of ML and Transformer Techniques”, 12(21): 4396 .
- [9] Nurliyana Abas, Hanani Hussain, 2023, “Biodoumoye George Bokolo, Qingzhong Liu, 2023, “Exploring the interconnection of social media, mental health and youth: A bibliometric analysis”, *Social and Management Research Journal* 20(2) 2023, Research Journal 185 – 206.
- [10] Veerpal Kaur, Aadrita Nandy, Jyoti Chaudhary, 2023, “Machine Learning for Early Detection of Child Depression: A Data Driven Approach”, *IEEE 2<sup>nd</sup> International Conference*.
- [11] Betul Keles, Annmarie Grealish, 2023, “A systematic review: The influence of Social Media on depression, anxiety and psychological distress in adolescents”, *International Journal of Adolescence and Youth*, Volume 25
- [12] Sriteja Kataru, Kathleen King, 2024, “Machine Learning Based Early Detection and Intervention for Mental Health Issues in Children”, *IEEE 48<sup>th</sup> Annual Computers, Software and Application Conference*.

# Plagiarism Report:

Impact of Social Media on Mental Health using Machine Learning			
ORIGINALITY REPORT			
6% SIMILARITY INDEX	5% INTERNET SOURCES	4% PUBLICATIONS	2% STUDENT PAPERS
PRIMARY SOURCES			
1	pdfs.semanticscholar.org Internet Source	2%	
2	www.ijcaonline.org Internet Source	1%	
3	www.mdpi.com Internet Source	1%	
4	H L Gururaj, Francesco Flammini, V Ravi Kumar, N S Prema. "Recent Trends in Healthcare Innovation", CRC Press, 2025 Publication	1%	

# Paper Submission:

7th International Conference on ADVANCES IN SCIENCE & TECHNOLOGY : Submission (130) has been created. Inbox x

Microsoft CMT <email@msr-cmt.org>  
to me ▾

12:19 PM (0 minutes ago) ☆ (

Hello,

The following submission has been created.

Track Name: Track I: Computer Sciences and Recent Trends

Paper ID: 130

Paper Title: Impact of Social Media on Mental Health using Machine Learning

# Review 1 Sheet:

Inhouse/ Industry \_Innovation/Research:

## Project Evaluation Sheet 2024 - 25

Class: D17/2025

Group No.: 35

Sustainable Goal:

Title of Project: The Impact of Social Media on Mental Health

Group Members: Prasad Chaudhari D17B08, Aryan Manghi D17B28, Devyansh Rastan D17B45

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
3	4	3	2	4	2	2	2	2	2	2	3	3	2	3	39

Comments: Consider more parameters & generate appropriate reports. Manu  
Name & Signature Reviewer1

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
3	4	3	2	4	2	2	2	2	2	2	3	3	3	3	40

Comments:

Date: 1st March, 2025

Abha Tewari  
Name & Signature Reviewer 2

## Review 2 Sheet:

Inhouse/ Industry\_Innovation/Research:

Sustainable Goal:

### Project Evaluation Sheet 2024 - 25

Class: D17 A(B)C

Group No.: 35

Title of Project: Impact of social media on Mental Health

Group Members: ~~Hitesh Kadam~~ Aryan Manghi (D17B/28), Prasad Chaudhari (D17B/8), Devyansh Razdan (D17B/45)

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
4	4	5	3	5	2	2	2	2	2	3	3	2	2	3	44

Comments:

Name & Signature Reviewer1

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
4	4	5	3	5	2	2	2	2	2	3	3	2	2	3	44

Comments:

Date: 1st April, 2025

P-10/30  
A-30

Name & Signature Reviewer 2

