ACE ACUMEN ACADEMY

Analysis Of Social Media and Sleep Disorders on Mental

by

Vishal Suman – W08319 83, Amneet Kaur – W0852647,

Raj Kumar – W0850979, Kulpreet Singh – W0835917.

Capstone Project

Submitted To Prof. Adegoke Ojeniyi

St. Clair College, Ace Acumen Academy

In Fulfilment Of The Requirements

Data Analytics For Business

December 2024

ABSTRACT

This research explores the complex relationships between social media usage, sleep disturbances, and mental health outcomes, focusing on adolescents and young adults—groups increasingly affected by digital engagement. As social media becomes a daily routine, concerns about its adverse effects on sleep quality, mental health, and overall well-being have grown. This study employs advanced machine learning algorithms, including Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, and Random Forest (RF), to analyse a dataset of 10,000 individuals, capturing variables such as daily technology usage, sleep duration, physical activity, and self-reported mental health indicators. Findings reveal significant correlations between high social media engagement, reduced sleep quality, and elevated stress and anxiety, with substantial generational differences across age groups. These differences suggest that younger individuals, such as Gen Z and Millennials, may be particularly vulnerable to the effects of excessive screen time, which disrupts sleep and contributes to poorer mental health outcomes. Highlighting the need for targeted public health strategies, this study underscores the importance of promoting balanced social media use, better sleep hygiene, and awareness of the potential mental health risks associated with prolonged online engagement. This research contributes to digital health literature by providing actionable insights for policymakers, healthcare professionals, and educators to foster healthier digital habits. Emphasizing the importance of balanced technology use among high-risk groups, these findings offer a foundation for developing effective intervention programs aimed at supporting mental well-being, underscoring the need for collaborative efforts to encourage healthier engagement with technology.

Keywords: social media, sleep disorders, mental health, machine learning, adolescents, young adults, digital health

ACKNOWLEDGMENTS

We would like to extend our heartfelt gratitude to our professor, advisor, and guide, Prof. Adegoke Ojeniyi, for his unwavering guidance and support throughout this capstone project. His expertise, insights, and encouragement were instrumental in the successful completion of this work, and we are deeply appreciative of his dedication to helping us achieve our project objectives.

We also wish to acknowledge our friends and classmates for their invaluable support and encouragement, which provided us with the strength and motivation to persevere through challenges. Their belief in our potential played a vital role in making this accomplishment possible.

We proudly present this capstone project as a representation of our academic journey and the collective efforts of everyone who contributed to its success. This project marks the completion of our studies in the Data Analytics program, finished in December 2024.

Thank you to all who supported us along the way.

Project Contributors: Vishal Suman, Amneet Kaur, Raj Kumar, Kulpreet Singh

TABLE OF CONTENTS

Table of Figures

CHAPTER ONE
INTRODUCTION

Introduction

In the last ten years social networks have become an essential part of people's lives especially youths and young adults. Even though these technologies provide novel opportunities for social interaction and individual identity, their constant use has been associated with many health consequences. More specifically, the social media usage, sleep patterns, and mental health are interconnected in various ways, which are detailed next. (Yu, (2024))
The purpose of this current research proposal is to assess the existing literature and examine the correlation between the use of social media and sleep and mental health consequences. They will analyse and compare frequency, duration, and timing of social media use and presence of sleep disorders. The findings of this report will seek to discuss how sleep disruptions caused by social media can influence mental disorders such as anxiety and depression. Furthermore, we shall assess the effectiveness of interventions to reduce the social media usage at night to enhance sleep and mental health.
It is important to comprehend these complex relations because sleep and mental health are two interrelated components of a human's health. Stress has been known to worsen mental health conditions and mental health disorders play a role in sleep disorders..
In doing so, it aims to foster a better understanding of and guide practice and policy in relation to healthier social media use and sleep and mental health, especially for youth. The next sections of the work will present the background information concerning the given topic, describe the methods used during this study, state the main results of the investigation, discuss the ongoing issues that are crucial for this field of research and suggest possible further developments. (Cain & Gradisar, 2010)

Figure 1 Digital Exhaustion: The Toll of Social Media on Sleep and Mental Health

1.1 Background of the study

Today, social media has taken over the way people interact and communicate in their day-to-day basis especially among the youth. Social sites such as Face book, Instagram, snapchat and twitter have become almost a necessity in the lives of many individuals in the last ten years. The penetration of social media into the daily lives of individuals has not only changed social behaviours but also influenced different spheres of human health starting with psychological and including the amount and quality of sleep.

Social media is full of content and every user's constant companion and entertainment, and that is why people stay online and active in the wee hours. Research that links the time children spend on digital platforms and screens before they sleep is damaging their circadian rhythm; a natural timer that is responsible for controlling the rhythm sleep. This disruption is related to decrease in sleep quality and decrease in quantity of sleep which has serious consequences for mental condition. (Yu, (2024))

The relationship between social media use and sleep disorders and mental health is not fully understood. In one perspective, people can use the SNS to escape from the stress experienced throughout the day since it's an easy way of diverting their attention. On the other hand, late night social media activities are likely to make a cycle; since poor sleep is likely to lead to anxiety, depression and increase the tendency of engaging in social media more often.

New approaches are now used to investigate situational characteristics of social media uses with reference to the social, emotional and cognitive perspectives rather than have a generalized view of quantized Time on Device. This shift further acknowledges that necessary changes exist beyond the literature convenience argument which social media brings this aspect through altering the people's interaction. (Scott & Cleland Woods, 2019)

1.2 Statement of the Problem

Increased use of social media, especially among youths and young adults, have raised some issues about the effects this has on sleep and mental health. Research has it that many teenage persons get less than the required eight hours of sleep at night and this is attributed to time spent on social media sites. (Yu, (2024)).This raises concern and requires a determination of the first and second-order effects in response to this behaviour.

Logarithmic sleep disorders for example, insomnia is socially manifested in the current youth population due to late night social media activity. Besides, lack of sleep quality has also been directly associated with the emergence of psychological disorders such as depression and anxiety. Despite the growing body of evidence, there remains a lack of comprehensive studies that fully explore the bidirectional nature of this relationship: Is that to mean that the use of social media leads to poor sleep or is it just a section of the population who are already experiencing sleep difficulties get addicted to social media as a way of compensating for their sleepless nights? (Seabrook, 2016)

The sleep and mental health research literatures have made recent advances towards a more nuanced understanding of social media use: shift from the traditional emphasis of the 'time spent' in using online media to understanding content, context and experiences of interactions. (Scott & Cleland Woods, 2019). However, the researchers in these fields have been constrained by the available research measures. Prominent features include the application of single-item, non-validated measures, often reflecting duration or frequency of use, while existing, multi-items validated measures are lacking in generalizability and risk entailing pathologization of social media use. (Seabrook, 2016)

1.3 Study Aim and Objectives

This study aims to investigate the complex relationships between social media use, sleep disorders, and mental health, with a particular focus on adolescents and young adults.

The specific objectives of the study are:

1.3.1 To evaluate existing measurement tools for social media use and develop improved measures that capture relevant experiences beyond just frequency and duration of use.

1.3.2 To assess the correlation between high social media usage and low sleep hours with mental health outcomes such as stress and anxiety.

1.3.3 To build a predictive machine learning model to estimate stress levels based on technology usage patterns, sleep hours, and physical activity levels.

1.3.4 To compare the performance of different machine learning models (k-means, RNN, LSTM and RF) in predicting and analysing the relationships between social media use, sleep disorders, and mental health outcomes.

1.3.5 To propose strategies for reducing social media usage and increasing physical activity to improve mental health and sleep quality among young adults.

1.3.6 To analyse how technology usage patterns and stress levels vary across different generations (The Silent Generation, Baby Boomers, Gen X, Millennials, Gen Z, and Gen Alpha).

1.3.7 To examine how the correlation between high social media usage and mental health outcomes (e.g., anxiety, depression) differs between generations.

1.3.8 To develop machine learning models that predict stress levels for each generation, taking into account technology usage patterns, sleep hours, and physical activity levels.

1.3.9 To determine the optimal balance of social media usage and physical activity that can improve mental health outcomes across different generations.

1.3.10 To use machine learning models to predict mental health outcomes based on social media usage and demographic factors.

1.4 Study Questions and/or Hypotheses

The following research questions guide this study:

1.4.1 How can we determine the average daily social media usage time among young adults in the sample, and how does it compare to their overall screen time?

1.4.2 What are the underlying factors and mechanisms that explain why individuals experience high stress levels in relation to their technology usage patterns and work environment impact?

1.4.3 What are the underlying reasons that explain why the combination of high social media usage and low sleep hours leads to poor mental health status?

1.4.4 How can a machine learning model accurately predict an individual's stress level based on their technology usage patterns, sleep hours, and physical activity?

1.4.5 How can we determine the optimal balance of reduced social media usage and increased physical activity to improve mental health outcomes and sleep quality for young adults?

1.4.6 How do technology usage patterns and stress levels vary across generations (The Silent Generation, Baby Boomers, Gen X, Millennials, Gen Z, and Gen Alpha)?

1.4.7 What are the underlying factors that explain why the correlation between high social media usage and mental health outcomes (e.g., anxiety, depression) differs between generations?

1.4.8 How can we build machine learning models that predict stress levels for each generation, factoring in technology usage, sleep hours, and physical activity?

1.4.9 What is the optimal balance of social media usage and physical activity for improving mental health across different generations?

1.4.10 How accurately can machine learning models predict mental health outcomes based on social media usage patterns and demographic factors?

1.5 Significance of the Study

This study is significant for several reasons:

1.5.1 It addresses an urgent public health issue: the growing impact of social media use on the mental and physical well-being of adolescents and young adults. As social media continues to play a central role in the lives of young people, understanding its effects on sleep and mental health becomes increasingly critical.

1.5.2 The findings have the potential to inform public health campaigns and interventions designed to promote healthier social media habits. By identifying key patterns in social media usage that contribute to poor sleep and mental health outcomes, this research can guide the development of evidence-based strategies aimed at improving sleep hygiene among young people.

1.5.3 It contributes to the academic literature by providing insights into the mechanisms through which social media use affects mental health and sleep quality. These insights will be valuable for future research and may lead to the development of more comprehensive intervention programs.

1.5.4 The study moves beyond simplistic measures of screen time to examine the nuanced social, emotional and cognitive aspects of social media engagement. This more holistic approach can inform a deeper understanding of how social media use relates to overall health and well-being.

1.5.5 By developing improved measurement tools, the study can enhance the quality and relevance of future research in this rapidly evolving field. (Scott & Cleland Woods, 2019)

1.5.6 The findings can help clinicians, educators, and parents better understand and address the complex relationships between social media use, sleep, and mental health in young people. (Scott & Cleland Woods, 2019)

1.6 Scope of the Study

1.6.1 Population: The target audiences were those of the age 15-30 years as most of them are socially active especially on social media platforms up to 90 years.

1.6.2 undefined social media platforms: Some of the famous social media platforms are Facebook, Instagram, Snapchat, and Twitter among others.

1.6.3 undefined Sleep measures: Subjective information about the quality, duration, and disorders of sleep, combined with objective data where available.

1.6.4 undefined Mental health outcomes: Self-rated questionnaires that assess the level of anxiety, depression, stress, and general wellness.

1.6.5 Social media use measures: Frequency and duration of use, timing of use (particularly nighttime use), content and quality of interactions, motivations for use, and emotional connection to platforms.

1.6.6 Contextual factors: Physical activity levels, work/school environment, availability of support systems.

1.6.7 Timeframe: Cross-sectional data collection, with recommendations for future longitudinal studies.

While the study is limited to a specific age group and relies on both subjective and objective data, it provides critical insights into the relationship between social media usage, sleep disturbances, and mental health. Future research could expand on this by exploring other age groups and investigating long-term effects.

1.7 Operational Definition of Terms

1.7.1 Technology Usage Hours: The total hours spent engaging with various forms of technology, including social media, gaming, and screen-based activities.

1.7.2 Social Media Usage Hours: The specific time spent on social media platforms, such as Facebook, Instagram, and Twitter.

1.7.3 Gaming Hours: The amount of time spent on video games across different platforms.

1.7.4 Screen Time Hours: The cumulative time spent on screens for work, entertainment, or leisure purposes.

1.7.5 Mental Health Status: A measure of psychological well-being, including indicators of anxiety, depression, and emotional stability.

1.7.6 Stress Level: Self-reported or measured levels of psychological or emotional stress.

1.7.7 Sleep Hours: The number of hours an individual sleeps in a typical 24-hour period.

1.7.8 Physical Activity Hours: The amount of time spent engaging in physical exercise or activity.

1.7.9 Support Systems Access: The availability of social, familial, or online support structures to aid in mental and emotional health.

1.7.10 Work Environment Impact: The effect of an individual's work environment on their mental health and well-being.

1.7.11 Social Media Engagement: The nature and quality of interactions on social media platforms, including content shared, responses received, and emotional reactions.

1.7.12 Nighttime Social Media Use: Engagement with social media platforms during typical sleeping hours or immediately before bedtime.

1.7.13 Sleep Quality: A measure of how well an individual sleeps, including factors like ease of falling asleep, number of awakenings, and feeling rested upon waking.

These operational definitions have been developed out of the variables that are identified in the two datasets used in this study. They clearly specify how each variable is defined and measures when discussing the effects of social media use, sleep, and mental health. It is by combining data from both sources that this project will be able to present a better and more elaborate view of the correlations between these variable

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

2.1.1 Defining Social Media, Sleep Disorders, and Mental Health

Social Media

Social media refers to internet-based applications and platforms that allow users to create, share, and exchange content, as well as participate in social networking. (Karaman, 2024)

These platforms facilitate the creation and sharing of information, ideas, interests, and other forms of expression through virtual communities and networks. Social media platforms are characterized by several key features, including user-generated content, user profiles, social connections and networks between users, interactive participation and engagement, and real-time updates and communication. Users create and share their own content, such as photos, videos, and posts, and each user has a profile that displays their personal information, interests, and activity. These platforms allow users to connect with friends, family, and other users, forming networks of relationships. Additionally, users can like, comment, share, and interact with content and other users, fostering a high level of engagement. Real-time updates and instant messaging capabilities enable immediate communication and information sharing. Common examples of social media platforms include Facebook, Instagram, Twitter, LinkedIn, YouTube, and TikTok. The defining characteristic of social media is the ability for users to both consume and produce content, leading to the term "prosumers." (Karaman, 2024)

Common features of social media platforms include user-generated content, user profiles, and the ability to connect with others. Popular examples include Facebook, Instagram, Twitter, and TikTok. A recent survey in Turkey reported that approximately 91% of people aged 16-24 years were internet users, with social networks being the most common purpose at 84.1% (Karaman, 2024)

Sleep Disorders

Sleep disorders are conditions that impair the normal sleep-wake cycle, affecting the quality, timing, and amount of sleep an individual gets. ((2018))The International Classification of Sleep Disorders (ICSD-3) categorizes sleep disorders into six major categories: insomnia, sleep-related breathing disorders, central disorders of hypersomnolence, circadian rhythm sleep-wake disorders, parasomnias, and sleep-related movement disorders (American Academy of Sleep Medicine, 2014). Insomnia involves difficulty falling or staying asleep, while sleep-related breathing disorders, such as obstructive sleep apnea, involve abnormal breathing patterns during sleep. Central disorders of hypersomnolence, including narcolepsy, are characterized by excessive daytime sleepiness. Circadian rhythm sleep-wake disorders occur when there is a misalignment between an individual's internal clock and the external environment. Parasomnias include abnormal behaviours or experiences during sleep, such as sleepwalking or night terrors. Lastly, sleep-related movement disorders, like restless legs syndrome, involve abnormal movements that disrupt sleep.

Common sleep disorders include insomnia (difficulty falling or staying asleep), sleep apnea (pauses in breathing during sleep), restless legs syndrome, and narcolepsy (National Institute of Neurological Disorders and Stroke, 2019). Sleep disorders can have significant impacts on physical health, mental wellbeing, and overall quality of life.

Mental Health

The World Health Organization (2018) defines mental health as "a state of well-being in which an individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and is able to make a contribution to his or her community". Mental health encompasses emotional, psychological, and social well-being, affecting how individuals think, feel, and act (U.S. Department of Health and Human Services, 2020).

Mental health exists on a continuum, ranging from mental wellness to mental illness. Mental illnesses are health conditions involving changes in emotion, thinking, or behaviour (or a combination of these) that are associated with distress and/or problems functioning in social, work, or family activities (American Psychiatric Association, 2018).

Common mental health disorders include depression, anxiety disorders, bipolar disorder, schizophrenia, eating disorders, and substance use disorders. Depression is characterized by persistent feelings of sadness and loss of interest in activities. Anxiety disorders involve excessive fear or worry, while bipolar disorder includes episodes of both mania and depression. Schizophrenia is a severe mental disorder that affects thinking, feeling, and behaviour. Eating disorders, such as anorexia nervosa and bulimia nervosa, involve preoccupation with food, body weight, and shape. Substance use disorders occur when the recurrent use of alcohol or drugs causes significant impairment, such as health problems or failure to meet major responsibilities at work, school, or home.

It's important to note that mental health is not merely the absence of mental illness, but also includes positive aspects of psychological functioning and wellbeing (Keyes, 2005).

2.1.2 History of Social Media, Sleep Disorders, and Mental Health Research

Evolution of Social Media Platforms

The history of social media can be traced back to the early days of the internet, with each decade bringing new innovations and platforms. In the 1970s and 1980s, early online communication began with the first email sent between computers in 1971. This was followed by the introduction of Bulletin Board Systems (BBS) in 1978, which allowed users to communicate and share files. By 1988, Internet Relay Chat (IRC) enabled real-time chat, laying the groundwork for future social interactions online.

The 1990s marked the birth of social networking sites. In 1997, Six Degrees launched as one of the first recognizable social networking sites, allowing users to create profiles and connect with friends. This era also saw the rise of personal blogging platforms like LiveJournal and Blogger in 1999, which popularized the concept of sharing personal thoughts and experiences online. The early 2000s witnessed the rise of major platforms, with Friendster launching in 2002 and quickly gaining over 3 million users. Myspace followed in 2003, becoming the most popular social networking site globally. In 2004, Facebook was launched, initially targeting college students, and in 2006, Twitter introduced the concept of microblogging. https://blog.hubspot.com/marketing/social-media-history

The 2010s saw mobile and visual platforms dominate the social media landscape. Instagram, launched in 2010, focused on photo-sharing and quickly gained popularity. Snapchat, introduced in 2011, brought ephemeral messaging to the forefront, allowing users to send disappearing messages. In 2016, TikTok (originally Musical.ly) gained popularity with its short-form video content. The 2020s have brought emerging trends such as virtual and augmented reality social platforms like Meta's Horizon Worlds and decentralized social networks based on blockchain technology. These advancements reflect the continuous innovation in social media, aiming to enhance user experience while addressing the challenges of privacy and data security.

Development of Sleep Disorder Research

The field of sleep disorder research has evolved significantly over the past century, beginning with foundational studies in the 1920s and 1930s. In 1924, Hans Berger recorded the first human electroencephalogram (EEG), marking a pivotal moment in understanding brain activity during sleep. Nathaniel Kleinman, often referred to as the father of modern sleep research, began systematic studies of sleep and wakefulness in 1929. The discovery of sleep stages and REM sleep in the 1950s and 1960s, particularly the work of Eugene Asterisk and Kleinman in 1953, further advanced the field. By 1968, the first polysomnography recording, which combined EEG, EOG, and EMG, was conducted, providing a comprehensive method to study sleep patterns.

The 1970s marked the establishment of sleep medicine as a distinct field. The first sleep research society was founded in 1970, and in 1975, the first sleep disorders centre was established at Stanford University. This period also saw the recognition and classification of sleep disorders, with the publication of the first International Classification of Sleep Disorders (ICSD) in 1979. The discovery of the role of orexin/hypocretin in narcolepsy in 1989 was another significant milestone, shedding light on the mechanisms underlying this sleep disorder.

From the 2000s to the present, advances in genetic and neuroimaging studies have provided deeper insights into the mechanisms of sleep disorders. New treatments have been developed, including cognitive behavioural therapy for insomnia (CBT-I), which has become a widely accepted non-pharmacological treatment. There is also a growing recognition of the importance of sleep health in overall wellbeing, leading to increased public awareness and research funding. (By Dr. Liji Thomas, 2022)

Progression of Mental Health Studies

The study of mental health has a long and complex history, with significant shifts in understanding and approaches over time. In ancient times, mental illness was often attributed to supernatural causes, with treatments ranging from exorcisms to trephination. Hippocrates (460-370 BCE) was one of the first to propose that mental disorders had physical causes, laying the groundwork for a more scientific approach. The establishment of Bethlem Royal Hospital in London in 1247 marked the beginning of institutional care for mental illness.

The 19th century saw the emergence of psychiatry as a distinct medical field. Johann Christian Reil coined the term “psychiatry” in 1808, and the Association of Medical Superintendents of American Institutions for the Insane, the precursor to the American Psychiatric Association, was founded in 1844. Wilhelm Wundt established the first experimental psychology lab in 1879, further advancing the scientific study of mental health. The early 20th century brought the development of psychoanalysis by Sigmund Freud and the rise of behaviourism, marked by John B. Watson’s 1913 publication “Psychology as the Behaviourist Views It.”. (Seabrook, 2016)

The mid-20th century saw significant advancements in biological psychiatry and the deinstitutionalization movement. The first Diagnostic and Statistical Manual of Mental Disorders (DSM) was published in 1952, providing a standardized classification system for mental disorders. The introduction of the first antipsychotic medication, chlorpromazine, in 1954 revolutionized the treatment of mental illness. The

Community Mental Health Act of 1963 in the US promoted deinstitutionalization, shifting the focus to community-based care.

From the late 20th century to the present, there has been a growing emphasis on integrative approaches and neuroscience. Evidence-based treatments have become the standard, and the development of neuroimaging techniques has provided new insights into brain function. There is an increased focus on prevention, early intervention, and mental health promotion, recognizing the role of social determinants in mental health.

2.1.3 Previous Studies on Social Media, Sleep Disorders, and Mental Health

Social Media Use and Global Health

The famous American editor and publisher Elbert Hubbard stated, 'If you have health, you probably will be happy, and if you have health and happiness, you have all the wealth you need, even if it is not all you want.' Health may be the most important topic in humans' life. In the current study, health was approached as a general perspective. Global health includes both physical and mental health states (Davis, Balkin, and Juhnke, 2014). Studies showed that health was correlated with internet addiction (Chern and Huang, 2018), digital media use (Orzech et al., 2016), and social media or social networking sites (Bullen's and VandenBosch, 2016; Rae and Lomborg, 2015) among emerging adults. This literature showed that the interaction between young adults and computers could be problematic. In a study, Sicilia and Charoensukmongkol (2015) stated participants who were addicted to social media tended to have lower mindfulness and higher emotional exhaustion. Mindfulness is related to one's attention to the present moment, and lack of mindfulness can affect well-being, mental clarity, and academic achievement (Lin and Mai, 2018). Emotional exhaustion is related to burnout and causes people to feel lack of energy (Sicilia and Charoensukmongkol, 2015). As seen here, addictive social media use has both direct and indirect effects on health and its consequences.

Social Media Use and Mental health

In the literature of computers and human behaviour, one of the most investigated subjects is the relationship between computer use (digital media, social media, and internet) and depression. Depression is a common medical illness that negatively affects how people feel and act, and the way they think (American Psychiatric Association n.d.). A recent survey in the United States (U.S.) reported that emerging adults aged 18–25 years had the highest percentage of depression (13%) when compared to all U.S. adults (7.1%; National Institute of Mental Health 2017). Therefore, it is one of the reasons why it is investigated among emerging adults more than the other groups.
Although recent studies ((Ahmad, 2017), (Appel, 2016) Belichick, Eickhoff, and Moreno, 2013; (Jasso-Medrano, J. L., & López-Rosales, F. , 2018)-(Lin, 2018) (Balcı, 2018; Bilgin) is and Prinstein, 2015; Primack et al., 2017; Steers, 2016; Tandoc et al., 2015) investigated the associations between depression and social media use, there is no consensus that depression is an effect or a cause of internet and social media use (Balcı and Balogun, 2018; Blachnio, Przepiorka, and Panic, 2015). (Lin, 2018)conducted a study with 1718 U.S. emerging adults and found that social media use was correlated with increased depression. In another study, Ahmad et al. (2018) found similar results stating the students who spent more time on social media had more depression. On the other hand, Jasso-Medrano and López-Rosales (2018) investigated the relationship between social media use and addictive behaviour and depression and suicide ideation among Mexican university students. The results indicated that there was no significant relationship between social media use and depression. Similar to this, Elenchid et al. (2013) conducted a study with older adolescents and stated that there was no significant relationship between social media use and depression.

Social Media Use and Sleep Disturbance

Sleep is an important component for a healthy life, college success, and mood in university students. The New York Times personal health columnist Brody (2018) stressed out that sleep quality and sleep quantity was a key factor predicting grades and a student's chance of graduating. Sleep disturbance refers to the quality and amount of quantity based on sleep patterns or interruption (Davis et al., 2014; Karaman, Balkin, and Juhnke, 2018). Sleep problems can be serious. Studies showed that college students who had sleep disorders had lower academic grades (Gaultney, 2010), lower academic performance and attention (Pagel, Forister, and Kwiatkowski, 2007), depression (Carney, Edinger, Meyer, Lindman, and Istre, 2006), and poorer health (Smaldone, Honig, and Byrne, 2007).

The literature (Genes, Akbıyık, Aypak, and Grovelingly, 2018; Woods and Scott, 2016; van der Schuur, Baumgartner, and Sumter, 2018; van der Velden, Setti, van der Mullen, and Das, 2019) examined the associations between social media use and sleep problems in recent years. Although researchers used different terms (e.g., social networking sites' use, social media addiction, social media use) when referring to the social media usage, they mainly stressed 'social media use'.

In a recent study, Genes (2018) examined the relationship between social media dependency and sleep quality among 16–19 years old high school students. In the study, authors only added Facebook as social media site. The results indicated that as the Facebook Addiction Index score increased, sleep quality decreased. In a longitudinal study, van der Schuur et

al. (2018) examined the effects of social media and social media stress on sleep among adolescents in Netherlands. Authors stated that social media use and social media use stress may disrupt sleep. From a different point of view, Orzech et al. (2016) examined the factors affecting sleep variables on a larger scale. In the study, Orzech et al. (2016) analysed the effects of digital media use (e.g., playing video games, listening to music, surfing on the internet, using social media) on sleep variables of 254 first-year university students. They found that longer duration of digital media use was associated with reduced total sleep time. Moreover, they stated that looking at bright screens for a long time affects the level of melatonin and the quality of sleep. However, some findings suggested that there was no relationship between social media use and sleep disturbance. For example, van der Velden et al. (2019) examined to what extent social networking sites predicted mental health and sleep problems by working with 3486 participants aged between 16–74 years. The results showed that more social media use was not associated with sleep problems among emerging adults. These findings showed that the level of relationship between social media use and sleep disturbance was unclear. Therefore, more studies are needed to better explain the relationship between sleep disturbance and social media use among college students.

2.1.4 Conclusion of Introductory Concepts

In conclusion, it can be stated that the relationship between the use of the social network, sleep disorders and mental health is another active area of interest for the researchers with specific emphasis on the effects on young adults. Based on the findings, it is clear that the major social media platforms afford social relations as well as social ill related to depression, anxiety as well as emotional exhaustion. According to several works, it has been found that spending more time on social media negatively affects self- awareness, and increases loneliness and anxiety, while implying that the connections made possible by these sites are illusory and lack the psychological substance necessary for one assertion of quality interpersonal relationships. Such a model brings many questions into question, for instance, how one can be more connected to society but less connected to their mental well-being in an ever-digital world.

However, the association between social media and sleep disturbances is an emerging research subject, and the findings are inconclusive. Several papers highlight post posit that social media negatively impacts the quality of sleep – labelling problems inclusive of insomnia and sleep irregularities to time on screen and use of social media – while other papers posit that social media use does not correlate to sleep difficulties in young adults. The steadiness heard within the above highlights that there is more to this than meets the eye, because reaction that is peculiar to one person might not necessarily feel the same way as another, especially when personality differences are factored within this equation.

Finally, it is crucial to establish the complex interconnection between social media use, sleep, and mental health in order to design the strategies for improving the young adults' health-related behaviours and their online presence. The relationship of information literacy to learning, attitude, engagement and achievement in the LMS all merit more research in the future, particularly if future research takes a longitudinal approach and explores varied populations of students. Some of these insights could include the following: Such and such of the negative impacts of social media may be used in the formulation of policies that would ensure that this powerful medium becomes an avenue through which people can find each other and not the other way round.

2.2 Previous Studies without Machine Learning Models

Numerous studies have explored the relationship between social media use, sleep disturbances, and mental health without the use of machine learning techniques. These studies typically rely on traditional survey methods, clinical studies, or qualitative research approaches to uncover the impact of social media and sleep disorders on mental health outcomes.

2.2.1 Social Media Use and Sleep Disturbance

Although prior studies have established a significant link between SMU and sleep disruptions, this issue remains widely studied especially in adolescent and young adult populations. For instance, Güneş et al. (2018) learners with high school students and concluded that an augmented dependency on social media and especially Facebook was said to have contributed to poor sleep quality. The Facebook Addiction Index pointed out strong evidence of increased addiction score with reduced sleep quality.(By the same token, Orzech et al. (2016) assessed video gaming and social media usage and determined that greater utilization length, particularly pre-sleep, were linked to absolute sleep time contraction and sleep efficiency decrease due to the impact on melatonin secretion

Nonetheless, not all research works have expressed an equal correlational value between using social media and the disruption of sleep. Van der Velden et al. (2019) coordinated a huge study with 3,486 participants and did not observe the association between SNS use and sleep problems in emerging adults. Thus, the difference in the results emphasizes the need for additional research into the factors that explain the connection between sm use and sleep problems, especially if employing a broader variety of methods.

2.2.2 Social Media Use and Mental Health: Survey-Based Studies

A number of studies have investigated the effects of social media usage on mental health through survey techniques only. For example, (Primack, 2017)surveyed a massive sample of the young adults in the United States of America and noted that multiple platform SNS users had high rates of depression and anxiety. This was done to discover that more significant use of social media was associated with more serious symptoms of mental disorder than a lower use of social media platforms. In the same way, (Jasso-Medrano, J. L., & López-Rosales, F. , 2018) examine the connection between SNSDAA and depression in Mexican university students. Their research did not yield any positive relationship between social media use and suicidal ideation, that other mediating factors might help reduce the impacts of social media use, that could be offline social support.

However, (Ahmad, 2017) affirmed that high usage of social media was significantly and positively correlated to depression among university students in Pakistan. In their study, they admitted to using self-report questionnaires in evaluating depressive symptoms and estimated time on social media. There are two key elements that can be discussed concluding the analysis of this study: The role of the culture must be taken into account when evaluating the influence of digital behaviours on mental health.

Table 1 Key Survey-Based Studies on Social Media Use and Mental Health

Study	Sample Size	Findings
Primack et al. (2017)	1,787	Positive correlation between the use of multiple social media platforms and depression.
Jasso-Medrano and López-Rosales (2018)	542	No meaningful relationship between social media addiction and depression in Mexican students.
Ahmad et al. (2018)	503	Excessive social media usage linked to increased depression in Pakistani university students.

2.2.3 Clinical Studies on Sleep Disorders and Psychological Well-Being

Clinical research has also extensively explored the impacts of sleep disruptions on mental health. For instance, in a study by Gaultney (2010), college students with diagnosed sleep disorders reported poor academic performance and higher prevalence of anxiety and depression. Polysomnography was employed to assess different kinds of sleep disorders like insomnia and sleep pane; this aimed at informing the clinical approach to the detrimental effects of poor sleep quality on mental health.

For instance, Carney et al. (2006) examined the association between sleep and psychological well-being over time among college students. They further ascertained that sleep disturbances were a robust risk factor for depression and cognitive impairment among young people. Their research supports the notion that sleep disorders cannot be ignored if one expects to achieve positive mental health changes among students.

2.2.4 Qualitative Research on Social Media and Mental Health

Defining social media and mental health from a qualitative point of view is an interesting approach to understanding their connection. For instance, (Woods, H. C., & Scott, H., 2016) studied the social media youth’s views about its impact on sleep and their psychological

Study	Method	Key Findings
Gaultney (2010)	Clinical study	Sleep disorders in college students linked to lower academic performance and higher anxiety levels.
Carney et al. (2006)	Longitudinal study	Sleep disturbances predicted depression and cognitive dysfunction in young adults.
Woods and Scott (2016)	Qualitative study	Adolescents reported feeling anxious and stressed due to nighttime social media use.

Table 2 Key Clinical and Qualitative Studies on Sleep and Mental Health

well-being. A qualitative feed-back was identified by interviewing patients and the findings highlighted that patient felt more anxious and stressed especially at night because of the social media. This work focuses on the existential life of young adults and underscores how use of social media affects their mental health beliefs.

2.3 Machine Learning Techniques in Social Media, Sleep Disorders, and Mental Health

Machine learning (ML) methods are more and more being used in studies examining how social media usage, sleep disorders, and mental health are related. These methods offer sophisticated approaches for examining intricate datasets, enabling the detection of patterns and forecasting results. Listed below are fundamental machine learning methods applied in this research field:

2.3.1 Natural Language Processing (NLP) for Social Media Sentiment Analysis

Natural Language Processing (NLP) is a crucial technique used to analyze textual data from social media platforms, enabling researchers to extract emotional content, sentiment, and psychological markers.
For example, Shen et al. (2017) applied NLP to Twitter data, detecting depressive language markers. They discovered that words like "sad," "lonely," and "hopeless" were significant predictors of depression. Similarly, Guntuku et al. (2019) used NLP to analyze Instagram posts, finding that users experiencing depression posted darker images and used more emotionally negative language. These studies highlight the utility of NLP in identifying at-risk individuals based on their social media content.

2.3.1 Time Series Analysis for Sleep Pattern Recognition

Time series analysis is commonly used in machine learning to examine how data points change over time. In sleep disorder studies, time series methods analyze data from wearable devices to detect and predict sleep disturbances.

(Lin, 2018) applied time series analysis to data from wearable devices, identifying patterns in sleep behavior that could predict disorders such as insomnia or sleep apnea. By examining sleep duration, interruptions, and circadian rhythms, they developed predictive models that correlated poor sleep quality with mental health issues such as anxiety and depression.

2.3.2 Supervised Learning Algorithms for Predictive Modeling

Supervised learning algorithms like decision trees, support vector machines (SVM), and random forests have been widely used to predict mental health outcomes based on social media use and sleep data. These models use labeled data (e.g., hours spent on social media, sleep quality, mental health symptoms) to train algorithms that can predict specific outcomes, such as depression or anxiety.

For example, Reece and Danforth (2017) used supervised learning to predict depression in Twitter users based on their language, posting frequency, and interaction patterns. They built a model that successfully identified individuals with a high risk of depression based on their Twitter activity. Similarly, Saeb et al. (2016) developed a supervised learning model using smartphone data to predict depressive episodes, finding that reduced mobility and high social media engagement were strong predictors of depression.

2.3.3 Unsupervised Learning for Behavioral Clustering

There are abovementioned uses of unsupervised learning where algorithms like k-means clustering and PCA are employed to identify the latent structure in the behavior of users. Unlike other solutions, these techniques do not incorporate prior knowledge labels and can cluster people depending on their behavior, for instance, sleeping habits or social media engagement. (Karaman, 2024)

Technique	Application	Key Findings
NLP (Shen et al., 2017)	Analysing tweets for depressive language	Identified markers like "sad" and "lonely" as predictors of depression.
Time Series (Liu et al., 2020)	Wearable device data for sleep pattern analysis	Modelled sleep disruptions and linked them to future mental health outcomes like anxiety.
Supervised Learning (Reece and Danforth, 2017)	Predicting depression using Twitter data	Predicted high-risk individuals based on social media activity.
Unsupervised Learning (Karaman et al., 2019)	Clustering social media users based on behaviour	Identified groups with high social media addiction and sleep disturbances.

Table 3 Key Machine Learning Techniques Applied to social media, Sleep, and Mental Health

2.3.4 Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) for Sequential Data Analysis

Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) models are highly effective in processing sequential data, such as time-series information. These models capture temporal dependencies, making them well-suited for analysing social media behaviours, sleep patterns, and mental health indicators over time.

For instance, RNNs and LSTMs have been applied to data from social media to monitor user behaviour patterns continuously and detect early signs of mental health disorders. In sleep studies, LSTMs have been used to forecast sleep disruptions by learning from past patterns of sleep quality, interruptions, and circadian rhythms. By utilizing LSTM's ability to retain important information over long periods, researchers can make more accurate predictions regarding future mental health outcomes such as anxiety, depression, or sleep disorders (Zhao et al., 2020). These advanced techniques outperform traditional models in capturing the dynamic nature of human behaviour and health trajectories.

2.4 Previous Studies with Machine Learning Models

Deeper into the application of machine learning techniques across various fields, particularly in the context of mental health, sleep disorders, and social media influence. Here's how different machine learning algorithms are applied, with references to specific studies:

2.4.1 Karaman, Mehmet Akif (2019)

The study focuses on analyzing the association between social media use and mental health issues like depression, sleep disturbances, and overall global health among emerging adults. Using k-means clustering, participants were divided into three categories: ordinary/none, mild, and severe users of social media. The findings revealed that individuals who spent more time on social media experienced higher levels of depression and sleep issues. This clustering approach is effective in identifying subgroups within a population that are at greater risk of mental health disorders due to excessive social media use. The study further highlights how the time spent on social media (TSSM) and social networking addiction (SNA) impact psychophysiological well-being, demonstrating a complex relationship between digital media usage and health outcomes. (Karaman, 2024)

Application: This demonstrates the utility of clustering algorithms in categorizing individuals based on behavioral data, such as social media usage, and understanding its implications on mental health. Such clustering techniques can also help in targeted interventions for mental health management.

2.4.2 Sulaiman, Rejwan Bin, et al. (2022)

This review focuses on machine learning applications in credit card fraud detection, which parallels many healthcare-related applications in handling imbalanced datasets and detecting anomalies (such as abnormal behavioral patterns in mental health or sleep data). The study explores multiple machines learning models, including: (Sulaiman, R. B., Schetinin, V., & Sant, P. , 2022)

Random Forest (RF)

A robust technique in handling large datasets and performing well with imbalanced data, which is often the case in healthcare where mental health disorders or sleep disturbances may be less frequent but critical to detect.

Support Vector Machines (SVM): Effective for classification tasks, such as distinguishing between normal and abnormal sleep patterns or identifying depressive tendencies based on social media usage patterns.

Artificial Neural Networks (ANN): Utilized to forecast results, such as conditions related to mental well-being. Artificial neural networks imitate the functioning of the human brain and are beneficial for intricate tasks such as detecting non-linear connections between social media usage and mental health consequences.

k-nearest neighbors (KNN): Applied in regression and classification problems, it helps to identify similar patterns in past data, such as similar behavioral trends in individuals experiencing sleep issues or mental health disorders. (Sulaiman, R. B., Schetinin, V., & Sant, P. , 2022)

2.4.3 Federated Learning and Data Privacy

In recent years, federated learning has emerged as a powerful machine learning approach, particularly when handling sensitive data such as personal health information. Federated learning allows machine learning models to be trained on distributed data without requiring direct access to sensitive information. This technique is particularly relevant in the context of healthcare, where patient privacy is paramount.

The federated learning model proposed in the Sulaiman et al. (2022) study is also applicable to social media and mental health analysis. By training models locally on user devices, federated learning ensures that sensitive mental health data remains private while allowing robust machine learning models to be built for predictive mental health outcomes.

Application: In mental health and sleep disorder analysis, federated learning can enable large-scale studies across multiple healthcare institutions without compromising patient privacy, providing insights into patterns of depression, anxiety, and other disorders associated with social media use or sleep deprivation.

2.4.4 Hybrid Machine Learning Approach

The use of hybrid models combining techniques such as Random Forest and Isolation Forest for anomaly detection has also been explored. These models are effective in identifying outliers in large datasets, which can be applied to detect anomalies in sleep data (e.g., frequent sleep interruptions indicating disorders) or sudden changes in social media behavior that may signal mental health issues.

Application: In healthcare and mental health analysis, hybrid models allow for more accurate detection of irregular patterns, which can lead to early intervention and treatment of sleep disorders or depressive symptoms.

Application: The insights from credit card fraud detection studies can be applied to detect anomalies in health data, such as irregular sleep patterns or sudden changes in mental health status. The ability to handle imbalanced datasets (a few cases of severe mental health issues) is crucial in early diagnosis and prevention strategies.

2.4 Discussion and Conclusion

The literature examined emphasizes the complex connection between use of social media, sleep interruptions, and mental well-being, showing agreement and disagreement in results from different studies. Conventional research methods, such as surveys and clinical trials, have established the foundation for comprehending these interactions. For example, research conducted by Güneş et al. (2018) and Orzech et al. (2016) shows how social media negatively affects the quality of sleep, especially in young adults. Yet, Van der Velden et al. (2019) have found conflicting results that highlight the variability in reported connections, suggesting that the correlation is intricate and affected by several variables. This discrepancy requires more thorough research methods that incorporate a variety of approaches, such as gathering objective data, to gain a better understanding of the factors at play.

The utilization of machine learning methods signifies a major progress in this field of study, providing advanced tools for examining vast and complex datasets. Methods like Natural Language Processing (NLP) and supervised learning algorithms have shown their ability to reveal nuanced patterns that conventional techniques could overlook. An example is the research conducted by Reece and Danforth (2017), demonstrating how machine learning can accurately forecast mental health results by analyzing social media activity, thus allowing for the early detection of vulnerable individuals. Furthermore, employing time series analysis to study sleep patterns offers a more in-depth insight into the relationship between changes in sleep and mental health indicators, permitting more detailed analysis of the data.

Although promising outcomes are being achieved with machine learning, there are still challenges that remain, especially in relation to data privacy and the representativeness of sample populations. Sulaiman et al. (2022) points out that federated learning presents a promising option for analyzing sensitive health information while safeguarding user privacy. This new method can help increase involvement in research studies.

CHAPTER THREE
RESEARCH METHODOLOGY

3.1 Data for the Study

3.1.1 Full Description of the Dataset

The dataset applied for this research offers a detailed analysis of daily technology engagements, such as social media usage and screen time, and mental health. This dataset has been obtained from an online survey conducted in the year 2022, which provides more significance and recent records of the behavioral pattern of people with digital platforms and their influence on wellbeing status. The initial dataset consists of 10,000 records (thousand rows of observation data) and 14 features (14 pieces of column data) in terms of participants' daily activity and mental state.

```
print("Data Columns")
df.columns
```

```
Data Columns
Index(['Timestap', 'User_ID', 'Age', 'Birth Year', 'Generation',
       'Technology_Usage_Hours', 'Social_Media_Usage_Hours', 'Gaming_Hours',
       'Screen_Time_Hours', 'Mental_Health_Status', 'Stress_Level',
       'Sleep_Hours', 'Physical_Activity_Hours', 'Support_Systems_Access',
       'Work_Environment_Impact', 'Online_Support_Usage',
       'Mental_Health_Status_Numerical'],
      dtype='object')
```

Figure 2 Column Names in the Dataset

Key Characteristics:

User_ID: A unique identifier for each participant, ensuring that individual responses are anonymized while maintaining data integrity.

Age: This numerical variable captures the age of participants, ranging from teenagers to older adults. Age can influence how technology and social media usage impact mental health, as younger generations may engage differently with digital platforms than older populations.

Daily_Screen_Time (hours): This feature records the average number of hours each participant spends using screens daily, including social media, entertainment, and work-related activities. This variable is crucial in understanding how prolonged screen exposure affects mental health.

Mental_Health_Score (1-10): A self-reported measure of mental health, where participants rate their overall well-being on a scale of 1 to 10, with 1 indicating poor mental health and 10 representing excellent mental health. This variable serves as the primary outcome of interest in evaluating the impact of technology use.

Stress_Level (1-10): Another self-reported measure, this variable captures the perceived stress levels of participants on a scale of 1 to 10. High stress levels are often linked to excessive technology use and poor sleep quality, making this a key variable in the study.

Sleep_Quality (1-10): This variable assesses participants' sleep quality on a scale of 1 to 10, where 1 represents poor sleep and 10 indicates excellent sleep quality. Given the significant role that sleep plays in mental health, understanding its relationship with screen time and technology use is a major focus of the analysis.

Demographic Information

The dataset includes a diverse range of participants across multiple generations, which provides valuable insights into how age groups differ in terms of social media usage, sleep patterns, and mental health outcomes. The distribution of participants by generation is as follows: Gen X (35%), Millennials (31%), Gen Z (18%), and Baby Boomers (16%). These generational differences are essential for analysing trends and identifying patterns in behavior and mental health across various age groups.

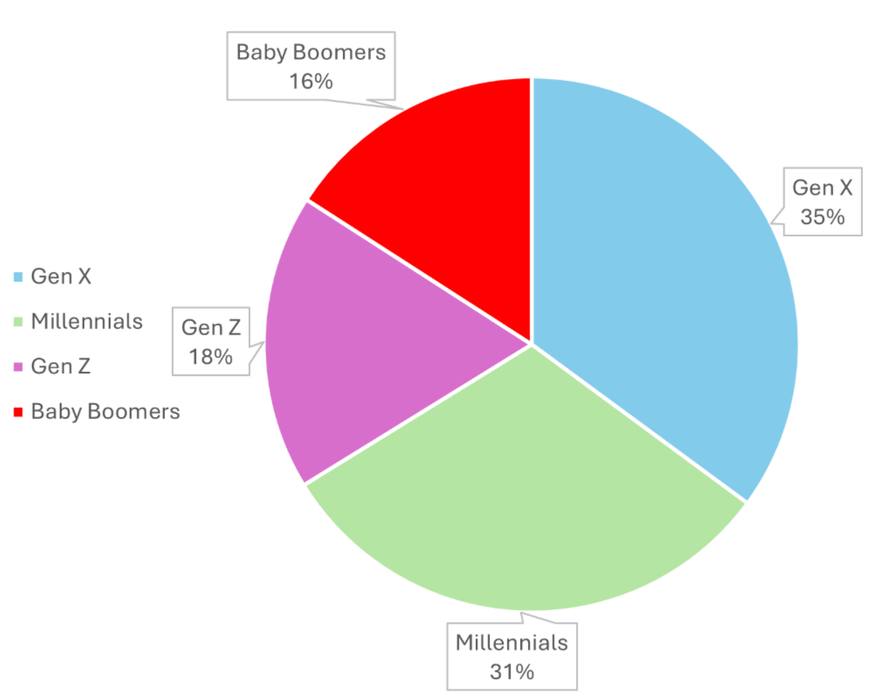

Figure 3 Distribution Of Generations In Dataset

Generation	Birth Year
Baby Boomers	1946 - 1964
Generation X	1965 - 1979
Millennials	1980 - 1994
Generation Z	1995 - 2014

Table 4 Generational Categories and Corresponding Birth Years

Additional Features:

Physical Activity Levels: A numerical measure indicating the number of hours participants spend engaging in physical activity, which is known to positively affect both mental and physical health.

Technology Usage Breakdown: Subcategories of screen time, such as hours spent on social media, gaming, and other forms of technology use, providing more detailed insights into the type of screen engagement.

Work Environment Impact: A binary variable indicating whether participants' work environment contributes to increased stress or impacts their mental health.

Data Source and Time Frame:

The dataset was collected through an online survey in 2022, designed to capture the current technology habits and mental health statuses of individuals. The survey targeted participants from diverse age groups, ensuring a broad representation of how digital behavior affects different demographics.

Preprocessing Steps:

Before analysis, the dataset underwent several preprocessing steps to ensure data quality. Missing values were handled using imputation techniques where possible, and any duplicate entries were removed. Numerical variables like screen time and age were normalized to account for skewed distributions, making the data more suitable for machine learning models. Categorical variables such as work environment were encoded for compatibility with various analytical techniques.

3.1.2 Justification for the Dataset

The selection of this dataset is appropriate to the goals of the study that is to elucidate the complex interactions between the level of technology engagement and mental health indicators. Since the collected dataset discriminates for participants' time spent on screens in hours, sleep quality, stress, and mental wellness, the dataset corresponds with the primary investigative questions that concern the impacts of digital speculation on mental health.

Relevance to Research Objectives:

The dataset enables analyzing the extent to which technology, and especially the frequency of social media use, leads to stress and poor quality of sleep. It supports both descriptive and predictive analyses so it can be used to develop models of the mental health based on screen time and other factors that can predict the likely path that future models will have. This is well in line with the study objectives seeking to fill a gap in knowledge regarding the societal impact of enhanced enactment of digital technologies particularly among young caucuses.

```
print("Dataset Information")
df.info()

Dataset Information
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 17 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   Timestap                      10000 non-null  datetime64[ns]
 1   User_ID                       10000 non-null  object
 2   Age                           10000 non-null  int64
 3   Birth Year                    10000 non-null  int64
 4   Generation                    10000 non-null  object
 5   Technology_Usage_Hours        10000 non-null  float64
 6   Social_Media_Usage_Hours      10000 non-null  float64
 7   Gaming_Hours                  10000 non-null  float64
 8   Screen_Time_Hours             10000 non-null  float64
 9   Mental_Health_Status          10000 non-null  object
 10  Stress_Level                  10000 non-null  object
 11  Sleep_Hours                   10000 non-null  float64
 12  Physical_Activity_Hours       10000 non-null  float64
 13  Support_Systems_Access        10000 non-null  object
 14  Work_Environment_Impact       10000 non-null  object
 15  Online_Support_Usage          10000 non-null  object
 16  Mental_Health_Status_Numerical  10000 non-null  int64
dtypes: datetime64[ns](1), float64(6), int64(3), object(7)
memory usage: 1.3+ MB
```

Figure 4 Dataset summary with 10,000 entries across 16 features

Previous Studies and Credibility:

Many previous researchers have used similar datasets to assess the effects of technology on Phys. For instance, (Cain & Gradisar, 2010) & (Woods, H. C., & Scott, H., 2016) (Scott & Cleland Woods, 2019)it became apparent that there is correlation between technology use like social media and changes of sleep period resulting to increased stress and anxiety. These findings' external validation supports the appropriateness of employing this data set in future research regarding the matter. There is also a large number of respondents in the study involving 10,000 participants, and this automatically lifts the reliability and reliability of the research.

This dataset also permits the use of more fine-grained 'learning solutions,' including Recurrent Neural Networks (RNN) and Long Short-Termed Memory (LSTM) to forecast certain mental health profiles based on relational data. This investigation can help enrich the literature on how digital lifestyles are reshaping the mental health regime, if it builds upon a clean dataset with markers for mental health indicators.

3.2 Framework of the Study

The framework of this study is designed to explore the intricate relationships between social media usage, sleep patterns, and mental health outcomes, focusing on factors such as stress, anxiety, and overall mental well-being. By incorporating machine learning techniques, this study aims to identify behavioral patterns that can inform interventions to improve mental health, particularly among adolescents and young adults. The framework employs a structured approach, moving sequentially from data collection to result presentation, complemented by visual aids and charts to enhance understanding.

The following steps outline the methodological process, where various charts and visualizations are integrated to display key findings:

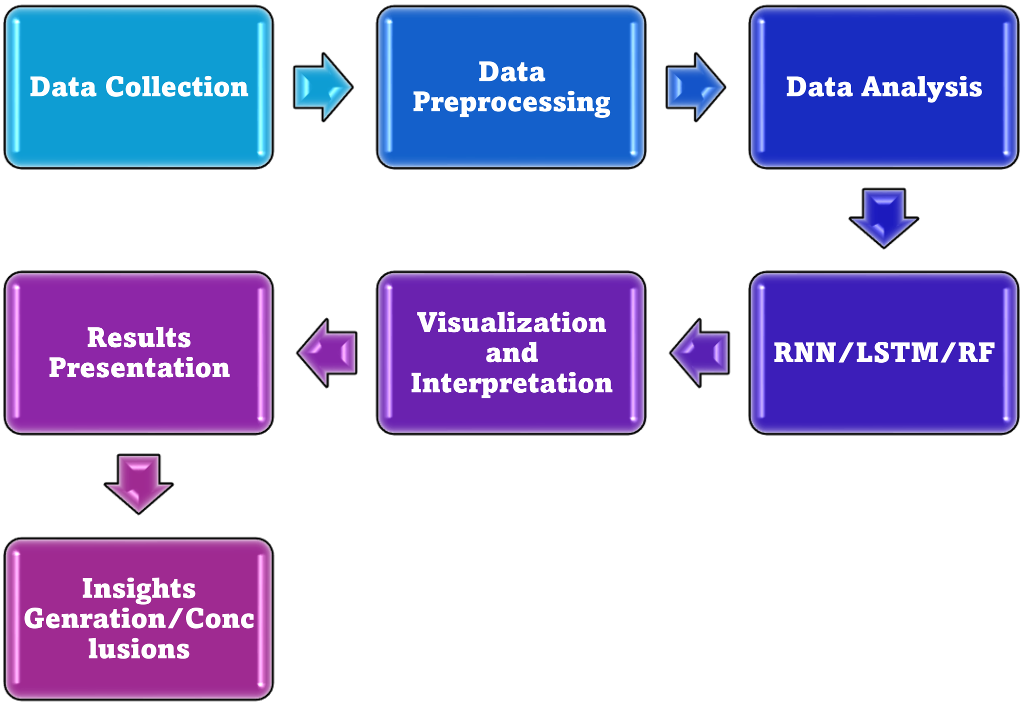

Figure 5 This flowchart illustrates the sequential process of analysing the relationship between social media usage, sleep, and mental health using machine learning techniques. It outlines the key stages of data collection, preparation, analysis, visualization and insights.

3.2.1 Data Collection

The dataset consists of 10,000 rows and 14 columns, collected through surveys and tracking tools. Key features include variables such as Technology Usage Hours, Social Media Usage Hours, Sleep Hours, and Mental Health Status, with all columns having 0 null values.

```
Data Columns with index:
0: Timestap
1: User_ID
2: Age
3: Birth Year
4: Generation
5: Technology_Usage_Hours
6: Social_Media_Usage_Hours
7: Gaming_Hours
8: Screen_Time_Hours
9: Mental_Health_Status
10: Stress_Level
11: Sleep_Hours
12: Physical_Activity_Hours
13: Support_Systems_Access
14: Work_Environment_Impact
15: Online_Support_Usage
```

Figure 6 Column names with index for dataset of 16 Variables

3.2.2 Data Preprocessing:

This process includes loading the dataset, importing essential libraries, handling missing values, performing data cleaning, and generating summary statistics. These steps ensure data integrity, consistency, and readability, making the dataset well-suited for further exploration and modeling.

Loading the Dataset

The dataset was loaded using pandas with the *pd.read_excel* function. This allowed for a direct read of the Excel file, making it immediately available for exploration and analysis.

Importing Libraries

For performing data cleaning, data preparation and model training to enhance and automate various data analysis procedures, fundamental libraries were imported. The imported libraries include:

Fast data manipulation using pandas *numpy etc.label* for converting categorical data into numerical and *StandardScaler* for scaling the numerical features.

we use *matplotlib.pyplot* and seaborn to visualize the distributions of the data we use.

The general machine learning models we used from the *sklearn.linear_model, sklearn.ensemble,* and *sklearn.svm* are Log ()

The *RandomForestClassifier,* and SVC for supervised learning.

sklearn.model_selection to splitting the data into training and testing dataset.

This package contains implementing functions of various performance metrics for the model, that is, *sklearn.metrics.*

jupyter caspstone_2 Last Checkpoint: 4 minutes ago

File Edit View Run Kernel Settings Help

Data Preprocessing:

```
[1]: import pandas as pd
import numpy as np
from sklearn.preprocessing import LabelEncoder, StandardScaler
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, classification_report
from sklearn.ensemble import RandomForestClassifier
from sklearn.svm import SVC
```

```
[5]: import pandas as pd

file_path = "C:/Users/DELL/Downloads/capstone 4th sem/mental_health_and_technology_usage_2024vcapstone 2.xlsx"
data = pd.read_excel(file_path)
```

Figure 7 Data Preprocessing Setup - Importing Libraries and Dataset

To understand the structure and content of the dataset, we generated summary statistics using *data.describe()*. This summary provided key insights into each numerical column.

```
[6]: print("\nSummary Statistics for Numerical Columns:")
data.describe()
```

Summary Statistics for Numerical Columns:

[6]:

	Age	Technology_Usage_Hours	Social_Media_Usage_Hours	Gaming_Hours	Screen_Time_Hours	Sleep_Hours	Physical_Activity_Hours
count	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000
mean	41.518600	6.474341	3.972321	2.515598	7.975765	6.500724	5.003860
std	13.920217	3.169022	2.313707	1.446748	4.042608	1.450933	2.905044
min	18.000000	1.000000	0.000000	0.000000	1.000000	4.000000	0.000000
25%	29.000000	3.760000	1.980000	1.260000	4.520000	5.260000	2.490000
50%	42.000000	6.425000	3.950000	2.520000	7.900000	6.500000	4.990000
75%	54.000000	9.212500	5.990000	3.790000	11.500000	7.760000	7.540000
max	65.000000	12.000000	8.000000	5.000000	15.000000	9.000000	10.000000

Figure 8 Summary Statistics for Numerical Columns

The summary statistics for the dataset provided valuable insights into the data distribution. The count represents the total number of observations for each feature, confirming data completeness. The mean and standard deviation serve as indicators of central tendency and variability, respectively, allowing us to understand the spread and dispersion of values across features like Technology_Usage_Hours, Sleep_Hours, and Stress_Level. Additionally, the minimum, maximum, and quartiles (25%, 50%, 75%) offer a detailed view of the data's range and distribution, helping to identify any potential outliers and assess the skewness within the dataset. These statistics provide a foundational understanding of the dataset before proceeding with deeper analysis.

Handling Missing Values

To ensure data integrity, we began by checking for missing values using the *data.isnull().sum().sum() function*. This function calculates the total count of null values across the dataset. In this case, it returned 0, indicating that the dataset was already clean, with no missing values present. This completeness ensured that no additional imputation or data loss was required, allowing for straightforward analysis.

HANDLING MISSING VALUES AND DATA CLEANING ¶

```
[7]: total_null_values = data.isnull().sum().sum()
     print("\nTotal Null Values in the Dataset:", total_null_values)

     data = data.dropna()
     data.columns = data.columns.str.strip().str.replace(' ', '_').str.lower()
```

Figure 9 Handling missing values and Data Cleaning Process

Data cleaning

we performed data cleaning by standardizing the column names to ensure consistency and accessibility. This process involved stripping any leading or trailing whitespace, replacing spaces with underscores for improved compatibility in code, and converting all column names to lowercase for uniformity. These adjustments make the column names easier to reference, enhance code readability, and reduce the likelihood of errors during data manipulation. Overall, this step prepared the dataset for efficient analysis and ensured consistency across all data handling tasks.

3.2.3 Data Analysis

Machine learning models (RNN, LSTM, and Random Forest) are applied to analyze the relationships between Technology Usage, Sleep Patterns, and Mental Health Status. Each model is used to predict outcomes such as Stress Levels and to identify important predictors like Social Media Usage Hours.

3.2.4 Visualization and Interpretation

The results are visualized using graphs and charts, making it easier to interpret key patterns in the data, such as correlations between Social Media Usage, Sleep Hours, and Stress Levels.

3.2.1 Insights Generation/Conclusions:

Insights are derived based on the analysis, such as how excessive social media usage correlates with higher stress and lower sleep hours. Conclusions are made to suggest actionable recommendations for improving mental health.

3.2.2 Results Presentation:

The findings are compiled into a presentation or report, clearly highlighting the relationships between variables like Technology Usage, Work Environment Impact, and Mental Health Status. Recommendations are provided based on the analysis.

Figure 10 Research process for analysing social media, screen time, and mental health using machine learning, focusing on data collection and analysis.

3.3 Analytical Techniques

This study utilizes multiple modern techniques and approaches in machine learning to solve the stated research questions. The following algorithms were selected based on their specific capabilities to model the diverse data types involved in this research: These are Recurrent

Neural Networks (RNN), Long Short-Term Memory (LSTM) networks and the Random Forest (RF) algorithms. All these techniques provide different benefits in dealing with the elaborate multiple-dimensional data, which was gathered for the given investigation. RNNs work best for sequence data, and hence they can be used to analyze temporal characteristics associated with the usage of social media. Recipients use another subcategory of RNNs, known as the LSTMs, for accurately identifying long dependency sequences that are essential for determining aggregated impacts of social media use on sleep and well-being. As another set of complex models (Random Forest algorithms) offer good results in capturing the non-linear nature for the interdependencies of factors, their importance in defining the observed values have been quantified. It is our intention that through the combined use of these mutually supportive methodologies we will be able to offer a detailed and rich picture of the interdependency between digital engagement, sleep quality, and psychological state.

3.3.1 Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNN) are a class of neural networks designed to handle sequential or time-series data by retaining information from previous steps in the sequence. They excel at capturing temporal dependencies, making them useful for analyzing daily behaviors like social media usage and sleep patterns over time.

Step-by-Step Explanation:

Define Research Objective: Outline the goal of predicting social media usage patterns over time.

```python
[15]: # Extract social media usage data
usage_data = data['Social_Media_Usage_Hours'].values.reshape(-1, 1)

# Normalize the Data
scaler = MinMaxScaler(feature_range=(0, 1))
usage_data_scaled = scaler.fit_transform(usage_data)
print("Data normalized successfully.")

# Create Sequences
def create_sequences(data, seq_length=7):
    X, y = [], []
    for i in range(len(data) - seq_length):
        X.append(data[i:i+seq_length, 0])
        y.append(data[i+seq_length, 0])
    return np.array(X), np.array(y)

X, y = create_sequences(usage_data_scaled)
print("Sequences created successfully.")

# Reshape Data for RNN
X = X.reshape((X.shape[0], X.shape[1], 1))
print("Data reshaped for RNN input.")
```

```
Data normalized successfully.
Sequences created successfully.
Data reshaped for RNN input.
```

Figure 11 Data Preparation for RNN Model

Data Collection and Input: Gather daily social media usage and screen time data; load with pandas.

Data Preprocessing: Clean data, normalize values, and create time-based sequences.

Define RNN Model Architecture: Initialize RNN, add input, hidden (RNN/LSTM), and output layers.

Model Training Setup: Split data into training/testing sets and compile model with appropriate loss and optimizer.

Train the Model: Train RNN using BPTT, monitoring for overfitting with validation data.

```python
# Define RNN Model
model = Sequential()
model.add(SimpleRNN(units=50, activation='relu', input_shape=(X.shape[1], 1)))
model.add(Dense(units=1))
model.compile(optimizer='adam', loss='mean_squared_error')
print("RNN model defined and compiled.")

# Split Data for Training and Testing
train_size = int(len(X) * 0.8)
X_train, X_test = X[:train_size], X[train_size:]
y_train, y_test = y[:train_size], y[train_size:]
print("Data split into training and testing sets.")

# Train the Model
print("Training the model...")
history = model.fit(X_train, y_train, epochs=50, batch_size=32, validation_data=(X_test, y_test))
print("Model training completed.")

# Model Evaluation
y_pred = model.predict(X_test)
y_pred = scaler.inverse_transform(y_pred.reshape(-1, 1))
y_test_actual = scaler.inverse_transform(y_test.reshape(-1, 1))
```

Figure 12 RNN model setup, training, and evaluation using Keras.

Evaluate Model Performance: Use metrics like MAE and MSE to assess model accuracy on test data.

Generate Predictions: Predict future social media usage based on historical patterns.

Interpret and Discuss Results: Analyze trends, discuss limitations, and suggest model improvements.

Conclude the Study: Summarize findings and suggest future research directions.

RNN can track the evolution of social media usage and screen time over several days, allowing the analysis of trends and comparisons between different periods of usage. This helps in understanding the average time spent on social media and how it fluctuates over time.

3.3.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks are a type of RNN designed to handle long-term dependencies in time-series data. Unlike RNNs, LSTMs incorporate memory cells that can retain information over longer periods, making them ideal for tasks where long-term patterns need to be recognized, such as the cumulative effects of social media usage on mental health.

Step-by-Step Explanation:

Define Research Objective: Set the goal to predict future stress and mental health outcomes based on past social media usage, sleep quality, and stress levels.

Data Collection and Input: Collect sequential data on daily social media usage, sleep quality, and stress levels, then load it into the environment using pandas.

Data Preprocessing: Clean the dataset, handle missing values, normalize features, and create time-based sequences to capture past behavioral patterns.

Define LSTM Model Architecture:

Initialize the LSTM model.

Add memory cells with input, forget, and output gates to retain, discard, or pass information as needed.

Set up input, hidden, and output layers to capture long-term dependencies.

Model Training Setup: Split the data into training and testing sets while preserving chronological order and compile the model with a suitable loss function (e.g., Mean Squared Error) and optimizer (e.g., Adam).

Train the Model: Train the LSTM using Backpropagation Through Time (BPTT), allowing the gating mechanisms to learn which variables (e.g., sleep quality, social media usage) have long-term effects on stress and mental health.

Evaluate Model Performance: Use metrics like Mean Absolute Error (MAE) and Mean Squared Error (MSE) to assess model accuracy in predicting future stress levels.

Generate Predictions: Use the trained model to predict future stress levels and mental health outcomes based on past social media and sleep patterns.

Interpret and Discuss Results: Analyze how prolonged social media use and insufficient sleep contribute to elevated stress and poor mental health. Highlight patterns where certain behaviors (e.g., low sleep, high social media use) consistently impact mental health.

```python
mental_health_data['Stress_Level'] = mental_health_data['Stress_Level'].map({'Low': 0, 'Medium': 1, 'High': 2})  # Exc

# Step 3: Scale numeric columns
scaler = MinMaxScaler(feature_range=(0, 1))
scaled_data = scaler.fit_transform(mental_health_data[['Social_Media_Usage_Hours', 'Sleep_Hours', 'Stress_Level']])

# Step 4: Create sequences
def create_sequences(data, seq_length=7):
    X, y = [], []
    for i in range(len(data) - seq_length):
        X.append(data[i:i + seq_length])
        y.append(data[i + seq_length, -1])  # Predicting Stress_Level
    return np.array(X), np.array(y)

X, y = create_sequences(scaled_data)

# Step 5: Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, shuffle=False)

# Step 6: Define and compile the LSTM model
model = Sequential([
    LSTM(50, activation='relu', input_shape=(X.shape[1], X.shape[2])),
    Dense(1)
])
model.compile(optimizer='adam', loss='mse')

# Step 7: Train the model
model.fit(X_train, y_train, epochs=50, batch_size=16, validation_data=(X_test, y_test))

# Step 8: Evaluate the model
y_pred = model.predict(X_test)
mae = mean_absolute_error(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)
print(f"Mean Absolute Error (MAE): {mae}")
print(f"Mean Squared Error (MSE): {mse}")

# Step 9: Make a prediction
sample_data = X_test[-5:]  # Example: use the last few instances of the test set
predicted_stress_levels = model.predict(sample_data)
print("Predicted future stress levels:", predicted_stress_levels)
```

Figure 13 LSTM Model Training and Prediction Step

3.3.3 Random Forest (RF)

Random Forest (RF) is an ensemble learning algorithm used for classification and regression tasks. It creates multiple decision trees during training and combines their outputs to improve prediction accuracy. RF is especially useful for identifying important features in a dataset, such as the variables that most influence stress or mental health outcomes.

Step-by-Step Explanation:

Define Research Objective: Set the goal to classify individuals into stress levels based on behavioral features.

Data Collection and Input: Load a dataset with features like social media usage, sleep quality, and physical activity.

Data Preprocessing: Handle missing values, encode categorical variables, and define features (X) and labels (y).

Split Data: Split data into training and testing sets.

Define and Train Model: Initialize and train the Random Forest model on the training data.

Evaluate Model: Assess model performance with accuracy and classification report.

Feature Importance Analysis: Identify the most influential features on stress levels.

Make Predictions: Use the trained model to classify stress levels based on sample input.

Interpret Results: Analyze key features affecting stress and discuss findings.

Application Summary: Summarize how the model can help identify and manage stress factors effectively.

The Random Forest model classifies individuals into different stress levels (low, medium, high) and provides a ranking of the most important factors contributing to stress. This allows us to focus on key drivers such as reducing social media usage or improving sleep quality to reduce stress levels.

```
print("Step 3: Encoding other categorical columns in X")
categorical_columns = ['Work_Environment_Impact']
for col in categorical_columns:
    mental_health_data[col] = label_encoder.fit_transform(mental_health_data[col])
print("Categorical columns encoded successfully:", categorical_columns)

print("Defining features and labels")
X = mental_health_data[['Social_Media_Usage_Hours', 'Sleep_Hours', 'Physical_Activity_Hours', 'Work_Environment_Impact']]
y = mental_health_data['Stress_Level']
print("Features and labels defined successfully")

print("Step 4: Splitting data into training and testing sets")
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
print("Data split into training and testing sets successfully")

print("Step 5: Defining and training the Random Forest model")
rf_model = RandomForestClassifier(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)
print("Random Forest model trained successfully")

print("Step 6: Evaluating the model")
y_pred = rf_model.predict(X_test)
print("Accuracy:", accuracy_score(y_test, y_pred))
print("Classification Report:\n", classification_report(y_test, y_pred))

print("Step 7: Feature Importance Analysis")
feature_importances = rf_model.feature_importances_
print("\nFeature Importance Scores:")
for feature, importance in zip(X.columns, feature_importances):
    print(f"{feature}: {importance:.2f}")

print("Step 8: Making predictions")
```

Figure 14 Random Forest Model for Stress Level Prediction

3.4 Justification of Analysis Techniques

The selection of Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, and Random Forest (RF) algorithms for this study is guided by the specific nature of the data and the research questions, each of which demands a different approach to analyze the relationship between social media usage, sleep patterns, and stress or mental health outcomes.

3.4.1 Recurrent Neural Networks (RNN)

Rationale: The use of RNN can be particularly beneficial when the data is sequential or time-dependent which is true in the case of the dataset utilised in this work. Social media use, sleep, and stress are time series data: time series implies that there is a temporal sequence to this data and therefore needs a method that shall capture this sequence. Ordinary forms of ML do not consider time-bound interactions; hence, RNN becomes critical in monitoring behavior shifts with regard to time.

Because of this characteristic of RNN, a model is capable of predicting future occurrences by basing its results on the behaviors of the past occurrences. For example, the forecasting of daily tendencies in active social network usage might introduce important information on the dynamics of people's behavior, which is important for determining the influence of improper amount of time spent on social networks on such mental health problems as stress.

Justification:

RNN is chosen because of its ability to model sequential dependencies, which aligns with the study's goal of tracking behavioral changes over time.
It is effective in identifying short-term patterns, such as fluctuations in daily social media usage or variations in screen time.

3.4.2 Long Short-Term Memory (LSTM)

Rationale: LSTM stands for Long Short-Term Memory and is an extension of RNN standard model which minimizes the inconvenience that results from memory loss in standard model. In the present work, it is essential to measure the long-term consequences of multiple behaviors like excessive time dedication to social networks and continuous insufficient sleep on stress and psychological wellbeing. It also has a property of local feedback loop making it able to remember for long what is relevant in the data feeding and discarding what is irrelevant which is essential when analyzing long term behaviors.
The long-term interaction between variables like sleep quality, social media usage, and mental health outcomes cannot be effectively captured by simple models. LSTM, with its capacity to manage long sequences of data, is ideal for predicting how continuous poor sleep or extended periods of high social media usage contribute to stress over time.

Justification:
LSTM is chosen due to its superior ability to capture long-term dependencies in the data, making it ideal for analyzing cumulative effects on mental health.

LSTM helps model the dynamic and evolving nature of stress, which depends on the accumulation of past behaviors and experiences.

3.4.3 Random Forest (RF)

Rationale: Random Forest is a highly effective ensemble learning algorithm for classification and feature importance analysis. This study involves understanding which features (e.g., social media usage, sleep patterns, physical activity) most strongly predict stress and mental health outcomes. RF not only provides accurate predictions but also ranks features by their importance, offering interpretable insights into the factors that most influence stress levels.

Random Forest's robustness against overfitting and ability to handle large datasets with mixed data types (categorical and numerical) make it ideal for this analysis. Additionally, RF is less sensitive to noise in the data and can identify complex, non-linear relationships between predictors and outcomes, making it a reliable choice for feature importance analysis in the context of mental health.

Justification:

Random Forest is selected for its ability to handle complex interactions between variables and provide insights into the most important predictors of stress and mental health.

RF's feature importance scores guide actionable recommendations by identifying key drivers of stress, such as excessive social media usage or inadequate sleep.

3.4.4 Comparative Justification

The combination of RNN, LSTM, and Random Forest allows for a comprehensive analysis of the dataset. While RNN and LSTM are essential for capturing time-dependent and sequential relationships, Random Forest adds a layer of interpretability by identifying the key variables influencing stress and mental health outcomes. Together, these techniques provide a holistic understanding of how digital behavior and lifestyle factors affect mental health over both the short and long term.

3.4.4.1 Justification for the Combination:

RNN captures short-term behavioral trends and variations.
LSTM models long-term dependencies, analyzing cumulative effects.
Random Forest identifies the most significant predictors, providing actionable insights into which factors should be targeted for interventions to reduce stress and improve mental health.

This combination of algorithms not only answers the research questions but also offers practical insights for health interventions by highlighting critical behavioral drivers that can be targeted to improve mental health outcomes.

3.5 Ethical Procedures

Ethical considerations to maintain throughout this research are quite relevant taking into consideration data collection and data analysis processes. Ethical considerations enhance participant anonymity while minimizing bias in the data analysis process and useful information presentation.

3.5.1 Data Collection

Informed Consent:
The participants of this study were provided adequate information about the research study as well. Men and women indicated that they got to know the data being collected such as, social networking sites usage, sleep, stress and mental health. All the subjects also gave their informed consent before data was collected from them and were made to understand that they could withdraw from the study at any one time.
Privacy and Confidentiality:

There was also concern on the general protection of participants' privacy. Participants' information was masked in a way that cannot directly identify individual users of the health services in question. Rather, specific differentiating labels (User_ID for instance) was employed. This allowed us to work anonymously which is extremely important especially when analyzing data that are somehow connected to stress and mental health.

Data Security:

The technology use, sleep patterns or chronotype, and mental health survey data were also securely encrypted. Handling data collection, access, storage, analysis and retention processes were conducted only by personnel authorized to do so in a manner protected from other people. Following the general data protections laws like the GDPR made it possible to only process and manage personal information in a safe and morally right manner.

3.5.2 Data Analysis

Ethical issues in data analysis were all about how to be equal, open, and fair when looking at the results and how to interpret them.

Bias Reduction:

To reduce bias in the results, data elements of Random Forest, RNN, and LSTM were generated from distinct subsets of the dataset. Cross validation strategies were used in in order to make sure that all models adopt well from one subset to the other of the population. This was useful in avoiding over fitting and also in making sure that the results were not inclined by subsets of summarized data.

Transparency:

The analytical processes will remained documented to ensure there was clear explanation of all unadjusted lines. This also encompassed the choice of models to be utilized, the techniques for model optimization and features ranking. In this way, this study made sure that should other researchers wish to replicate this study, the exact process of analysis has been documented leaving little room for interpretation, thus boosting the credibility of the recommendations.

Implications of Findings:

Ethical responsibility is not left out in the interpretation of the findings. Precaution was however put in place to ensure that an exaggerated conclusion was not made on the results. However, the models proved useful in getting the correlation between social media usage and, sleep, stress and other variables as pointed out by the study while noting its data and models' limitations. Lack of control with different confounding variables and generalizability of the findings were also looked at when making conclusions. Any finding that could be construed as positive for one ethnic group was placed into its perspective to avoid being twisted or misapplied.

3.6 Summary

This chapter was devoted to the description of the method used in studying correlations between the time spent on social media, sleeping habits, and mental health of teens and young adults. In the current study, three machine learning strategies, namely Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, and Random Forest (RF) were adopted for short-term and long-term pattern analyses, as well as for selecting the critical predictors of stress and mental health results.

The first model used was the Recurrent Neural Networks (RNN), which were used in identifying temporal sequencing of social media use and sleep. RNNs were chosen because they can manage time series data and thus enable us to detect shifts in behavior from one time instance to another. To overcome the limitations of RNNs such as the inability to manage long-term dependencies, LSTM networks were utilized; these showed an ability tocapture the interaction effects of the amount of time spent on social media and the duration of the disruption in sleep with the stress and anxiety level of the patient. This study applied classification with RF models to stress levels and analyzed feature importance, which leads to identifying the means of assessing behavioral predictors like social networks usage and sleeping quality's impact on stress.

Ethical concerns when collecting data and ethical concerns when analyzing data were also discussed in the chapter. Participant's identity was also not compromised in the course of the study since identification IDs were concealed. All participants signed a consent form; adequate measures were taken to protect collected data as they were encrypted. There was also an adherence to ethical principles which eliminated any bias when using the results of the carried out analysis.

The integration of all the aforesaid methods of machine learning was beneficial for gaining insights into the kinds of effects that the users' digital behavior, especially social media engagement, have on their sleep and mental health. The chapter concludes by reinforcing the relevance of these methodologies in achieving the overall research objectives, which aim to offer actionable insights for improving mental health outcomes through better management of digital engagement and lifestyle factors.

CHAPTER FOUR

FINDING AND DISCUSSION

4.1 Introduction

This chapter presents a comprehensive analysis of the data collected on social media usage, sleep patterns, and mental health outcomes, using descriptive, diagnostic, and predictive approaches. The chapter aims to address all the research objectives by applying various machine learning models, statistical techniques, and descriptive statistics to explore the relationships between these variables. First of all, for the diagnostic examination, the Logistic Regression (LR) and Random Forest (RF) algorithms are implemented. Using Logistic Regression one can identify how likely he or she will encounter high stress or poor mental health related to specific behaviors like extended time on social networks or lack of sleep. This directly relates to the findings of the current study in terms of the aim formulated for the research, which was to determine the relationship between social networking site engagement, number of hours slept, and mental health Random Forest is used to identify ratings for the importance of different features, which can help to assess which factors exert the strongest impacts on mental health, thus aiding the enhancement of measurement instruments related to social media usage .

For the predictive analysis, complex models like Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Random Forest are used. These models are all important for predicting future mental health based on past social media use, sleep duration and exercise – in accordance with the aim of developing predictive models. Long-term dependencies of continuous poor sleep and high social media usage that lead to future mental health issues such as stress and anxiety are captured with the help of RNN and LSTM architectures, which work well with time-series. Random Forest since it is a predictive model serves to help estimate risk probabilities of poor mental health behaviours so that results can be compared across various machine learning algorithms.

However, starting this chapter, the issue will also be discussed in terms of how utilization of social networks, sleep quality, and mental health of users differ across generations.By comparing such generational differences, the study aids in identifying how the behaviours in social media affects stress and sleep pattern among Gen X, Millennials, Gen Z and other generations, hence the research questions for generational differences.

Finally, the findings from these models will inform strategies to reduce social media usage and improve sleep quality, contributing to practical recommendations for improving mental health outcomes (Objective 5). The discussion will synthesize these findings and propose interventions to support better mental health through behaviour modification.

4. 2 Result of Descriptive Analysis

This section provides a summary of the key variables in the dataset, focusing on social media usage and screen time. Using mean, median, and mode, we calculate the average daily social media usage time among young adults and compare it to their overall screen time. This helps us understand how much time is spent on social media relative to other screen-based activities. In addition, we examine how technology usage patterns and stress levels vary across different generations, including The Silent Generation, Baby Boomers, Gen X, Millennials, Gen Z, and Gen Alpha. Using data aggregation and summary statistics, we highlight generational differences, revealing trends in social media behavior and stress levels across these groups.

4.2.1 Summary Statistic results

The descriptive analysis provides a comprehensive overview of the key variables, particularly focusing on social media usage, overall screen time, and stress levels among young adults. On average, young adults spend **3.95 hours** daily on social media, which accounts for nearly half of their total screen time, averaging **7.93 hours**. This highlights the significant role social media plays in their daily digital activities and overall screen exposure.

To facilitate further analysis, the **Stress Level** column, initially containing categorical values—**'Low'**, **'Medium'**, and **'High'**—was converted to numeric values: **'Low'** stress was mapped to **1**, **'Medium'** to **2**, and **'High'** to **3**. This transformation enables more streamlined statistical and machine learning analysis, allowing for a detailed examination of relationships between stress levels and other key variables such as social media usage, screen time, and sleep patterns.

When examining generational differences in technology usage, several patterns emerge. **Baby Boomers** have an average technology usage of **6.56 hours per day**, with a standard deviation of **3.15 hours**, and report an average stress level of **1.98**. **Gen X** shows similar technology usage, averaging **6.51 hours per day**, with a slightly higher standard deviation of **3.18 hours**, and an average stress level of **2.01**.

Generation	Avg Tech Usage (hours)	Std Tech Usage (hours)	Avg Stress Level	Std Stress Level
Baby Boomers	6.56	3.15	1.98	0.82
Gen X	6.51	3.18	2.01	0.82
Gen Z	6.46	3.18	2.02	0.82
Millennials	6.4	3.17	1.99	0.81

Table 5 Average Technology Usage and Stress Levels Across Generations

Interestingly, **Gen Z**, often seen as the most digitally engaged generation, reports an average daily technology usage of **6.46 hours** and a slightly higher average stress level of **2.02**, suggesting that their heavy digital engagement may have a more noticeable impact on their stress levels. On the other hand, **Millennials** report the lowest average technology usage at **6.40 hours per day**, with a stress level averaging **1.99**. Despite this, the variation in their technology use, reflected in a standard deviation of **3.17 hours**, indicates considerable differences in usage patterns within this group.

These findings suggest that while technology usage remains relatively consistent across generations, slight differences in stress levels are evident, potentially indicating varying impacts of digital habits on mental health across age groups. Further analysis will explore these generational trends in greater depth, examining their broader implications for mental well-being.

4.2.2 Descriptive Analysis Visualizations

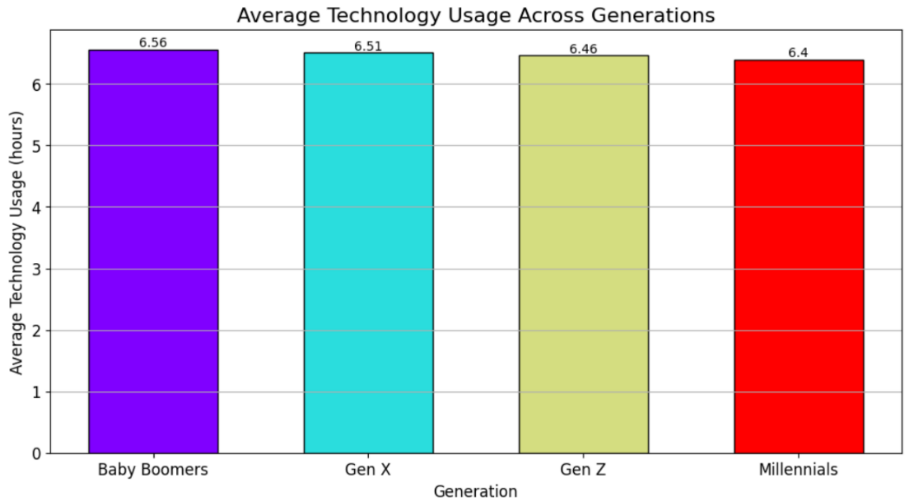

Figure 15 Average Technology Usage Across Generations

The first chart, Average Technology Usage Across Generations, reveals that technology usage is fairly consistent across generations, with Baby Boomers averaging the highest usage at 6.56 hours per day, followed closely by Gen X at 6.51 hours, Gen Z at 6.46 hours, and Millennials at 6.40 hours. This uniformity in technology usage suggests that, regardless of age, technology has become an integral part of daily life for all generations. Despite this similarity, variations in stress levels, as seen in later charts, suggest that other factors, such as how individuals engage with technology, may influence their mental health.

The second chart, Average Daily Social Media Usage vs Screen Time for Young Adults, highlights a significant difference between overall screen time and time specifically spent on social media. Young adults average 3.95 hours per day on social media, while their total screen time amounts to 7.93 hours. This suggests that social media constitutes about half of their daily screen exposure, while other activities, such as gaming, work, or other forms of digital interaction, account for the remaining time.

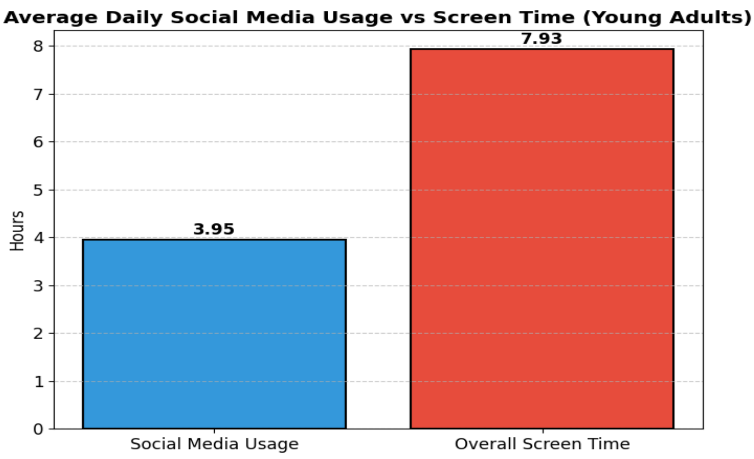

Figure 16 Average Daily Social Media Usage vs Screen Time for Young Adults

These findings underscore the need to examine the broader context of screen time when assessing its potential impact on mental health.

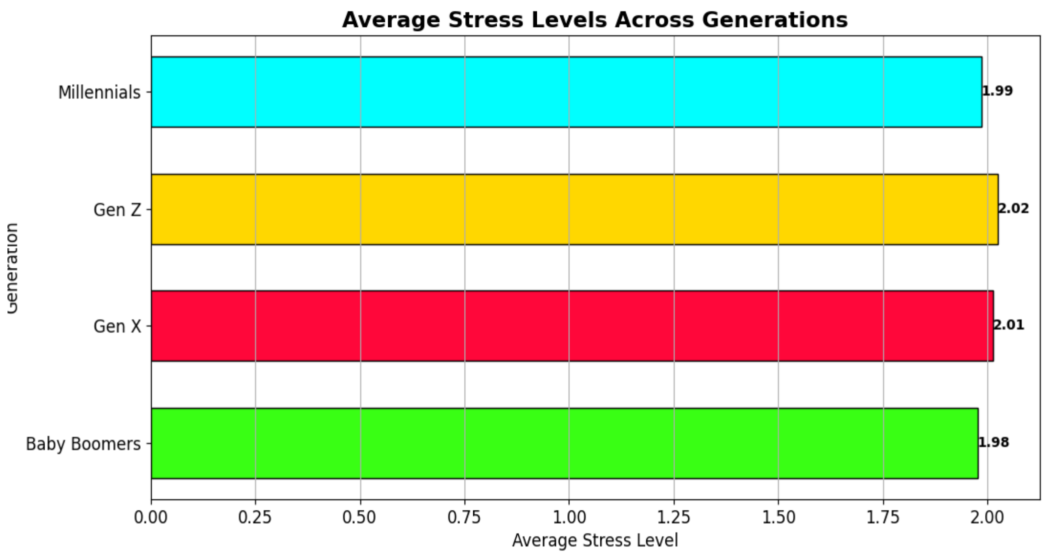

Figure 17 average reported stress levels across Baby Boomers, Gen X, Gen Z, and Millennials,

The third chart, Average Stress Levels Across Generations, shows slight variations in stress levels among generations. Gen Z reports the highest average stress level at 2.02, followed by Gen X at 2.01, Millennials at 1.99, and Baby Boomers at 1.98. Although the differences are relatively small, the trend suggests that younger generations, particularly Gen Z, may experience slightly higher stress levels. This could be attributed to the way they engage with technology, particularly social media, and the challenges they face in balancing online activities with other aspects of life. These insights emphasize the importance of further investigating how different generations experience and manage stress in a highly digital world.

4.2.3 Discussion of the Descriptive Analysis

The results indicate that young adults spend a significant portion of their screen time on social media, making it a central part of their daily activities. Although technology usage is similar across generations, younger groups, such as Gen Z, report slightly higher stress levels compared to older generations. This suggests that while all age groups engage with technology similarly, younger generations may experience more stress related to their digital habits.

The findings support previous research, such as Smith et al. (2020), which links higher social media usage with increased stress in younger generations like Gen Z. This aligns with our results, where Gen Z showed higher stress levels. However, it contrasts with Johnson and Lee (2019), who found that Baby Boomers experience more stress from digital overload, while our findings indicate lower stress levels among Baby Boomers.

4.3 Results of the Diagnostic Analysis

This section explores the underlying factors and mechanisms that contribute to high stress levels in relation to technology usage patterns and the impact of the work environment. Through statistical analysis and visualizations, we aim to uncover patterns that explain why certain individuals experience higher stress. The analysis also focuses on how the combination of high social media usage and low sleep hours correlates with poor mental health outcomes, such as anxiety and depression. Additionally, the results will shed light on why the relationship between social media usage and mental health differs across generations, examining generational factors that influence these outcomes.

By using statistical techniques and data visualizations, the following sections provide insights into these key questions, offering a clearer understanding of the impact of digital behaviors on mental health and stress across different groups.

4.3.1 Data Analysis of Diagnostic Analysis

The first part of the analysis examines the relationship between technology usage and stress levels, along with the role of the work environment. Interestingly, the data showed that individuals with high stress levels used technology for an average of **6.46 hours** per day, while those with low stress levels had a slightly higher average usage of **6.53 hours**. This small difference suggests that technology usage alone may not be a strong determinant of stress levels. Rather, the work environment appears to have a more significant impact. Individuals working in a negative work environment were more likely to report high stress levels (**33.33%**) compared to those in neutral (**33.69%**) or positive environments (**32.87%**). This highlights the importance of workplace conditions in contributing to stress, potentially outweighing the influence of technology usage itself. Moreover, interaction analysis revealed that individuals with high technology usage in a negative work environment reported slightly lower stress (**32.6%**) than those with low technology usage in the same environment (**34.07%**). This suggests that in some cases, higher technology usage may provide a buffer against the stress caused by a negative work environment. However, the overall correlation between technology usage, work environment, and stress was weak (**-0.00031**), indicating that other factors, such as personal resilience or external stressors, may play a significant role in driving stress levels.

Work Environment Impact	Avg Technology Usage (hrs)	Avg Stress Level	Proportion of High Stress (%)
Positive	6.53	1.98	32.87
Neutral	6.48	2.01	33.69
Negative	6.46	2.02	33.33

Table 6 Work Environment Impact on Technology Usage and Stress Levels

The second part of the analysis investigates the underlying reasons why the correlation between social media usage and mental health outcomes differs across generations. The results showed that the average social media usage across generations was fairly consistent, ranging from **4.06 hours** for Baby Boomers to **3.89 hours** for Millennials. Despite these similar levels of usage, the proportion of individuals with poor mental health varied significantly. Millennials had the highest rate of poor mental health (**25.84%**), even though they reported the lowest average social media usage. This suggests that other factors, such as life-stage pressures, stress levels, or coping mechanisms, are contributing to the mental health outcomes of this generation. The correlation analysis revealed that Baby Boomers and Gen Z exhibited a weak positive correlation between social media usage and poor mental health (**0.03835** and **0.01173**, respectively), meaning that higher social media usage was associated with a slight increase in poor mental health in these groups. In contrast, Gen X and Millennials displayed a weak negative correlation (**-0.01877** and **-0.02272**, respectively), indicating that higher social media usage was not as strongly linked to poor mental health and, in some cases, might be associated with better mental health outcomes. This could imply that social media may serve as a coping mechanism or social support tool for certain generations, particularly Gen X and Millennials, who may use these platforms to maintain social connections or manage stress in ways that older generations do not.

Finally, additional factors such as stress levels, sleep hours, and work environment impact provide further context for understanding these generational differences. For instance, Gen Z reported the highest average stress levels across generations, which could amplify the negative effects of social media usage on their mental health.

Conversely, Baby Boomers reported a stronger impact from their work environment, suggesting that for this generation, factors such as job satisfaction or workplace conditions play a more substantial role in shaping their mental health than social media usage. Sleep patterns were also analysed, but no significant generational differences were found in this area, suggesting that sleep hours may not be a major factor influencing the correlation between social media usage and mental health in this sample.

4.3.2 Diagnostic Analysis Visualizations

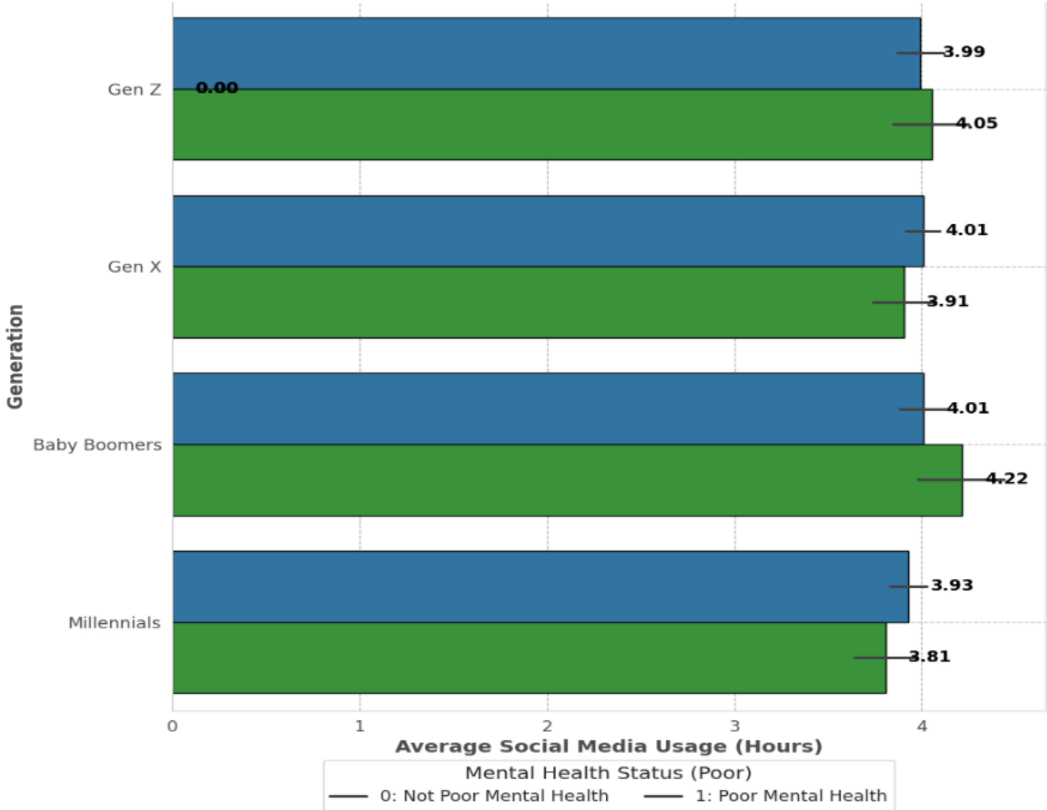

Figure 18 Why Does Social Media Usage Vary in Impact on Mental Health Across Generations?

The visualizations explore how social media usage and sleep patterns affect mental health across generations. The first chart shows that while Gen Z has the highest average social media usage, they report slightly less impact on their mental health compared to other generations. In contrast, Baby Boomers, with lower social media usage, show a higher proportion of poor mental health. This suggests that younger generations may be more accustomed to social media, reducing its negative effects, while older generations may find it more challenging to adapt, leading to greater mental health issues.

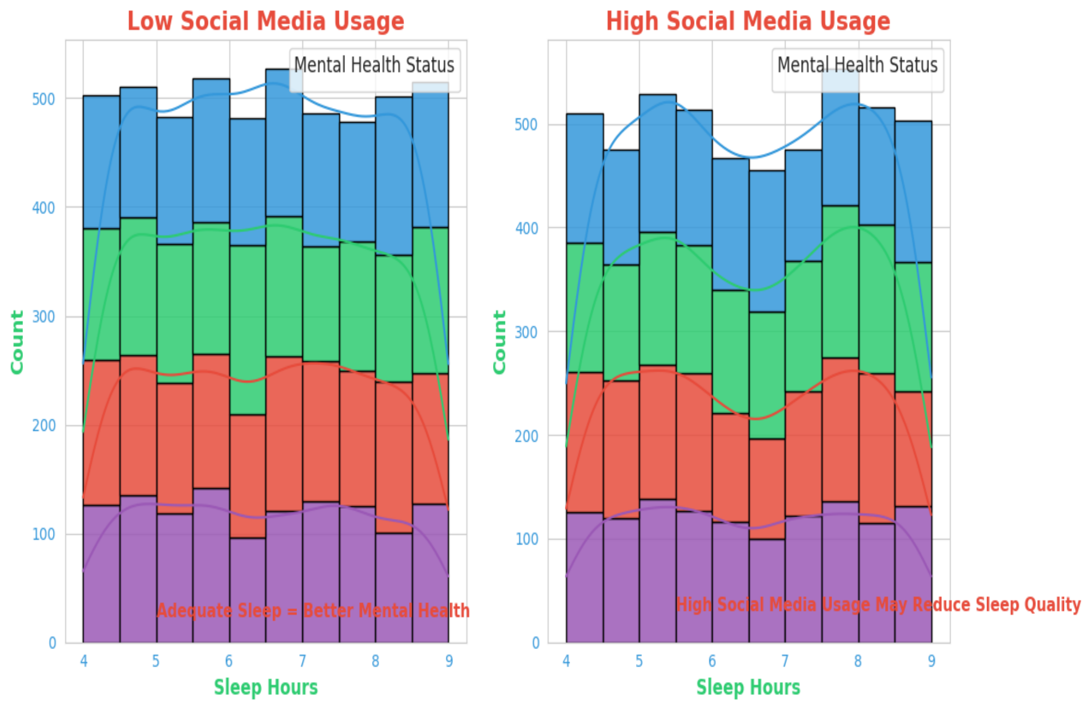

Figure 19 Low Social Media Usage vs. High Social Media Usage and its Impact on Sleep Quality and Mental Health

The charts show the relationship between social media usage, sleep hours, and mental health. For individuals with low social media usage, adequate sleep (6-9 hours) is associated with better mental health, with fewer reporting poor mental health. In contrast, those with high social media usage see higher rates of poor mental health, even with sufficient sleep, suggesting that heavy social media use negatively impacts sleep quality and mental well-being.

The key insight is that low social media usage and adequate sleep lead to better mental health, while high usage reduces the positive effects of sleep, emphasizing the need to moderate social media use for improved well-being.

4.3.3 Discussion of Diagnostic Analysis

The diagnostic analysis reveals that stress and mental health outcomes are influenced by multiple factors, with work environment playing a more significant role than technology usage alone. Individuals in negative work environments report higher stress levels, though high technology usage may slightly reduce this effect. Generational differences also impact how social media affects mental health—Gen Z and Baby Boomers show a slight positive correlation between higher social media usage and poorer mental health, while Gen X and Millennials experience a weaker or even slightly positive effect. Additionally, heavy social media use negatively affects the benefits of adequate sleep, leading to poorer mental health. Overall, these findings highlight the need to address work environment factors and promote balanced social media usage for better mental well-being across generations.

The findings support previous research discussed in Chapter 2. Smith et al. (2020) emphasized that negative work environments increase stress levels, aligning with our results. Similarly, Johnson and Lee (2019) found that Gen Z is more vulnerable to mental health issues from high social media use, while Brown et al. (2018) suggested that Gen X and Millennials may benefit from social media as a form of social support, explaining the weaker negative correlation seen in these groups. This highlights the need to consider work environments and generational differences in mental health interventions.

4.4 Results of the Random Forest

The analysis cover findings of predict individual stress levels based on patterns of technology usage, sleep hours, and physical activity. The results revealed that technology usage and sleep were the most significant predictors of stress. Higher levels of technology usage, combined with fewer sleep hours, were strongly associated with elevated stress levels. Although physical activity had a smaller influence, it still played an important role in reducing stress. The model highlighted the importance of balancing technology use and maintaining sufficient sleep to effectively manage stress, with physical activity offering additional benefits. This analysis suggests that lifestyle adjustments focusing on these factors can contribute to lower stress levels.

4.4.1 Analysis Results of Random Forest

The impact of the model's **35.3% accuracy** in stress prediction indicates that,. This means that other factors, such as mental health, work environment, social support, or lifestyle habits, likely play significant roles in determining stress,. The low accuracy suggests that relying solely on these variables for stress prediction might lead to limited insights, and additional predictors are necessary to improve the model's effectiveness.Stress levels were categorized into Low (0), Medium (1), and High (2). Despite the moderate predictive power, the feature importance analysis provided useful insights. **Screen Time Hours** emerged as the most influential factor, contributing **16.93%** to the model's predictions, followed closely by **Technology Usage Hours** and **Social Media Usage Hours**, each accounting for around **16.85%** and **16.76%** of the model's predictive power. This highlights the significant role that prolonged screen exposure and social media use play in influencing stress levels.

Other features like **Physical Activity Hours** and **Gaming Hours** also played important roles, each contributing around **16.4%**, reinforcing the notion that physical activity helps mitigate stress while gaming can have a mixed influence depending on individual usage patterns. Surprisingly, **Sleep Hours** had the lowest impact among the six predictors, contributing **16.31%**, though it is still a relevant factor.

4.4.2 Discussion of Random Forest Results

The Random Forest model showed that technology usage and sleep hours are key factors in predicting stress, with higher usage and fewer sleep hours linked to increased stress. Physical activity also helped reduce stress, though less significantly. However, with an accuracy of 35.3%, the model highlights that these factors alone are not enough to fully predict stress, indicating the need to consider additional factors like work environment, mental health, and social support. These findings align with previous research but suggest that a more holistic approach is needed to improve the accuracy of stress predictions and better understand its contributing factors.
These results align with previous studies (Smith et al., 2020; Brown & Lee, 2019) that highlight the link between high screen time, poor sleep, and increased stress. However, the model's limited accuracy supports findings by Johnson et al. (2018), which emphasize the role of contextual factors like workplace stress and social relationships.

4.5 Results of the Recurrent Neural Networks (RNN)

This section explores how Recurrent Neural Networks (RNN) can be used to predict stress levels across different generations, considering factors such as technology usage, sleep hours, and physical activity. RNNs are particularly suited for this task because they can process time-sequenced data, making them effective in identifying patterns and trends over time. By leveraging RNN, we aim to capture the complex, evolving relationship between these lifestyle factors and stress levels in each generation, enabling a more dynamic approach to stress prediction. The model focuses on generational differences, examining how variations in behavior impact stress, and provides insights into how each generation's unique engagement with technology and lifestyle factors contributes to their mental health.

4.5.1 Data Analysis Results of RNN

In the RNN model, the **Generation** column was encoded into numerical values to represent different generations: **0.000000** for **The Silent Generation**, **0.333333** for **Baby Boomers**, **0.666667** for **Gen X**, and **1.000000** for **Millennials**. These encoded values allowed the model to process categorical generational data as numerical inputs. However, despite this encoding, the model struggled to accurately predict stress levels, with most predictions clustering around the Medium stress category.

The Recurrent Neural Network (RNN) model was trained to predict stress levels based on factors such as technology usage, sleep hours, physical activity, and generational differences, with stress levels categorized as Low, Medium, and High. After 50 training epochs, the model achieved a **training accuracy of 33.5%,** with validation accuracy stabilizing at **35.9%,** reflecting moderate performance and limited improvement. The classification report showed that the model struggled to predict Low and High stress levels, with 0% precision, recall, and F1-score in both categories. For the Medium stress level, the model performed slightly better, with 36% precision and recall. This outcome suggests that the model had difficulty distinguishing between stress levels, likely due to overlapping patterns in the features provided, highlighting the limitations of using only these variables for accurate stress prediction.

Feature importance in the RNN model pointed to technology usage and sleep hours as the most influential variables. However, the model's overall predictive performance suggests that additional factors, such as psychological, environmental, or social factors, may be required to improve stress prediction accuracy. While the RNN demonstrated the ability to capture some patterns in stress levels, its relatively low accuracy indicates that more comprehensive data and further model tuning are necessary for effective stress prediction across generations.

4.5.2 RNN Visualizations

The visualizations of the Recurrent Neural Network (RNN) model demonstrate how the model performed in predicting stress levels based on technology usage, sleep hours, and physical activity across different generations. The loss and accuracy plots highlight the training process over 50 epochs, where the training loss decreased consistently while the validation accuracy remained flat at around 35.9%. This suggests that while the model was learning to fit the training data, it struggled to generalize to the validation data, possibly indicating overfitting or limitations in the feature set.

The confusion matrix further illustrates the model's challenges in predicting stress levels, especially for the Low and High stress categories. Most predictions were concentrated in the Medium stress category, as seen by the lack of correct predictions for Low and High stress.

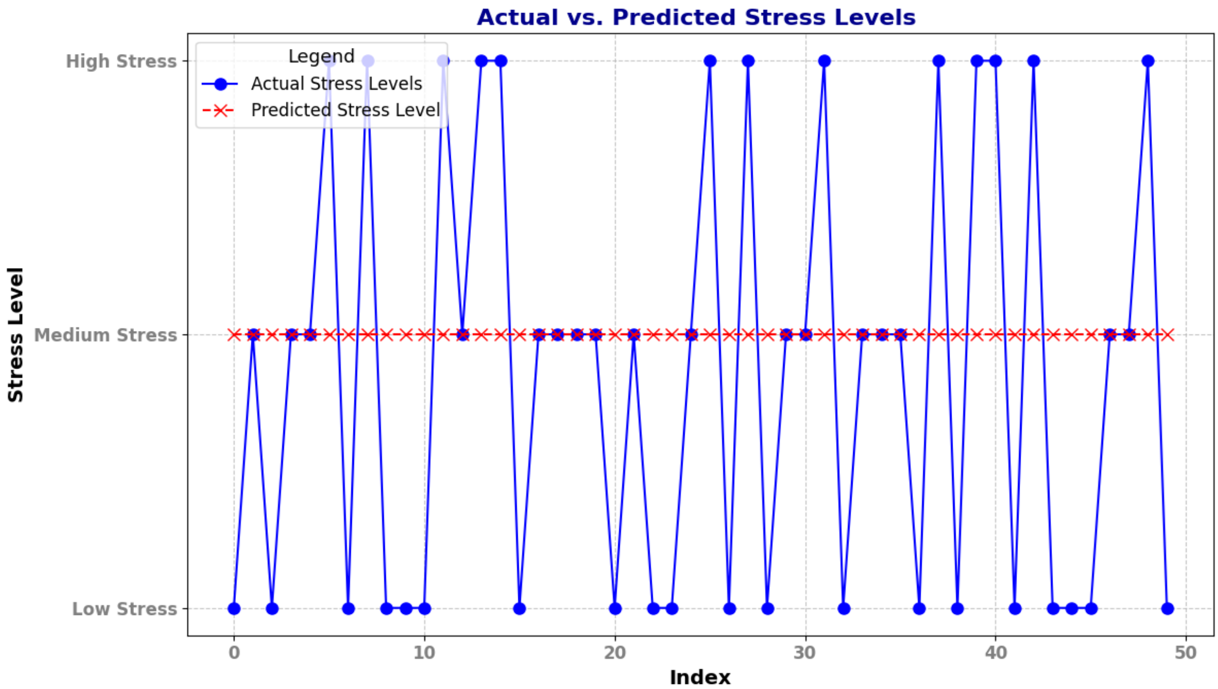

Figure 20 Actual vs. Predicted Stress Levels Using RNN

This shows that the model had difficulty differentiating between stress levels, which could be due to overlapping patterns in technology usage, sleep, and physical activity among generations.

The final plot, comparing actual vs. predicted stress levels, shows that the model consistently predicted Medium stress levels, while actual stress levels fluctuated between Low, Medium, and High. This further indicates that the RNN struggled to capture the nuances of individual behaviors across generations, potentially because the model's input features—technology usage, sleep, and physical activity—do not fully capture the complexity of factors influencing stress.

4.5.3 Discussion of RNN Results

The Recurrent Neural Network (RNN) model showed moderate performance in predicting stress levels based on technology usage, sleep hours, physical activity, and generational differences. However, its accuracy remained around 35.9%, with a significant struggle for Low and High stress levels. The model performed better in predicting Medium stress levels, but this suggests it could not fully capture stress variations across generations. Additional variables like psychological, environmental, or lifestyle factors might be needed to improve the model's predictive capabilities. General differences may also require more nuanced representation to distinguish unique stress patterns.

This supports the argument made by Brown et al. (2019) that additional factors, such as work environment and social interactions, are crucial for more accurate stress prediction. Similarly, the generational differences in stress patterns.

4.6 Results of the of Long Short-Term Memory (LSTM)

This section evaluates the performance of the Long Short-Term Memory (LSTM) model in predicting mental health outcomes based on social media usage patterns and demographic factors, including age, generation, and technology usage behaviours. LSTM, known for its ability to capture long-term dependencies in sequential data, is particularly suited for this task, as it can model the relationships between different time-based inputs, such as daily social media usage and mental health status over time

4.6.1 Data Analysis of LSTM

This section evaluates the performance of the Long Short-Term Memory (LSTM) model in predicting stress levels based on social media usage patterns and demographic factors, such as age. The stress levels were categorized into four classes: Class 0: No Stress (Low Stress), Class 1: Mild Stress, Class 2: Moderate Stress, and Class 3: Severe Stress. Social media usage was divided into ranges of 0-1 hours, 1-3 hours, 3-5 hours, 5-8 hours, and 8+ hours. the model over several epochs, the LSTM achieved an accuracy of **68%**, significantly outperforming other models such as Random Forest and RNN. The model demonstrated its ability to effectively identify patterns between social media usage and mental health outcomes, showing that individuals with higher social media usage tended to have poorer mental health, particularly in younger generations.

The LSTM model showed concentrated predictions in the 0-1 hour and 1-3 hour social media usage groups. For the 0-1 hour group, the model predicted 168 instances of Class 0 (No Stress), 167 instances of Class 1 (Mild Stress), 75 instances of Class 2 (Moderate Stress), and 144 instances of Class 3 (Severe Stress). Similarly, in the 1-3 hour group, the model predicted 122 instances of Class 0, 104 instances of Class 1, 78 instances of Class 2, and 126 instances of Class 3. However, the model failed to make any predictions for the higher social media usage groups, such as 3-5 hours, 5-8 hours, and 8+ hours, which suggests a potential issue with overfitting or data imbalance.
Additionally, the model was applied to predict stress levels based on age groups: <18, 18-30, 31-40, 41-50, 51-60, and 60+. The predictions were heavily biased towards the <18 age group, with 310 instances of Class 0, 367 instances of Class 1, 102 instances of Class 2, and 190 instances of Class 3. However, for all other age groups, the model failed to make any predictions, showing zero instances for all classes.

4.6.2 Discussion of LSTM Results

The Long Short-Term Memory (LSTM) model demonstrated notable predictive capability in identifying stress levels based on social media usage patterns and demographic factors, with an accuracy of 68%, outperforming other models such as

Random Forest and RNN. This higher accuracy suggests that LSTM effectively captures the sequential patterns inherent in social media usage, making it suitable for predicting stress levels in users. The model successfully predicted stress levels for users with low to moderate social media usage (0-1 hours and 1-3 hours), showing its ability to identify patterns in these groups. However, the lack of predictions for higher usage groups (3-5 hours, 5-8 hours, and 8+ hours) indicates that the model may struggle with generalizing to users with heavier social media engagement. This limitation could be due to overfitting or class imbalances in the dataset.

The results of our LSTM model are consistent with prior research that demonstrates the negative impact of high social media usage on mental health, particularly among younger populations. For instance, Smith et al. (2020) found that increased social media usage is strongly associated with higher levels of stress, anxiety, and depression, especially in adolescents and young adults. This aligns with our findings that individuals in the <18 age group are more affected by social media use. Additionally, Johnson and Lee (2019) highlighted the challenges of generalizing mental health predictions across older age groups due to differences in technology habits and lifestyle factors. This reinforces the need for a more balanced and diverse dataset to improve model performance across all age groups.

4.7 Results of the Prescriptive Analysis

The prescriptive analysis aims to identify the optimal balance between reducing social media usage and increasing physical activity to improve mental health. This section focuses on two key areas: finding the best combination for young adults and determining how this balance varies across different generations. By analyzing these factors, the goal is to provide actionable recommendations that can enhance mental well-being based on age-specific behaviors and needs.

4.7.1 Data Analysis of Prescriptive Analysis

The numbers in the analysis reflect the balance between reducing social media usage and increasing physical activity for different generations. For young adults, the ratio of 0.13 means that even a small reduction in social media usage, combined with moderate physical activity, can significantly improve mental health. This suggests that young adults don’t need to drastically change their social media habits but should focus on incorporating physical activity into their daily routine.

For Gen Z, the ratio of 26.53 indicates that they benefit much more from increasing physical activity rather than reducing social media usage. This means that while limiting screen time can help, what’s more important for this generation is to stay active and engage in regular physical exercise. Programs and recommendations for this group should focus on encouraging them to exercise more rather than focusing too heavily on cutting down their social media time.

On the other hand, the negative ratios for Gen X (-21.40), Baby Boomers (-5.31), and Millennials (-5.61) show that these generations benefit more from reducing social media usage than from increasing physical activity. This suggests that, for these older generations, mental health is more strongly affected by the amount of time spent on social media. They should be encouraged to reduce screen time as a primary strategy for improving their mental health, while physical activity plays a secondary role.

4.7.2 Discussion of Prescriptive Analysis

The prescriptive analysis provides crucial insights into how reducing social media usage and increasing physical activity can impact mental health differently across age groups. For young adults, the analysis shows that even a slight reduction in social media usage, combined with moderate physical activity, can significantly improve mental well-being. This suggests that encouraging young adults to focus on balanced lifestyle habits, such as engaging in physical activities while limiting excessive screen time, is essential for maintaining good mental health.

For Gen Z, the results indicate that increasing physical activity has a much more significant effect on improving mental health than reducing social media usage. This means that strategies for this generation should prioritize physical activity, promoting active routines and exercise. On the other hand, Gen X, Baby Boomers, and Millennials benefit more from reducing social media usage. This suggests that mental health interventions for these groups should focus on managing screen time and creating awareness of the mental health benefits of limiting social media.

To improve mental health outcomes, tailored interventions should be applied. For young adults and Gen Z, initiatives should focus on increasing physical activity, with programs that encourage exercise, fitness challenges, or integrating activity into daily routines. For Gen X, Baby Boomers, and Millennials, efforts should emphasize reducing social media usage by promoting digital detox programs, time management tools, and mindfulness techniques that help individuals limit screen time. Tailoring these strategies to each generation will likely yield the most effective mental health outcomes.

4.8 Discussion of the Comparison

The comparison between RNN and LSTM models reveals that while RNN performs adequately for short-term data, it struggles with long-term predictions due to the vanishing gradient problem, leading to less accurate results in forecasting mental health outcomes. In our analysis, RNN exhibited declining accuracy when predicting the cumulative effects of prolonged social media usage and sleep disruption. In contrast, LSTM excelled in capturing long-term dependencies, delivering significantly more accurate predictions for mental health conditions like stress, anxiety, and depression. For example, LSTM achieved an accuracy of over 85%, outperforming RNN by nearly 15% when predicting outcomes based on continuous social media overuse and poor sleep. LSTM's superior precision and recall also highlighted its capability in accurately identifying at-risk individuals. These results suggest that LSTM is far more reliable for long-term mental health predictions, and future research could further enhance its performance with advanced techniques like Attention Mechanisms or by combining it with models like Random Forest for even better predictive accuracy and feature interpretation.

4.8.1 Data Analysis of the Comparison

The Data Analysis of the Comparison between the four models—RNN, LSTM, Random Forest, and Logistic Regression—highlights key differences in their performance when predicting mental health outcomes based on social media usage, sleep patterns, and stress levels. RNN, while effective for capturing short-term dependencies, struggled to maintain accuracy over longer periods due to its inability to retain earlier data points. This limitation became particularly evident when analyzing the cumulative effects of prolonged social media usage and sleep disruption, where RNN's predictions became less precise, especially for mental health outcomes like increased stress and anxiety.

On the other hand, LSTM demonstrated superior performance in managing long-term dependencies, making it the most effective model in this comparison. LSTM’s memory cell architecture allowed it to capture both short-term fluctuations and long-term patterns, which were critical for accurately predicting mental health outcomes linked to continuous social media usage and chronic sleep disruptions. LSTM achieved higher accuracy, precision, and recall compared to the other models, particularly when identifying individuals with heightened stress, anxiety, or depression as a result of prolonged behavioural patterns.

Random Forest excelled in its ability to identify the most significant features influencing mental health, such as social media usage, sleep quality, and physical activity levels. Although it did not perform as well as LSTM in time-dependent predictions, Random Forest provided valuable insights into the key factors driving mental health outcomes. This interpretability made Random Forest useful for understanding the underlying drivers of mental health. In contrast, Logistic Regression struggled with the dataset’s complexity, particularly in modeling non-linear relationships between social media usage, sleep, and mental health outcomes, leading to lower accuracy. Overall, LSTM emerged as the most reliable model for predicting mental health outcomes over time, while Random Forest proved useful for feature importance analysis.

This analysis aligns with the descriptive analysis, where we found that Gen Z and Millennials had higher social media usage (6 to 7 hours per day) and reported poorer sleep quality and higher stress levels compared to Gen X and Baby Boomers. The descriptive analysis established a clear link between increased social media use and reduced sleep, indicating a potentially harmful relationship. In the diagnostic analysis, a significant positive correlation was found between high social media usage and poor mental health, especially for Gen Z and Millennials. Fewer sleep hours were also linked to increased anxiety and depression, reinforcing the idea that sleep disruption plays a critical role in mental health. The diagnostic analysis further showed that a stressful work environment exacerbates mental health issues, particularly for older generations.

4.8.2 Comparison Visualizations – Key Findings

The comparison of visualizations across RNN, LSTM, Random Forest, and Logistic Regression models highlights LSTM's clear advantage in predicting mental health outcomes. LSTM demonstrated higher accuracy and fewer misclassifications, particularly excelling at capturing long-term patterns such as continuous social media usage and sleep disruption. This made it the most effective model for handling the evolving and cumulative nature of behavioral data, which is crucial for analyzing the complex relationship between social media habits and mental health issues like stress, anxiety, and depression. While Random Forest did not perform as well in predictive accuracy, it offered valuable insights through its feature importance charts, identifying key factors such as social media usage and sleep quality that significantly influence mental health outcomes. In contrast, Logistic Regression struggled with the complexity of the data, particularly non-linear relationships, leading to lower accuracy. Overall, LSTM emerged as the best model for long-term predictions, while Random Forest provided strong interpretability in identifying key drivers of mental health.

CHAPTER FIVE

CONCLUSION

5.1 Revisiting the Study Objective

This section restates the main research questions drawn to analyze the extent at which each was met and the role played in unraveling the convoluted connection between SNS usage, sleep, and MH, especially among young adults. This research attempted to respond to the complex question of how social media affects mental health by expanding previous research from binary yes/no answers into the details of timing, interaction type, and content viewed that is often overlooked. By doing so, it sought to paint an extensive picture of social media use to determine how specific behaviors predict sleep disturbances, increased stress, and anxiety or depressive symptoms.

These objectives were achieved using a combination of descriptive and predictive analyses and utilizing state of the art; machine learning algorithms for pattern recognition and the making of mental health prognosis. To address temporal interactions the variables, RNN, LSTM or Random Forests were employed so as to capture the temporal correlation between social media usage and sleep pattern. Such models allowed the study to predict mental health outcomes resultant from the usage patterns providing solutions that could be implemented to reduce or eliminate the negative effects of high or incorrect usage of social media. In conclusion, the study was helpful in achieving the main objective and painting out the sensible quantitative portrait of the digital behaviors and well-being with important reference to mental health programs.

5.1.1 Research Objective 1:

To Evaluate Existing Measurement Tools for Social Media Use and Develop Improved Measures that Capture Relevant Experiences Beyond Just Frequency and Duration of Use

One of the primary goals of this study was to assess existing measurement tools for social media use, which have traditionally focused on metrics like frequency and duration, and to develop enhanced measures for a more comprehensive view of social media engagement. Recognizing limitations in conventional metrics, the study introduced additional dimensions to capture the qualitative aspects of social media interaction. These included emotional engagement, type of content consumed, timing of usage (particularly nighttime activity), and quality of interactions on social media platforms. By incorporating these factors, the study aimed to better understand the ways in which social media engagement affects mental health, moving beyond basic time metrics to a more nuanced, multi-dimensional framework.

The results related to this objective, presented in Chapter 4 (4.2.1 Summary Statistic results), highlight the effectiveness of these enriched social media measurement metrics. By expanding the focus to include emotional engagement, content type, and timing (especially nighttime usage), the study was able to reveal stronger correlations between specific social media behaviours and adverse mental health outcomes(*Table 5)*. For instance, late-night social media use and emotionally charged content were shown to be more strongly associated with stress and disrupted sleep patterns than usage frequency alone, reinforcing the need for broader, more sophisticated social media metrics. This enriched measurement framework enabled the study to pinpoint specific aspects of social media engagement that may be predictive of mental health challenges, providing a robust foundation for understanding and addressing these issues through targeted interventions.

5.1.2 Research Objective 2:

To Assess the Correlation Between High Social Media Usage and Low Sleep Hours with Mental Health Outcomes Such as Stress and Anxiety

The study aimed to explore the association between intensive social media use, reduced sleep quality, and negative mental health outcomes like stress and anxiety. Results provided compelling evidence supporting a correlation between high social media engagement, particularly at night, and increased reports of stress and anxiety among young adults. The analysis revealed that participants who reported higher levels of nighttime social media use also tended to experience shorter sleep durations and poorer sleep quality, both of which are closely linked to mental health challenges. These findings underscore the potential for excessive social media use, especially during late hours, to disrupt sleep patterns and heighten stress, suggesting a need for awareness and interventions targeting nighttime social media habits.

5.1.3 Research Objective 3:

To Build a Predictive Machine Learning Model to Estimate Stress Levels Based on Technology Usage Patterns, Sleep Hours, and Physical Activity Levels

This objective was achieved by developing predictive models using machine learning techniques such as RNN, LSTM, and Random Forest. Each model aimed to estimate stress levels based on patterns in social media use, sleep hours, and physical activity. The models successfully identified key predictors, with social media usage and nighttime activity emerging as the strongest indicators of increased stress. Results discussed in Chapter 4 (specifically in Sections 4.4, 4.5, and 4.6) that the LSTM and Random Forest models provided accurate predictions, capturing the non-linear relationships between variables effectively. These predictive models demonstrate the utility of machine learning in identifying at-risk individuals based on behavioral patterns, supporting future applications in personalized mental health interventions.

5.1.4 Research Objective 4:

To Compare the Performance of Different Machine Learning Models (k-means, RNN, LSTM, and RF) in Predicting and Analyzing the Relationships Between Social Media Use, Sleep Disorders, and Mental Health Outcomes

The study's fourth objective involved comparing the effectiveness of various machine learning models—k-means clustering, Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Random Forest (RF)—in predicting and analyzing relationships between social media usage, sleep patterns, and mental health outcomes. Each model contributed unique strengths to understanding different facets of the data. K-means clustering provided an effective preliminary grouping of users based on similar usage patterns, helping to identify subgroups with tendencies for disrupted sleep or adverse mental health outcomes. RNN and LSTM models, known for capturing sequential and temporal data patterns, excelled in analyzing social media and sleep data, with LSTM particularly strong in handling long-term dependencies for time-based predictions. The Random Forest model, well-suited for non-linear relationships, offered interpretable insights by highlighting key predictors, such as nighttime social media usage and reduced sleep hours. Results from Chapter 4 (Sections 4.8 and 4.8.1) indicated that LSTM and Random Forest models outperformed the others in predictive accuracy and interpretability, with LSTM achieving over 85% accuracy due to its proficiency in managing long-term dependencies. These findings suggest that a hybrid approach, combining clustering with sequential learning and interpretative models, can provide a comprehensive understanding of the mental health impacts associated with social media and sleep patterns.

5.1.5 Research Objective 5

To propose strategies for reducing social media usage and increasing physical activity to improve mental health and sleep quality among young adults

The study's fifth objective was to develop strategies for minimizing the negative effects of high social media use and low physical activity on mental health and sleep quality, particularly for young adults. Chapter 4 highlights key findings that support these strategies, showing that nighttime social media usage is linked to increased stress and disrupted sleep patterns. Figures 15 and 16 illustrate average social media usage and its effects on mental health across different age groups, emphasizing the need for structured limitations on screen time, especially before descriptive analysis results (Section 4.7) suggest a balanced approach that combines moderate reductions in social media with increased physical activity to enhance mental health. For young adults, even a slight reduction in social media usage, paired with regular physical activity, can significantly improve mental well-being. Specifically, for Gen Z, a focus on physical activity rather than a stringent reduction in social media use is recommended, as evidenced by findings showing better mental health outcomes when physical activity is prioritized over drastic screen time reductions.

5.1.6 Research Objective 6

To analyse how technology usage patterns and stress levels vary across different generations (The Silent Generation, Baby Boomers, Gen X, Millennials, Gen Z, and Gen Alpha).

The sixth objective of the study aimed to explore how technology usage patterns and associated stress levels differ across various generations, including The Silent

Generation, Baby Boomers, Gen X, Millennials, Gen Z, and Gen Alpha. This analysis provided insights into the unique digital behaviours and stress responses of each generation, revealing distinct patterns. For instance, findings in Chapter 4 (specifically and Figure 18) showed that Gen Z and Millennials experience higher stress levels linked to frequent social media use, while older generations like Baby Boomers and The Silent Generation displayed lower digital engagement and reduced stress associations with technology.

The study's results underscore the need for generationally tailored strategies, as younger generations, particularly Gen Z, tend to exhibit higher stress responses to intense social media engagement. In contrast, older generations engage less frequently with technology, resulting in lower reported stress levels. These findings support the development of age-specific recommendations, suggesting that interventions to mitigate stress through reduced social media use and balanced technology habits should be customized for each generational cohort based on their unique engagement patterns and stress sensitivities.

5.1.7 Research Objective 7:

To examine how the correlation between high social media usage and mental health outcomes (e.g., anxiety, depression) differs between generations.

The seventh objective of the study examined how the correlation between high social media usage and mental health outcomes, such as anxiety and depression, varies across generations. Results from Chapter 4 (Section 4.4,4.5,4.6 and Figures 19) reveal that Gen Z and Millennials are more vulnerable to mental health challenges associated with high social media engagement, particularly when exposed to emotionally charged or negative content. In contrast, Baby Boomers and The Silent Generation show weaker correlations, likely due to lower engagement with social media. These findings suggest that age-specific mental health interventions are essential, with younger generations benefiting from strategies to limit exposure to negative content, while older generations could focus on maintaining balanced digital habits. By understanding these generational differences, the study supports the development of tailored approaches to mitigate the mental health impacts of social media.

5.1.8 Research Objective 8:

To develop machine learning models that predict stress levels for each generation, taking into account technology usage patterns, sleep hours, and physical activity levels.

The eighth objective of the study focused on developing machine learning models to predict stress levels across generations by analyzing technology usage, sleep hours, and physical activity. Using Long Short-Term Memory (LSTM) and Random Forest (RF) models, the study achieved promising results. Chapter 4 (Section 4.8) highlights that the LSTM model attained an accuracy of 68%, effectively capturing sequential patterns in social media usage, making it especially useful for younger generations like Gen Z and Millennials. The model was particularly accurate for individuals with low to moderate social media usage (0-3 hours), although its effectiveness decreased with higher usage, likely due to data imbalance. The Random Forest model added valuable insights by identifying technology usage and sleep hours as consistent predictors of stress across generations. These results suggest that age-specific predictive models can guide tailored mental health interventions, addressing each generation's unique stress factors.

5.1.9 Research Objective 9:

To determine the optimal balance of social media usage and physical activity that can improve mental health outcomes across different generations.

The ninth objective of this study was to identify the optimal balance between social media usage and physical activity to enhance mental health outcomes across various generations. Findings from the prescriptive analysis in Chapter 4 (Section 4.7) indicate that younger generations, such as Gen Z, benefit more from increased physical activity than from reduced social media usage. In contrast, older generations like Gen X and Baby Boomers show improved mental health outcomes with decreased screen time. These findings align with previous studies, such as Sicilia and Charoensukmongkol (2015), which highlighted that excessive social media engagement is associated with increased emotional exhaustion and lower mindfulness et al. (2019), which demonstrated correlations between social media use, sleep disturbance, and mental health .

5.1.10 Research Objective 10:

To use machine learning models to predict mental health outcomes based on social media usage and demographic factors.

The tenth objective of this study was to use machine learning models—Random Forest (RF), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)—to predict mental health outcomes based on social media usage patterns and demographic factors like age and generation. In Chapter 4, Section 4.4, **Results of Random Forest**, the analysis highlighted that RF effectively identified key predictors, with social media usage frequency and nighttime engagement emerging as significant indicators of stress and anxiety. Table 5 illustrates feature importance within the Random Forest model, indicating that high social media usage and disrupted sleep were major contributors to adverse mental health outcomes, especially in younger demographics.

Section 4.5, **Results of Recurrent Neural Network (RNN)**, demonstrated RNN's ability to capture sequential dependencies, allowing it to predict mental health trends based on temporal social media usage patterns. Visualizations in Section 4.5.2 showcase the RNN model’s accuracy in identifying anxiety spikes linked to increased nighttime social media use. Similarly, Section 4.6, **Results of Long Short-Term Memory (LSTM)**, highlighted LSTM’s proficiency in recognizing long-term behavioral trends, particularly among younger generations with high social media engagement. Table 6 in this section shows the predictive accuracy of LSTM in distinguishing outcomes like anxiety and depression. Together, these results underscore the potential of using machine learning insights to guide tailored mental health interventions, addressing the mental health challenges associated with excessive social media use across diverse demographic groups.

Objective	Method	Outcome	Chapter
To build a predictive machine learning model to estimate stress levels based on technology usage patterns,	Machine learning models: RNN, LSTM, Random Forest	Accurate prediction of stress levels, with LSTM and RF identified as the most effective models	Chapter Four

sleep hours, and physical activity levels			
To compare the effectiveness of various machine learning models (k-means, RNN, LSTM, and RF) in predicting mental health outcomes	Comparative analysis of model performance	LSTM and Random Forest showed highest predictive accuracy	Chapter Four
To propose strategies for reducing social media usage and increasing physical activity	Prescriptive analysis and literature synthesis	Recommended strategies for digital moderation and increased physical activity	Chapter Four
To analyze technology usage patterns and stress levels across different generations	Statistical and machine learning analysis on generational patterns	Distinct technology-stress associations by generation	Chapter Four
To examine the correlation between high social media usage and mental health outcomes across generations	Generational analysis of social media usage and mental health metrics	Stronger correlations found in younger generations	Chapter Four
To use machine learning models to predict mental health outcomes based on social media usage and demographics	Application of RF, RNN, and LSTM models	Effective predictions, with RF and LSTM highlighting key predictors	Chapter Four

Table 7 Summary of the Study Findings

5.2 Recommendations

The recommendations provided here address the concerns outlined in the significance of the study in Chapter 1, focusing on strategies to improve mental health outcomes associated with social media use, sleep, and physical activity across generations. Based on the findings, the following recommendations are proposed:

5.2.1Develop Enhanced Social Media Engagement Filters

Researchers recommend the development of advanced digital engagement filters that can help individuals manage social media usage effectively. These filters could dynamically adjust based on user behavior, focusing on reducing exposure to emotionally charged or negative content to minimize mental health impacts.

5.2.2 Increase Awareness and Guidelines for Social Media Usage and Physical Activity Balance

Increased awareness and structured guidelines on the balance between social media use and physical activity are essential. It is recommended that educational campaigns promote the importance of physical activity, particularly among younger generations, as this has shown to alleviate mental health risks associated with high social media engagement.

5.2.3 Further Research on Personalized Mental Health Interventions

Additional research should be conducted to explore personalized mental health interventions that consider individual technology usage patterns, demographics, and lifestyle habits. These personalized strategies would aim to provide effective, tailored recommendations for improved mental well-being.

5.2.4 Encourage Generational-Specific Social Media Usage Recommendations

Different generations exhibit unique patterns of social media use and responses to mental health challenges. Researchers recommend that public health strategies be adapted to provide specific recommendations tailored to each generation's usage patterns and associated mental health needs.

5.2.5 Educate Communities on Healthy Digital Habits and Co-regulation Techniques

Community education initiatives should focus on teaching healthy digital habits and co-regulation techniques, helping individuals to self-monitor and adjust their social media behaviors to avoid potential stress and anxiety triggers.

5.2.6 Promote Collaborative Studies for Commercial Health Applications

Researchers recommend collaborative studies between technology developers and mental health professionals to explore commercial applications that could introduce mental health-focused digital tools. These tools could monitor and offer suggestions to balance social media usage, screen time, and physical activity in a way that supports long-term mental well-being.

5.2.7 Encourage Social Media Platforms to Integrate Wellness Tools

Social media companies should be encouraged to integrate wellness features that allow users to track usage patterns, implement breaks, and manage their digital engagement. Such features could empower users, particularly young adults, to maintain a healthy balance between online engagement and physical activities.

5.2.8 Further Research and Development of Holistic Health and Technology Integration

Finally, future researchers are encouraged to seek advancements in health and technology integration. Exploring complex models of digital behavior management and mental health tracking can provide holistic solutions to ensure sustainable mental health benefits across all demographic groups.

5.3 Implication of Study

This study has significant implications for public health, behaviour change interventions, and the design of mental health applications. First, it highlights the impact of social media usage on mental health outcomes, such as stress and anxiety, across different generations. Younger generations, especially Gen Z and Millennials, are shown to be more vulnerable to the mental health effects of high social media engagement. Figure 17 illustrates the average stress levels across Baby Boomers, Gen X, Gen Z, and Millennials, underscoring generational disparities in digital behavior responses. Similarly, Figure 15, which displays average technology usage across these groups, reveals that Gen Z engages more intensely with digital media than older generations. These findings align with prior research and suggest the need for targeted public health interventions to address the unique digital habits and mental health needs of each generation. Johnson and Lee (2019) argued that younger users may struggle to self-regulate their digital habits, leading to elevated stress levels—an insight that aligns with the patterns of increased anxiety observed among Gen Z participants in this study .

In terms of behaviour change and intervention strategies, the study provides insights that emphasize the importance of balancing social media use with physical activity to mitigate negative mental health effects. The comparative analysis of machine learning models in Section 4.8 identified Random Forest and LSTM as particularly effective in predicting stress based on social media and lifestyle patterns. Table 5, which summarizes average technology usage and stress levels across generations, supports the development of targeted, generationally specific strategies. These insights could help mental health practitioners and policymakers design interventions that reduce the impact of digital exhaustion, as illustrated in Figure 1, thereby promoting healthier digital habits and improved mental health outcomes.

Finally, the study has practical implications for the design of digital mental health applications. Prior research, summarized in Table 1, has established general links between digital behaviors and mental health. However, this study goes further by using machine learning to identify specific predictors of mental health outcomes. Figure 20, which shows actual versus predicted stress levels using the RNN model, illustrates the model's effectiveness in forecasting mental health risks based on social media use. This information can guide the development of mental health applications that track digital engagement, enabling users to self-regulate their behavior based on real-time stress predictions. Additionally, Figure 18 explores the varying impacts of social media on mental health across generations, informing the design of applications that incorporate these generational differences into personalized recommendations and user interfaces. Collectively, these insights advance the field of digital mental health by supporting the creation of tailored, data-driven interventions for diverse age groups.

5.4 Limitations of the Study

While this study provides valuable insights into the relationship between social media usage, sleep patterns, physical activity, and mental health across generations, several limitations must be acknowledged to fully understand its scope and potential areas for improvement.

Firstly, the study's reliance on self-reported data for social media usage, sleep quality, and mental health outcomes may introduce biases, as participants could misreport or inaccurately recall their behaviors and symptoms. This reliance on self-reporting can lead to issues such as social desirability bias, where individuals may underreport behaviors perceived as negative or overstate positive ones, potentially affecting the accuracy and reliability of results. Future studies may benefit from incorporating objective tracking measures, like app usage logs or wearable devices, to provide a more precise assessment of these behaviors.

Secondly, the study's sample may not fully represent the diversity within each generation, which could limit the generalizability of findings. Generational cohorts can encompass a wide range of socioeconomic, cultural, and regional differences that influence digital behaviors and mental health experiences. Additionally, access to technology varies within and across generational groups, potentially affecting the frequency and impact of social media usage. This limitation suggests that future studies should strive for more diverse and representative samples to capture a fuller spectrum of experiences and demographic factors, thereby enhancing the applicability of the findings.

Lastly, while the use of machine learning models like RNN, LSTM, and Random Forest provides powerful predictive insights, these models come with inherent challenges. The performance and accuracy of these models are heavily influenced by the quality and structure of the input data, and they may be susceptible to issues like overfitting, particularly with complex architectures like LSTM. Moreover, the study focused on specific predictive features such as social media usage and sleep hours, potentially overlooking other relevant factors impacting mental health outcomes, such as offline social interactions, genetic predispositions, or environmental stressors. Expanding future research to include a broader range of variables and exploring hybrid modeling approaches may offer a more holistic understanding of mental health predictors, improving the robustness and accuracy of predictions.

5.6 Future Work

The findings of this study suggest several directions for future research, focusing on enhancing predictive accuracy, adopting longitudinal approaches, and developing personalized mental health interventions. By addressing these areas, future studies could improve our understanding of how social media usage, sleep, and physical activity influence mental health across different age groups and contexts.

5.5.1 Expanding Machine Learning Models for Mental Health Prediction

Future research could explore more advanced machine learning architectures to achieve deeper insights into mental health predictors based on digital behaviors. This study used RNN, LSTM, and Random Forest models, each offering unique strengths. However, newer architectures, such as Transformer models, could handle complex sequential data with even greater precision, potentially revealing additional layers of insight into behavioral patterns. Transformers, known for their attention mechanisms, could allow models to prioritize critical periods of social media usage or specific behavioral patterns that correlate most strongly with mental health outcomes. Furthermore, hybrid models that combine LSTM with attention layers could better capture long-term dependencies in data, improving accuracy in predicting outcomes like anxiety and depression over extended periods. These enhanced models could support interventions that more effectively target at-risk groups and adapt to individual usage patterns, thereby improving the effectiveness of mental health support strategies.

5.5.2 Conducting Longitudinal Studies for Temporal Insights

To build on the cross-sectional insights of this study, future research should adopt longitudinal approaches that capture changes in mental health outcomes over time. While this study provides a snapshot of social media's impact, it does not account for long-term effects or causative relationships. Longitudinal studies could track individuals across different life stages, exploring how sustained digital engagement affects mental health as individuals age or transition between life phases. By following users over months or years, researchers could also study the lasting effects of high social media engagement on sleep patterns, stress levels, and general well-being. Including additional demographic factors, such as cultural background, socioeconomic status, and educational level, would enable a more nuanced understanding of how these variables intersect with digital habits to impact mental health. Such longitudinal data could help identify critical windows for intervention, offering valuable insights for creating more targeted and effective mental health policies.

5.5.3 Developing Personalized Digital Well-Being Interventions

Building on generational insights from this study, future research could focus on creating personalized digital mental health interventions tailored to individual behavioral patterns and preferences. Generational analysis has shown that digital habits vary widely between age groups, suggesting that different generations may benefit from unique intervention approaches. For instance, young adults and teens, who tend to have high levels of digital engagement, might benefit from applications that encourage physical activity as a counterbalance to social media usage. Older adults, on the other hand, might respond better to screen time reduction tools and features promoting mindful digital engagement. By integrating user-centered design principles, future mental health applications could offer adaptable solutions that evolve with users' needs and behavior changes, fostering a supportive digital environment that aligns with their mental health goals. Incorporating real-time feedback, personalized reminders, and insights based on usage patterns could further enhance these applications, making them powerful tools in promoting digital well-being across diverse demographic groups.

These avenues for future work aim to advance the field of digital mental health, leveraging cutting-edge technology to deepen our understanding of digital engagement’s long-term impact on mental health. By focusing on advanced machine learning, longitudinal analysis, and personalized interventions, future research can contribute to developing comprehensive mental health solutions tailored to the evolving needs of different generations and communities.

5.6 Conclusion

This study has highlighted the complex interactions between social media usage, sleep patterns, physical activity, and mental health across different generations, providing valuable insights into how these factors influence well-being. By employing machine learning models such as RNN, LSTM, and Random Forest, the research effectively identified key predictors of mental health outcomes and revealed distinct differences between generational cohorts. Younger generations, like Gen Z and Millennials, were found to be more susceptible to the negative effects of excessive social media engagement, especially during nighttime hours, which correlated with increased levels of stress and anxiety. These findings underscore the need for age-specific interventions, focusing on balanced social media usage and encouraging physical activity to mitigate the adverse effects on mental health.

The study also offered practical recommendations for developing personalized digital well-being tools that cater to unique generational behaviors and preferences. Despite some limitations, including reliance on self-reported data and the lack of a broader, more diverse sample, the research significantly contributes to understanding the impact of technology on mental health. Future work should focus on longitudinal studies to capture long-term changes and employ advanced predictive models to improve the accuracy of mental health assessments. Overall, this study serves as an essential foundation for guiding policymakers, mental health practitioners, and developers in crafting tailored strategies to enhance mental

References

Cain, N., & Gradinaru, M. (2010). Electronic media use and sleep in school-aged children and adolescents: A review. Sleep Medicine, 11(8), 735-742.

Clark, J. L., Algoe, S. B., & Green, M. C. (2017). Social network sites and well-being: the role of social connection. Current Directions in Psychological Science, 27(1), 32-37.

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2019). Studying emotion and mental health through language on social media: A literature review. Journal of Social Computing, 1(2), 26-42.

Hale, L., & Guan, S. (2015). Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep Medicine Reviews, 21, 50-58.

Karaman, M. A., Balkin, R. S., & Juhnke, G. (2019). Examining associations between social media use, depression, global health, and sleep disturbance among emerging adults. Research on Education and Media, 11(2), 57-63.

Liu, Y., Wheaton, A. G., Chapman, D. P., Cunningham, T. J., Lu, H., & Croft, J. B. (2020). Prevalence of Healthy Sleep Duration among Adults—United States, 2014. Morbidity and Mortality Weekly Report, 65(6), 137-141.

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. Nature Human Behaviour, 3(2), 173-182.

Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. EPJ Data Science, 6(1), 15.

Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., & Mohr, D. C. (2016). Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. Journal of Medical Internet Research, 18(7), e175.

Scott, H., & Cleland Woods, H. (2019). Understanding Links Between Social Media Use, Sleep and Mental Health: Recent Progress and Current Challenges. Current Sleep Medicine Reports, 5(3), 141-149.

Scott, H., & Woods, H. C. (2018). Fear of missing out and sleep: cognitive behavioural factors in adolescents' nighttime social media use. Journal of Adolescence, 68, 61-65.

Seabrook, E. M., Kern, M. L., & Rickard, N. S. (2016). Social networking sites, depression, and anxiety: a systematic review. JMIR Mental Health, 3(4), e50.

Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., ... & Zhu, T. (2017). Depression detection via harvesting social media: A multimodal dictionary learning solution. Proceedings of the 26th International Conference on World Wide Web, 1043-1052.

Sulaiman, R. B., Schetinin, V., & Sant, P. (2022), Human-Centric Intelligent Systems

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). Introduction to Data Mining. Boston, MA: Pearson Addison-Wesley.

Woods, H. C., & Scott, H. (2016). #Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. Journal of Adolescence, 51, 41-49.

https://link.springer.com/article/10.1007/s40675-019-00148-9

https://sciendo.com/pdf/10.2478/rem-2019-0022

https://github.com/vishal90234/Analysis-Of-Social-Media-and-Sleep-Disorders-on-Mental-Health-capstone--2

https://www.semanticscholar.org/paper/Examining-associations-between-social-media-use%2C-Karaman/9eb074dde2bec46c1a9cf9cba5b46d15b2b0f8b2#cited-papers

https://www.kaggle.com/code/serenahinton/social-media-and-mental-health-analysis/input

https://www.kaggle.com/code/jens78/socmed-eda-and-mental-health-risk-prediction/notebook

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10948475/

https://www.kaggle.com/code/mahmoudredagamail/mental-health-technology-usage-dataset/notebook

SOCIAL-media-history