Evaluating Mental Health and Well-Being through Social Media Analysis

Muneera Al-Dossary[1], Mashael Saeed Alshahrani[1], Amal Furaih Altamimi[1], Mohammed Gollapalli[2], Atta Rahman[3,*]

[1]Department of Computer Information Systems, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[2]Kaplan Business School, Gold Coast, Queensland, 4215, Australia. Mohammed.gollapalli@kbs.edu.au
[3]Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia
[*]Correspondence: aaurrahman@iau.edu.sa

Abstract
As social media platforms continue to gain popularity, there is increasing interest in understanding how online interactions affect users' mental health and overall well-being. This study takes a novel approach by collecting and analyzing social media data to assess mental health comprehensively. By utilizing data from various platforms, we employed thorough analytical techniques to gain valuable insights into users' mental states. The study addresses the crucial issue of the impact of social media on mental health and well-being by providing a detailed dataset and analytical framework for future research and intervention development. Our findings reveal significant correlations between social media activity patterns, sentiment analysis, and self-reported mental health indicators. We demonstrate the potential of using online discussions for early mental health intervention and monitoring. This research not only advances the field of mental health assessment but also offers practical implications for mental health professionals, policymakers, and developers of social media platforms in enhancing online mental health support. By highlighting the effectiveness of data-driven approaches, this study reveals innovative ways to improve mental health and well-being assessments in the digital age.

Keywords:
Social media; mental health; wellbeing; R programming language; data mining; data wrangling.

1. Introduction

The World Health Organization (WHO) describes mental health as the "self-actualization of one's intellectual and emotional ability, subjective well-being, perceived self-efficacy, autonomy, competence, inter-generational reliance, and other elements. However, it is commonly acknowledged that mental health encompasses more than the absence of mental diseases. As a result, there is a wide range of mental health that includes both positive and negative effects. Happiness and a high mental quality of life are beneficial aspects of mental health. Conversely stated, the absence of disorders like anxiety and sadness is a sign of excellent mental health. Literature also demonstrates and hints that a person's total health, including physical well-being and social pressure, can be a foundation for and a mutually beneficial consequence of their mental health [1]. The use of the internet has progressed astoundingly over the years in such a way that it is now easy to utilize. The internet is now more accessible than ever thanks to the development of user-friendly apps for websites, Smartphones, and mobile phones. Traditional static websites are losing popularity to Web 2.0-style interactive Web Pages. In a web-based community, any user can create material and share it on the internet using various websites and programs known as "Web 2.0." User-generated content is developed and easily shared with other users using social media, a sort of Web 2.0 that has only recently been offered as internet-based websites and apps [2]. The daily lives of about half of the world's population have changed because of social media use making the average user spend 2 hours and 25 minutes per day on social media, which is more than they do in a day for eating, reading, and exercise combined. The concept of social media has been defined in a variety of ways due to its explosive growth, with early interpretations emphasizing Internet applications for user-generated content creation and sharing and more recent definitions focusing on virtual communities that offer sophisticated opportunities for social interaction. Social media is described in this study as web-based social networks with cutting-edge tools for producing and sharing content that promotes user interaction. We specifically look at how social media use affects attitudes and behaviors by concentrating on heavy

Manuscript received November 5, 2024
Manuscript revised November 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.11.3

users [3]. Social media has been found to have an impact on habits related to health, such as food and exercise, as well as public opinion, political beliefs, and purchase decisions. It now poses a hazard to public health since it has developed into a powerful platform for the broad spreading of false information. Despite many studies examining the risks of false information using social media data to assess public views, there is a lack of specific information on how mental health is effectively addressed and improved through social media for health promotion [4]. This project aims to utilize social media data to gain insight into individuals' emotional states, concerns, and coping mechanisms. However, to make this unstructured data useful, a crucial first step is required—data wrangling. Data wrangling involves collecting, cleaning, and transforming raw data into a structured format suitable for analysis. In this study, data wrangling is essential for aggregating and preparing diverse social media data sources for in-depth analysis. The main objectives are to find a dataset containing textual content, metadata, and temporal information from various social media platforms, and to use rigorous data cleaning and pre-processing techniques to ensure data quality, remove noise, and anonymize sensitive information. This carefully curated dataset will form the basis for comprehensive mental health and well-being assessments.

The remaining sections are organized as follows. Section 2 contains the literature review, while Section 3 contains the results of all the designated significant research. Section 4 represents dataset description, statistical analysis, and data wrangling, section 5 demonstrates the experiment results and discussion, whereas the last section contains our conclusions and future work recommendations.

2. Review of related literature

In a study, Lin et al. [1] aimed to understand the impact of social media usage on individuals' mental health. They explored the mediating roles of generalized trust (the belief in the trustworthiness of others) and perceived social support (the feeling of being supported by others) in this relationship. The methodology used in the paper involved a cross-sectional study design, gathering data at a single point in time from a sample of participants. Data collection included survey questionnaires assessing problematic social media use, mental health status, generalized trust, and perceived social support. The key findings indicate that problematic social media use is associated with poorer mental health outcomes, including increased stress, anxiety, and depressive symptoms. Generalized trust plays a mediating role, partially explaining the relationship between problematic social media use and mental health. In addition, lower generalized trust is linked to higher problematic use and, in turn, poorer mental health outcomes. Perceived social support also mediates the relationship between problematic social media use and mental health. Individuals with lower perceived social support tend to engage in more problematic social media use, which negatively affects their mental health. The study highlights the complex interplay between elusive social media use, generalized trust, perceived social support, and mental health. It suggests that addressing issues related to generalized trust and perceived social support could be important when designing interventions to mitigate the negative effects of excessive social media use on mental health. The findings emphasize the need for individuals, healthcare professionals, and policymakers to be aware of the potential mental health consequences of problematic social media use and to consider strategies that promote trust and social support in digital spaces. The paper’s clear limitations revolve around the fact that as a cross-sectional study, causal relationships cannot be established. Longitudinal research may provide deeper insights into the direction of these associations over time, and the data in this study relies on self-report measures, which may introduce response bias. In conclusion, this study contributes to understanding of how problematic social media use can impact mental health, shedding light on the mediating roles of generalized trust and perceived social support in this relationship. It underscores the importance of considering these factors when addressing the mental health implications of excessive social media use.

Hill et al. [2] indicate in their recent research that a range of effects of social media use on mental health, including negative, positive, and nothing at all. The quality of knowledge and training of doctors in screening and identifying these connections, however, has not undergone much research. For this reason, the authors developed a study that evaluated the understanding of potential physicians, or present

medical students, on the links between mental health and social media. The methodology used in the study was a 12-question poll that was used for the investigation. The questions were intended to measure knowledge gained through prior social media interactions, education regarding its usage and potential effects, and self-reported capacity to screen, identify, and advise patients on these associations. All classes from MS1 to MS4 were included in the survey, which went out to a total of 634 medical students. This survey received 148 responses overall (23.3% response rate). The major key finding was that most of the medical students indicated that they utilized social media for the first time between the ages of 13 and 18, and it was also found that this age group experienced bullying the most frequently. Even though most students thought social media might have both beneficial and bad effects on mental health, only a few of them claimed to be aware of specific social media usage patterns that are linked to mental health. Furthermore, only a few students reported feeling confident enough to screen, inform, and counsel patients about these relationships. The study highlights that even though the study's findings only represent a small portion of medical students, they nonetheless show the need for a pertinent teaching program for future doctors. Early in their education, workplace learning among medical students has been proven to have a favorable impact on elements including professional identity, attitude, skills, team experience, and task performance. Therefore, it is essential that such problems as social media use and mental health be appropriately addressed and studied because the knowledge provided in medical school serves as the foundation for the aspiring physician. Impressively, the study's limitations have been recognized as it was limited by the absence of data on the demographics of the medical students who engaged with the study, such as sex, age, and year in medical school. Additionally, the results from the survey data are those of a specific organization. Furthermore, using a yes/no response format rather than a Likert-style scale could have restricted the range of agreement and disagreement. In conclusion, we notice that the study's findings point is the need for educational materials to teach aspiring doctors how to screen, identify, and counsel patients on the connections between social media and mental health, while also taking into consideration the validity of the result interpretation in this study was limited by the small sample size. In the research conducted by Farsi et al. [3], tried to gauge the effect on the individual who are integrated in a global web-based environment where the internet is considered as the primary source of health-related information for the current populace, specifically the younger age. Patients can utilize social media in the healthcare industry for more special purposes like telemedicine, choosing a physician, and peer support. The purpose of the research is to explore how social media has been used in the healthcare sector from the viewpoint of patients and to outline the main concerns related to that use. The methodology used in the research was in the form of a study of the literature was done between March and June 2020 on PubMed, Google Scholar, and Web of Science for English-language articles that had been released since 2007 and addressed regarding the use of social media in healthcare. Publications relating to ethical and legal issues surrounding the use of social media were also included, in addition to solely English-language publications that covered patients' use of social media. The studies were then divided into categories such as health education, telemedicine, locating a medical professional, peer support and experience sharing, and promoting healthy behavior. The review has also been expanded with two new sections on ethical issues and problems with social media use in healthcare. The major key finding in the research was in the form of a case study where 75 studies were initially listed. A total of 91 studies were reviewed, along with an article and 13 web references, as the study went on to involve further studies. Reviews were created for around half of the studies. More than half of the research was published in the last five years, with the first one appearing in 2009 and the last in 2021. Most of the studies (n=40) were from the United States, followed by those from Europe (n=13), while the fewest (n=1) came from India. The social media platform that was under the most scrutiny was WhatsApp or WeChat. The study's limitation, despite being accurate, was that this review was observational in nature and did not provide a formal evaluation of the included papers because it is a narrative review. The included studies' data were reorganized and summarized, but not evaluated.

Despite the thorough search, it's possible that some pertinent studies went undiscovered. Given the nature of the review, there may have been bias in the literature selection and evaluation since they weren't done in a systematic way. In conclusion, the public

and patients can use social media to advance their knowledge and health and to evaluate the validity of the information gathered and its source, however, careful investigation is required. The risks connected with using social media should not be ignored by healthcare professionals, patients, or the public. It is important to comprehend the restrictions and disadvantages associated with patient social media use. The research shows that multifaceted health services have proven to be quite effective, especially when combining medical care with social networking sites and other kinds of communication. Through the audience is reached repeatedly, in diverse contexts, and through many sources, the outcome is maximized. The purpose of this research is to examine the mental health practitioners and young people's experiences of discussing young people's web-based activities related to mental health, identify barriers to effective communication, and highlight examples of good practice. The research design is a qualitative research study using interviews with mental health practitioners and young people and focus groups based on thematic analysis. This study's findings will shed light on the practices and perceptions of professionals and young people for exploring web-based activities in mental health consultations is concerned [4].

The study [5] looks at the use of digital technologies, particularly web-based activities such as social media, websites, and messaging when addressing mental health concerns for adolescents. Methodology The study used a qualitative design, which included focus group discussions and interviews that involved mental health professionals and adolescents/young adults between 16 to 24 years. The data collected from the focus groups and interviews were analyzed using thematic analysis. Under the Lens: Mental Health Practitioners and Young People in Discussing Web-Based Activities Linked to Mental Health during Clinical Consultations. The study took place in the UK and included young people who had experienced mental health problems and received support from statutory mental health services or third-sector organizations, as well as mental health practitioners working in different settings. Data for the study was obtained using focus groups conducted among young people and interviews with mental health practitioners. Discussing the web-based activities of young people and their mental health implications for young people. Some practitioners report having different levels of confidence in addressing web-based activities; they also indicated that more guidance and training are required. Practitioners rarely solicit information about young people's web-based behavior, and when they do youth say they feel judgmental or misinterpreted. Generally, the study's findings point out that although there is awareness of the relevance of Web-based activities in mental health conversations, there are difficulties in effectively dealing with this issue. Practitioners say there is a demand for more direction and instruction and some young people feel that their online behaviors are misunderstood or judged. Similarly, conversations about the extreme aspects of the web-based world are limited by the constraints of time and low assurance in addressing them. The study designed and validated the social media-induced Tendency (SMIDT) scale to measure social media-induced depression tendency among young people in Nigeria using an online survey (Google Forms). The SMIDT scale, consisting of 15 items, showed high internal consistency and reasonably valid, informing about the reasons behind depression in the use of social media. To construct and validate the SMIDT scale, a standardized multi-step method involving exploration factor analysis (EFA) and confirmatory factor analysis (CFA) was applied. A cross-sectional design was adopted for the research design. The data analysis used Descriptive Statistics and EFA on SPSS. Exploratory Factor Analysis was used to make this possible. To establish the validity of the SMIDT scale, the confirmatory factor analysis was done to examine internal consistency, construct validity, discriminant validity, and concurrent. The research used an online survey (Google Forms) along with purposive sampling to gather information from young Nigerians. The scale had high internal consistency and showed promising validity, suggesting that it could help identify those at risk of becoming socially media-induced depression cases and come up with ways to prevent or minimize the potential. As suggested by the study, it is advisable to recruit people of different age brackets, including older adults, for future investigations utilizing the SMIDT inventory to increase the generalization [5].

In recent years, the rise of social media platforms has significantly impacted how individuals communicate, share information, and perceive themselves and others. Social media, encompassing a variety of platforms like Facebook, Twitter, Instagram, and others, has become an integral part of daily life,

provided quick electronic communication, and facilitating the sharing of thoughts, ideas, and personal information. This digital revolution, accelerated by smartphones and high-speed internet, has transformed communication and social interactions globally. While social media offers numerous benefits, including enhanced connectivity, access to information, and opportunities for self-expression, it also poses several mental health challenges [6].

This research delves into the psychological underpinnings of social media use, explores the merits and adverse effects, examines the relationship between social media and mental health disorders like depression and anxiety, discusses the phenomenon of cyberbullying, and addresses the impact of social media on self-esteem and narcissism. The study emphasizes the need for continued research to identify harmful effects, develop prevention strategies, and optimize the coexistence of technology and mental well-being. The research investigates the intricate relationship between social media usage patterns and mental health outcomes. It explores various aspects, such as the psychological motivations behind social media use, the positive impacts of social media on mental health, and the adverse effects, including depression, anxiety, and cyberbullying. The study also examines the connection between social media, self-esteem, and narcissism, shedding light on the complex interplay between online interactions and individuals' psychological well-being. The research highlights the multifaceted nature of social media's impact on mental health. While it provides avenues for social inclusion, peer support, and access to information, excessive use and negative online interactions can lead to increased levels of depression, anxiety, and feelings of loneliness [7].

Cyberbullying, facilitated by social media platforms, poses significant risks, contributing to decreased self-esteem and emotional distress among victims. Moreover, the study identifies a correlation between social media use and anxiety symptoms, emphasizing the need for awareness regarding excessive smartphone usage and its potential adverse effects. The main obstacle facing the evolution of research in this field is the rapid pace of technological advancement. With the continuous spread and development of social media and digital media, it becomes challenging to keep up with all the changes and understand their psychological effects on individuals and society. These rapid developments require researchers to stay updated on the latest technologies and applications and how they impact mental health. Furthermore, the quick changes in social and digital media use pose a challenge regarding data collection and analysis. Researchers must be familiar with modern data collection methods and analytical tools to ensure the accuracy of results and the validity of conclusions. Additionally, it is difficult to determine the nature and prospects of research in this field due to the continuous nature of technology and its increasing impact on individuals' lives. Therefore, future research efforts will require more diligence in monitoring technological developments and fully understanding their effects on mental and social health. While the research provides valuable insights into the psychological aspects of social media use, it lacks a detailed exploration of potential interventions and coping mechanisms for individuals affected by the negative consequences of excessive social media usage. Additionally, the study does not delve into the role of social media in specific vulnerable populations, such as adolescents, and how socio-cultural factors might influence their experiences. Future research could benefit from addressing these gaps, focusing on preventive strategies, and examining the cultural nuances that shape the impact of social media on mental health [7]. The research aims to establish a connection between smartphone and social media use and the rise in adolescent mental distress, self-injurious behavior, and suicidality. It investigates how these digital technologies influence adolescents' self-perception, interpersonal relationships, and overall mental well-being. The findings of the research underline a significant correlation between excessive smartphone and social media use and adverse mental health outcomes among adolescents. Factors like social comparison, cyberbullying, addiction, normalization of self-harm, impaired social skills, and sleep deprivation were identified as contributors to mental distress in young people. Despite comprehensive analysis, the research falls short in exploring potential solutions and interventions to alleviate the negative impacts of social media and smartphone use on youth mental health. Furthermore, the study does not incorporate the perspectives of mental health professionals and educators, who are vital in addressing these issues. One of the primary challenges faced by researchers is the swiftly evolving landscape of social media and technology. As new platforms and

features emerge, understanding their influence on adolescent mental health necessitates ongoing research and adaptation. Additionally, the absence of standardized definitions and scales for smartphone and internet addiction poses a challenge in accurately measuring and comparing prevalence rates [8].

The study [9] investigates the relationships between social media use, purchase behavior, and mental health among Malaysia's X, Y, and Z generations. It explores how social media influences purchase behavior and affects users' mental health. The research utilizes quantitative methods and examines the associations between these variables. Findings Social Media Use and Purchase Behavior A significant positive association exists between social media use and purchase behavior. Social media platforms play a crucial role in influencing users' purchasing decisions. Prolonged use of social media is positively associated with users' mental health issues. Spending more time on social media can increase stress, anxiety, and other mental health problems. Surprisingly, there is no significant association between users' purchase behavior and their mental health. Engaging in purchase activities does not necessarily lead to increased stress or mental health issues. The study's cross-sectional design limits establishing cause-and-effect relationships.

Longitudinal studies could provide deeper insights into the dynamics between social media use, purchase behavior, and mental health over time. The use of convenience sampling might introduce biases, impacting the generalizability of the findings. A more diverse and representative sample could enhance the study's validity. The study relies solely on quantitative data, potentially missing nuanced aspects of participants' experiences. Qualitative methods, such as in-depth interviews, could better understand the psychological mechanisms. The findings emphasize the need for regulatory agencies to establish guidelines on social media usage, mainly focusing on limiting excessive screen time to safeguard users' mental well-being. Future research could explore other dimensions of mental health, such as depression and anxiety, to provide a more holistic understanding of the impact of social media and purchase behavior on users' psychological states. The study sheds light on the complex interplay between social media use, purchase behavior, and mental health. While social media platforms significantly influence users' purchasing decisions, excessive use can adversely affect mental health. The absence of a direct link between purchase behavior and mental health suggests that engaging in consumer activities might not inherently lead to increased stress or anxiety. Further research endeavors could delve deeper into the nuances of these relationships, addressing the limitations and enhancing the overall understanding of this intricate interlink [10].

The study conducted an umbrella review to bridge the gaps in existing research, synthesizing 25 studies published between 2019 and mid-2021 [11]. This encompassed meta-analyses systematic and narrative reviews, analyzing the intricate relationship between SMU and various mental health outcomes among adolescents. The study aimed to delineate the general characteristics of existing assessments, explore diverse interpretations of SMU effects on mental health, pinpoint gaps in the evidence base, and propose future research directions. This research delved into the nuanced interplay between social media use (SMU) and adolescent mental health. Specifically, it synthesized literature reviews from 2019 to mid-2021, investigating the manifold outcomes of SMU on mental health, encompassing positive and negative indicators of ill-being. The study highlighted the complexities in interpreting SMU's impact, emphasizing its varied effects on well-being and ill-being among adolescents.

Furthermore, it underscored the limitations of existing studies, particularly their reliance on cross-sectional data and the need for more precise definitions of SMU and mental health outcomes. The umbrella review uncovered a spectrum of interpretations regarding SMU's impact on adolescent mental health, ranging from weak and inconsistent associations to substantial and harmful effects. Meta-analyses indicated invalid associations between SMU and higher levels of well-being and ill-being among adolescents. The study emphasized the necessity of a nuanced approach, recognizing that SMU's effects on mental health are multifaceted and individual-specific. It underscored the need for more sophisticated research methods, including interdisciplinary collaborations, to navigate the complexities of understanding SMU's effects on mental health outcomes. The study identified several gaps in the existing literature, including the need for precise definitions of SMU and mental health outcomes, capturing the content and quality of social media interactions, and the necessity for a person-specific

approach to understanding SMU's effects. Challenges related to self-report measures and the lack of specificity in studying different types of SMU activities were also highlighted. The researchers emphasized the importance of adopting advanced research methods and interdisciplinary collaborations to address these challenges effectively [12].

Several studies in the literature have been conducted to assess the social media role in mental and emotional health of individuals during and after the pandemic [13-16]. Such as in [17] authors analyzed tweets in this regard using deep learning techniques. Convolutional neural networks (CNN) and bi-directional long-short memory (BiLSTM) deep learning algorithms were harnessed for categorizing the sentiment of Arabic tweets. This experimentation flashed that the performance of CNN achieved 92.80% accuracy. The performance of BiLSTM was marked at 91.99% in terms of accuracy. The scheme emphasizes the sentiment analysis before, during and after the pandemic. Another study has been conducted for user behavior prediction and classification using neuro-fuzzy modeling. In this regard, the dataset was collected from an organization's proxy server log files related to the user activities over the network, web and machine [18]. Similarly, several studies have been conducted in the literature to determine the social media impact on the population wellbeing [19-30].

3. Materials and Methods

The research's goal is to study the effect of social media on the mental health of its users, to achieve this result the R language program is used to help prepare and clean the data by using the various studio abilities like handling missing data, cleaning the data, and variable changing to modify the data or create a new variable from the data. The reason for choosing R language is how highly it is recommended for creating frameworks and conducting data analysis, as analysts and data miners utilize it extensively. In the research, we test various classification algorithms and identify the key characteristics that are most useful for assessing how social media affects the people who use it. To achieve these things, a dataset is used to pinpoint characteristics that are essential for analyzing user traits, media consumption patterns, and media types used. The subsequent sections provide more details on the 4-stage methodological approach used with the dataset:

1. Initial processing of the data to ensure accuracy by synchronizing the data, detecting extreme outlines, analyzing the data, and eliminating anomalies.
2. Creating specific classes by defining the feature attributes and data example sets.
3. Choosing the best method from the available dataset by contrasting the commonly used techniques for classification in the research.
4. Identifying information collections about the traits of users, the frequency of time used in media, and the state of their mental health are essential to finding a solution to the current issue.

3.1 Clinical data

The study conducted has 482 test subjects who participated in a survey about the possible relationship between a person's mental health and the quantity of time they dedicate to social media.

Table 1: Feature description

Variables	Description
Gender	Male or Female or others.
Age	Age of the user when analyzed.
Relationship	Married/divorced/dating/single.
Occupation	Retired/salary worker/school student/university student.
Access to social media	Accessibility to social media (yes/no)
Social media types	Types of social media used by the users. (e.g. Twitter / Facebook)
Self-validation	Is knowing and affirming a person's own feelings without depending on outside factors, will it be affected by obsessive media usage or not.
Depressive episodes	Are periods of severe sadness, low energy, and avoid interest or joy, some stays for two weeks and might lead to clinical depression. The long hours of media use might help with that.
Activity Fluctuation	The different levels in a person's involvement in tasks, activities, observing how the repeated use of social media would change it.
Sleep Deprivation	It happens from insufficient sleep creating mood swings, future health

issues, the amount of time users spend on social media.

4. Empirical studies

4.1 Study data

The social media and Mental Health datasets were obtained from Kaggle website [31] It contains 7 variables, and twelve Likert scale-based questions providing points that measure either frequency or intensity of various aspects of Mental Health. A low score of 1 generally indicates low frequency or intensity, and a high score of 5 typically indicates high frequency or intensity. The patients' private details, including their national identifications and phone numbers, were discarded to maintain their privacy. Table 2 contains the basic details of the nominal attributes of the dataset.

Table 2: Nominal attribute distribution

Feature	Count	Values
gender	481	Female (263), Male (211), other (3)
relationship	481	Married (283), dating (202), divorced (18), single (348)
occupation	481	Retired (8), salary worker (130), University student (292), school student (47)
access to social media	481	Yes (474), No (3)
social media types	481	Facebook (388), Instagram (335), YouTube (395), discord (194), Pinterest (143), TikTok (93), Twitter (110), Snapchat (181), Reddit (121)

4.2 Statistical analysis

In dissecting social media data, we unveiled temporal trends, conducted sentiment analysis, explored correlations, and dissected demographics, uncovering disruptive insights into mental health discussions, paving the way for a nuanced understanding of well-being in the digital age, so this statistical analysis provides important tools for visualizing and comprehending a data pattern to elevate the data per-processing and modeling process. Table 3 provides the statistical analysis of the numerical attributes presented in the social media and Mental Health dataset, like the missing values, central tendency, standard deviation, minimum, and maximum.

Table 2: Numerical attributes statistical description

Feature	count	mean	std	min	25%	50%	75%	max
Self-validation	481	3	18.04	1	1.25	3	10.5	82
Depressive episodes	481	3	18.04	1	1.25	3	10.5	82
Activity Fluctuation	481	3	18.04	1	1.25	3	10.5	82
Sleep Deprivation	481	3	18.04	1	1.25	3	10.5	82

4.3 Experimental setup

This section deals with explaining how experimental research was carried out and conducting various assessments to understand the impact of social media on mental health and psychological well-being. The study used the R language to extract data and correlate attributes for each category of user based on age, gender, social status, and what platforms are used in social media and others to predict their mental health.

5. Retros pective analysis

In this section, we are using descriptive statistics that provide a thorough retrospective study which was carried out as an initial step on each of the target questions. The purpose of the study is to comprehend the variables related to a person's use of social media and how it affects their mental health.

5.1 User epidemiology

Every year, more people are using social media than ever before. The total amount of registered profiles climbed from 1.48 billion in 2012 to 3.2 billion in 2018 and 4.76 billion in 2023. These numbers demonstrate both the rising usage of social media and the growing number of troubling users, who suffer from Intoxicating habits, a persistent need to spend time on social media for many hours every day, and an overwhelming sense of lacking are typical characteristics of troubling social media users. Additionally, the World Health Organization (WHO) highlighted a rise in mental illness, as evidenced by their studies revealing that troubling social media users have almost comparable outcomes in terms of their impact on mental health [31].

In this study, we studied 482 unique users (from pre-teens till adulthood) who use different types of media (in the study 9 types of social media are used), Figure 1 shows the ages of these social media users and their occupation status. As it's shown the users information available were between the ages of 13 and 91; all age groups have an equal distribution of salary workers, demonstrating the stability and diversity of this profession in media usage, the graph also shows students that usually begin their schooling in their teens, as evidenced by the fact that the majority of university students and school students are young in the data provided to us [31].

Figure 2 shows the occupational status distribution of the population considered for the analysis. It is apparent that university students have been the most into social media usage, followed by the salary workers from various domains. School students appear in the third position while the retired people are smallest among all in this distribution.

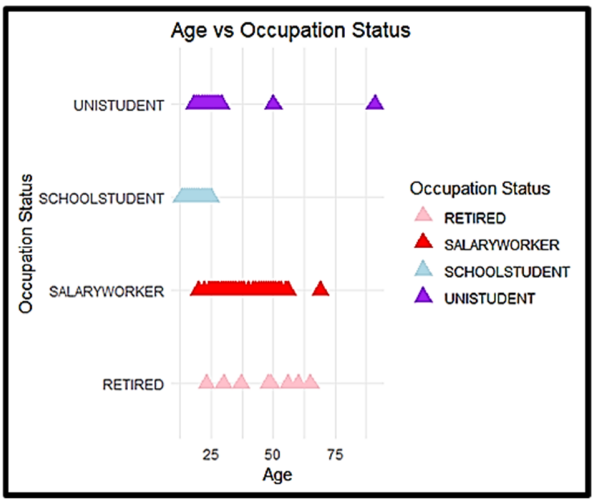

Figure 1: Age group comparison in occupation

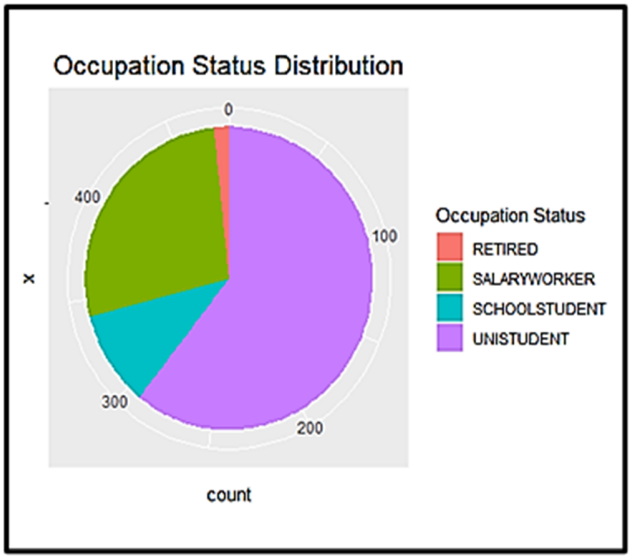

Figure 2: Occupation status distribution

5.2 Social Media Apps usage

The analysis of the types of social media was performed on users to see and determine the highest usage. As can be seen from Figure 3, it's clear that Facebook is the highest among his peers, after it comes to YouTube, we notice also how both apps are more female-centered usage (Facebook 100 individuals are males while 150 individuals are females), (YouTube 100 individuals are males while 150 individual's females). The data and figures also show how little males are interested in Twitter (Twitter 50 individuals are males versus 200

individuals are females), most female users from ages 20 to 40 that are married who are 128 users (26.50%) use social media more than single female users 120 users (24.84%).

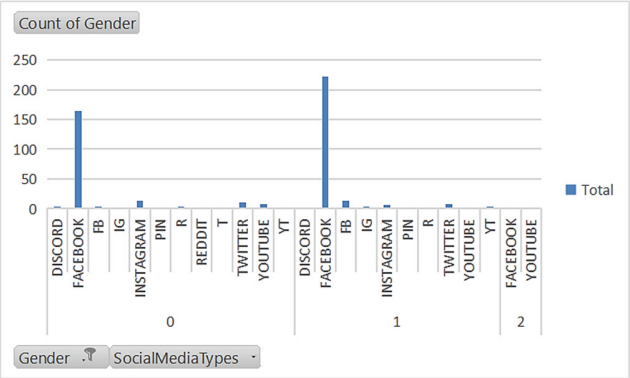

Figure 3. Social media usage distribution

5.3 Gender demographics

Figure 4 showcases all the users and their relation states. The results indicate that working females 26.9% (130) are more addicted to social media than working males 12% (59).

Additionally, Figure 5 goes more into the effect of both occupation and gender on the type of media any user consumes, the figure shows us the correlation between each pair of variables, The darker the color, the stronger the correlation becomes, as shown by the figure the type of correlation between male and female is a negative type correlation, the meaning on this type is that as one increases, the other decreases. This is expected because these variables represent the genders of the social media users, and in most cases, one gender is predominant over the other.

We can also see how every social media type per gender is affecting them, and the results show after detailed observation that female users are more prone to using Instagram, Facebook, and YouTube compared to male users, as given in Table 4.

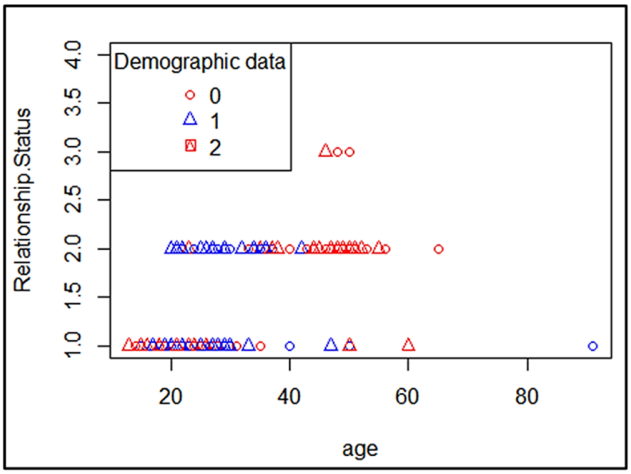

Figure 4. Relationships, age, gender distribution

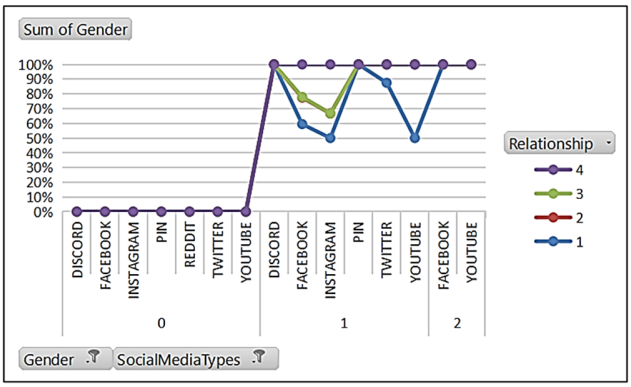

Figure 5: Social media and gender distribution

Table 4: Gender and social media statistics

Social media	Females	Males
Facebook	236 (81%)	168 (68.8%)
Instagram	12 (4.1%)	16 (6.5%)
YouTube	8 (2.7%)	13 (5.3%)
discord	1 (0.3%)	5 (2.0%)
Pinterest	7 (2.4%)	8 (3.2%)
TikTok	7 (2.4%)	8 (3.2%)
Twitter (X)	9 (3.0)	12 (4.9%)
Snapchat	4 (1.3%)	4 (1.6%)
Reddit	7 (2.4%)	10 (4.0%)
Total	291	244

5.4 Types of dysfunctions

Up until the last few years, dysfunctions had been underappreciated as a disorder, troubling for the people afflicted, and lacking in successful therapies. Before the COVID-19 pandemic, mental health disorders were thought to affect a small percentage of the general population. The number of dysfunction incidents has risen since the COVID-19 outbreak [16]. Based on the data obtained from the users using social media, four types of dysfunctions were regular in a lot of cases as shown in Table 5, these types start as a regular annoyance but if left untreated with the continuation of media usage increase, they eventually would develop into chronicle illnesses, as some lead to death (by illness or self-inflicted).

From Table 4, we notice that the ones who are most likely to develop these issues are females, as they tend to spend more time in media and can be easily affected by their emotions, especially in their youthful years.

Table 5: Types of dysfunctions

Symbols	Types of dysfunctions
DP	Depressive episodes
SD	Sleep deprivation
SV	Self-validation
AF	Activity fluctuation

6. Discussion

A different perspective on people's psychological well-being can be gained by collecting and analyzing social media data. This can provide valuable insights for assessing mental health and guiding the development of targeted interventions. However, ethical concerns come into play as we delve into this area. It's important to consider gender-specific patterns that may emerge in social media data. Research shows that women often share their mental health issues on these platforms, emphasizing the need for gender-specific and sensitive approaches to identify and support them promptly. The study suggests that providing specific guidance on seeking therapy would be beneficial for females, who are more likely to share their mental health struggles online. It's crucial to carefully frame this advice to highlight the strength and independence of seeking treatment, to reduce stigma and encourage proactive mental health care [32-50]. Nonetheless, it's essential to approach these recommendations with cultural sensitivity, recognizing that attitudes toward mental health and seeking help vary widely. It's important to ensure that the recommendations are realistic and inclusive by providing accessible and culturally competent resources and information about therapy options. Furthermore, it should be emphasized that social media data is not a substitute for professional diagnosis and treatment, even though it can be valuable. A comprehensive mental health awareness campaign should advocate for a holistic approach, integrating digital insights with traditional medical practices, and should explicitly encourage women to consider therapy [51-65].

7. Conclusion and recommendations

In conclusion, there are both advantages and disadvantages to gathering and evaluating social media data for a comprehensive assessment of mental health and well-being. Social media platforms provide a wealth of data that can provide insights into people's emotional states and the ongoing conversations about mental health. However, using this data for mental health assessment is complex and raises moral concerns, confidentiality issues, and the need for accurate analysis. To protect people's privacy in the future, strong ethical frameworks and privacy protections must be established. It is crucial to find a balance between safeguarding user privacy and maximizing the potential of social media data. Responsible guidelines for using data in mental health assessments can be developed through collaboration between social media platforms and data. This may involve investing in advanced machine learning and data analytics methods to enhance the accuracy and reliability of mental health assessments. Research on the relationship between mental health indicators and online behaviors should be continuous and consider cultural and contextual factors. Public awareness campaigns should also educate social media users about the implications of sharing mental health-related information online, promoting a stigma-free and supportive online community. Psychologists should be involved in developing methodologies and tools to maximize the usefulness of social media data, ensuring that assessments are clinically relevant and contextually informed. Integrating social media data into existing mental health frameworks can offer a more comprehensive view of people's mental health,

allowing for prompt intervention and personalized care.

Acknowledgment

The authors like to acknowledge College of Computer Science and Information Technology (CCSIT) at Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia for using the resources.

References

[1] C. Y. Lin, P. Namdar, M. D. Griffiths, and A. H. Pakpour, "Mediated roles of generalized trust and perceived social support in the effects of problematic social media use on mental health: A cross-sectional study," Health Expect., vol 24, no. 1, pp. 165-173, 2021, doi:10.1111/hex.13169.

[2] H. Hill, C. Hill, and J. W. Kim, "Prospective Physician Awareness of the Associations Between social media and Mental Health," Acad Psychiatry, vol. 44, no. 1, pp. 78–81, 2020, doi: 10.1007/s40596-019-01116-9

[3] D. Farsi, H. R. Martinez-Menchaca, M. Ahmed, and N. Farsi, "Social Media and Health Care (Part II): Narrative Review of Social Media Use by Patients," J Med Internet Res., vol. 24, no. 1, e30379, Jan 2022 doi: 10.2196/30379. PMID: 34994706; PMCID: PMC8783277.

[4] K. Zahrai, E. Veer, P. W. Ballantine, H. P. deVries, and G. Prayag, "Either you control social media or social media controls you: Understanding the impact of self-control on excessive social media use from the dual-system perspective," Journal of Consumer Affairs, vol. 56, no. 2, pp. 806–848, 2022, doi: 10.1111/joca.12449.

[5] V. H. Isabella and E. Linos, "Social media for public health: Framework for social media–Based public health campaigns," Journal of Medical Internet Research, vol. 24, no. 12, e42179, Dec 2022, doi: 10.2196/42179. PMID: 36515995; PMCID: PMC9798262.

[6] K. Srivastava, S. Chaudhury, J. Prakash, and S. Dhamija, "Social media and mental health challenges," Industrial Psychiatry Journal, vol. 28, no. 2, pp. 155-159, Jul–Dec 2019, doi: 10.4103/ipj.ipj_154_20.

[7] E. Abi-Jaoude, K. T. Naylor, and A. Pignatiello, "Smartphones, social media use, and youth mental health, CMAJ, vol. 192, no. 6, e136-E141, Feb 2020, doi: 10.1503/cmaj.190434.

[8] D. Musleh et al., "A Machine Learning Approach to Cyberbullying Detection in Arabic Tweets," Computers, Materials & Continua, vol. 79, no. 3, pp. 1-22, 2024.

[9] J. Derges et al., "Mental Health Practitioners' and Young People's Experiences of Talking About Social Media During Mental Health Consultations: Qualitative Focus Group and Interview Study," JMIR Form Res, vol. 7, 2023, doi: 10.2196/43115.

[10] H. Hassan, H. M. Hsbollah, and R. Mohamad, "We examine the interlink of social media use, purchase behavior, and mental health," Procedia Computer Science, vol. 196, no. 1, pp. 85-92, 2022.

[11] P. M. Valkenburg, A. Meier, and I. Beyens, "Social media use and its impact on adolescent mental health: An umbrella review of the evidence," Current opinion in psychology, vol. 44, no. 1, pp. 58-68, 2022.

[12] L. E. Ugwu, E. S. Idemudia, O. O. Chukwu, and M. C. C. Onyedibe, "Measuring the Impact of Social Media on Young People's Mental Health: Development and Validation of the Social Media-Induced Tendency Scale," Depress Res Treat, vol. 2023, 2023, doi: 10.1155/2023/8677521.

[13] S Abbas, SA Raza, MA Khan, A Rahman, K Sultan, A Mosavi, "Automated File Labeling for Heterogeneous Files Organization Using Machine Learning." Computers, Materials & Continua 74 (2), 3263-3278, 2023.

[14] A. Rahman et al., "Supervised machine learning-based prediction of COVID-19," Computers, Materials and Continua, vol. 69, no. 1, pp. 21-34, 2021.

[15] R. Zagrouba et al., "Modelling and Simulation of COVID-19 Outbreak Prediction Using Supervised Machine Learning." Computers, Materials & Continua, vol. 66, no. 3, pp. 2397-2407, 2021.

[16] R. A. Naqvi et al., "Coronavirus: A "Mild" Virus Turned Deadly Infection," Computers, Materials & Continua, vol. 67, no. 2, pp. 2631-2646, 2021.

[17] A. Alqarni and A. Rahman, "Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach," Big Data and Cognitive Computing, vol. 7, no. 1, pp. 1-16, 2023.

[18] A. Rahman et al., "A neuro-fuzzy approach for user behaviour classification and prediction," Journal of Cloud Computing, vol. 8, no. 1, pp. 1-15, 2019.

[19] G. Zaman, H. Mahdin, K. Hussain, A. Rahman, J. Abawajy and S. A. Mostafa, "An ontological framework for information extraction from diverse scientific sources," IEEE Access, vol. 9, pp. 42111-42124, 2021.

[20] N.A. Sajid, M. Ahmad, A. Rahman, G. Zaman, M.S. Ahmed, N. Ibrahim et al., "A Novel Metadata Based Multi-Label Document Classification Technique," Computer Systems Science and Engineering 46 (2), 2195-2214, 2023.

[21] S. Arooj, M. F. Khan, T. Shahzad, M. A. Khan, M. U. Nasir et al., "Data fusion architecture empowered with deep learning for breast cancer classification," Computers, Materials & Continua, vol. 77, no.3, pp. 2813–2831, 2023.

[22] F. Jan et al., "Assessing acetabular index angle in infants: a deep learning-based novel approach," J. Imaging, vol. 9, no. 1, p. 242, 2023, doi: 10.3390/jimaging9110242.

[23] M. I. B. Ahmed et al., "Personal protective equipment detection: a deep-learning-based sustainable approach," Sustainability, vol. 15, no. 1, p. 13990, 2023.

[24] MS Farooq, S Abbas, A Rahman, K Sultan, MA Khan, A Mosavi, "A Fused Machine Learning Approach for Intrusion Detection System," Computers, Materials & Continua 74 (2), 2607–2623, 2023.
[25] M. I. B. Ahmed et al., "A deep-learning approach to driver drowsiness detection," Safety, vol. 9, no. 65, 2023.
[26] A. Rahman et al., "Geo-Spatial Disease Clustering for Public Health Decision Making," Informatica, vol. 46, no. 6, pp. 21-32, 2022.
[27] MA Qureshi et al., "Aspect Level Songs Rating Based Upon Reviews in English." Computers, Materials & Continua 74 (2), 2589-2605, 2023.
[28] T. A. Khan et al., "Secure IoMT for Disease Prediction Empowered with Transfer Learning in Healthcare 5.0, the Concept and Case Study," IEEE Access, vol. 11, pp. 39418-39430, 2023, doi: 10.1109/ACCESS.2023.3266156.
[29] D. Weber, "Mental Health and Social Media. An Explorative Mediation Analysis about the Relationship between Social Media Use, Value Concepts, Fear of Missing out, Sociodemographic Variables, and Mental Health," Journal of Consumer Health on the Internet, 27(4): 393-411, 2023.
[30] J. K. Parker et al., "Patients' Perspectives on Qualitative Olfactory Dysfunction: Thematic Analysis of Social Media Posts," JMIR Formative Research, 5(12), pp. e29086, 2021.
[31] https://www.kaggle.com/datasets/souvikahmed071/social-media-and-mental-health/data
[32] W. H. Hantom and A. Rahman, "Arabic Spam Tweets Classification: A Comprehensive Machine Learning Approach," AI, vol. 5, no. 3, pp. 1049-1065, 2024.
[33] A. Rahman et al., "Rainfall Prediction System Using Machine Learning Fusion for Smart Cities." Sensors, vol. 22, no. 9, p. 3504, 2022, https://doi.org/10.3390/s22093504.
[34] M. Farooqui et al., "A Deep Learning Approach to Industrial Corrosion Detection," CMC-Computers, Materials & Continua, vol. 81, no. 1, pp. 1-19, 2024.
[35] J. A. Naslund, A. Bondre, J. Torous, and K. A. Aschbrenner, "Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice," J Technol Behav Sci., vol. 5, no. 3, pp. 245-257, Sep. 2020.
[36] M. N. Alnuaimi et al., "Transfer Learning Empowered Skin Diseases Detection in Children," Computer Modeling in Engineering & Sciences, vol. 141, no. 3, pp. 1-15, 2024.
[37] M. S. Ahmed et al., "Joint Diagnosis of Pneumonia, COVID-19, and Tuberculosis from Chest X-ray Images: A Deep Learning Approach." *Diagnostics*, vol. 13(1), p. 2562, 2023.
[38] M. Gollapalli et al., "Appendicitis Diagnosis: Ensemble Machine Learning and Explainable Artificial Intelligence-Based Comprehensive Approach." Big Data and Cognitive Computing, vol. 8, no. 9, p. 108, 2024.
[39] S.M. Alotaibi, Atta-ur-Rahman, M.I. Basheer, and M.A. Khan, "Ensemble Machine Learning Based Identification of Pediatric Epilepsy," Comput. Mater. Contin., vol. 68, no. 1, pp. 149-165, 2021.
[40] A. Rahman et al., "Diabetic Retinopathy Detection: A Hybrid Intelligent Approach," Computers, Materials and Continua, vol. 80, no. 3, pp. 4561-4576, 2024.
[41] A. Rehman et al., "Modelling, simulation, and optimization of diabetes type II prediction using deep extreme learning machine," Journal of Ambient Intelligence and Smart Environments, vol. 12, no. 2, pp. 125-138, 2020.
[42] N. M. Ibrahim et al., "Transfer Learning Approach to Seed Taxonomy: A Wild Plant Case Study." Big Data Cogn. Comput., vol. 7, no. 128, 2023.
[43] A. Alhashem et al., "Diabetes Detection and Forecasting using Machine Learning Approaches: Current State-of-the-art," IJCSNS - International Journal of Computer Science and Network Security, vol. 23, no. 10, pp. 199-208, 2023.
[44] M Gul, IA Khan, G Zaman, A Rahman et al., "A Game-Theoretic Approach to Safe Crowd Evacuation in Emergencies," CMC-Computers, Materials & Continua 79 (1), 1631-1657, 2024.
[45] Alabbad, D.A.; Ajibi, S.Y.; Alotaibi, R.B.; Alsqer, N.K.; Alqahtani, R.A.; Felemban, N.M.; Rahman, A.; Aljameel, S.S.; Ahmed, M.I.B.; Youldash, M.M. Birthweight Range Prediction and Classification: A Machine Learning-Based Sustainable Approach. Mach. Learn. Knowl. Extr. 2024, 6, 770-788. https://doi.org/10.3390/make6020036.
[46] IA Qureshi et al., "GFuCWO: A genetic fuzzy logic technique to optimize contention window of IEEE-802.15. 6 WBAN," Ain Shams Engineering Journal, 102681, 2024.
[47] D.M. Althawadi et al., "Exploring Efficient Solutions for the 0/1 Knapsack Problem," IJCSNS - International Journal of Computer Science and Network Security 24(2): 15-24, 2024.
[48] M. M. Qureshi, F. B. Yunus, J. Li, A. Ur-Rahman, T. Mahmood and Y. A. A. Ali, "Future Prospects and Challenges of On-Demand Mobility Management Solutions," in IEEE Access, vol. 11, pp. 114864-114879, 2023.
[49] I. Alrashide et al., "AIMS: AI based Mental Healthcare System," IJCSNS - International Journal of Computer Science and Network Security 23(12): 225-234, 2023.
[50] M. Gollapalli, A. -U. Rahman, A. Osama, A. Alfaify, M. Yassin and A. Alabdullah, "Data Mining and Visualization to Understand Employee Attrition and Work Performance," 2023 3rd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia, 2023, pp. 149-154.
[51] Musleh DA, Olatunji SO, Almajed AA, Alghamdi AS, Alamoudi BK, Almousa FS, Aleid RA, Alamoudi SK, Jan F, Al-Mofeez KA, et al. Ensemble Learning Based Sustainable Approach to Carbonate Reservoirs Permeability Prediction. Sustainability. 2023; 15(19):14403.
[52] R.A. Qamar, M. Sarfraz, A. Rahman, S.A. Ghauri, "Multi-Criterion Multi-UAV Task Allocation under Dynamic Conditions," Journal of King Saud University-Computer and Information Sciences 35 (9), 101734, 2023.
[53] N. AlDossary, S. AlQahtani, R. Alzaher, A. Rahman, "SYN Flood DoS Detection System Using Time Dependent Finite Automata," International Journal of Computer Science & Network Security 23 (6), 147-154, 2023.
[54] Z. Alsadeq, H. Alubaidan, A. Aldweesh, A. Rahman, T. Iqbal, "A Proposed Model for Supply Chain using Blockchain Framework," IJCSNS - International Journal of Computer Science and Network Security 23(6): 91-98, 2023.

[55] A. Albassam et al., "Integration of Blockchain and Cloud Computing in Telemedicine and Healthcare," IJCSNS - International Journal of Computer Science and Network Security, 23 (6): 17-26, 2023.
[56] Sajid, N.A.; Rahman, A.; Ahmad, M.; Musleh, D.; Basheer Ahmed, M.I.; Alassaf, R.; Chabani, S.; Ahmed, M.S.; Salam, A.A.; AlKhulaifi, D. Single vs. Multi-Label: The Issues, Challenges and Insights of Contemporary Classification Schemes. Appl. Sci. 2023, 13, 6804.
[57] Gollapalli, M.; Rahman, A.; Alkharraa, M.; Saraireh, L.; AlKhulaifi, D.; Salam, A.A.; Krishnasamy, G.; Alam Khan, M.A.; Farooqui, M.; Mahmud, M.; et al. SUNFIT: A Machine Learning-Based Sustainable University Field Training Framework for Higher Education. Sustainability 2023, 15, 8057. https://doi.org/10.3390/su15108057.
[58] Talha, M.; Sarfraz, M.; Rahman, A.; Ghauri, S.A.; Mohammad, R.M.; Krishnasamy, G.; Alkharraa, M. Voting-Based Deep Convolutional Neural Networks (VB-DCNNs) for M-QAM and M-PSK Signals Classification. Electronics 2023, 12, 1913. https://doi.org/10.3390/electronics12081913.
[59] Musleh, D.; Alotaibi, M.; Alhaidari, F.; Rahman, A.; Mohammad, R.M. Intrusion Detection System Using Feature Extraction with Machine Learning Algorithms in IoT. J. Sens. Actuator Netw. 2023, 12, 29.
[60] Alghamdi, A.S.; Rahman, A. Data Mining Approach to Predict Success of Secondary School Students: A Saudi Arabian Case Study. Educ. Sci. 2023, 13, 293. https://doi.org/10.3390/educsci13030293.
[61] Alhaidari, F., Rahman, A. & Zagrouba, R. Cloud of Things: architecture, applications and challenges. J Ambient Intell Human Comput 14, 5957–5975 (2023).
[62] Rahman, A. GRBF-NN based ambient aware realtime adaptive communication in DVB-S2. J Ambient Intell Human Comput 14, 5929–5939 (2023).
[63] Ahmad, M. et al. Enhanced query processing over semantic cache for cloud based relational databases. J Ambient Intell Human Comput 14, 5853–5871 (2023).
[64] Basheer Ahmed, M.I et al., A Real-Time Computer Vision Based Approach to Detection and Classification of Traffic Incidents. Big Data Cogn. Comput. 2023, 7, 22.
[65] M Jamal, NA Zafar, D Musleh, MA Gollapalli, S Chabani, "Modeling and Verification of Aircraft Takeoff Through Novel Quantum Nets." Computers, Materials & Continua 72 (2), 3331-3348, 2022.