How social media impacts the well-being of students, including mental health

Table of Contents

Table of figures

List of Tables

Acknowledgements

I would like to thank Mr. Faraz Janan for guiding me with all the support and dedication for completion of this major project as my supervisor. Overall, a special gratitude to Anglia Ruskin University and the students who helped with filling the survey as participants for giving their valuable feedback and time. The readings and review papers are collected and utilized from Google scholar and appropriate references are attached at the end of this file.

This achievement would have not been possible without all of you.

Thank you.

Abstract

Human life after social media use has rapidly changed to a certain extent and the usage level is increasing daily. The social media use survey conducted throughout this study focuses on complications where students who are staying in the U.K. on international basis deal with their daily life on social media. The process followed to gain a conclusion was started with collecting data in terms of responses and then pre-processing the information for hypothesis testing and emotional analysis. Machine learning technique was also a part of this research which was used to generate emotional analysis which is performed in the methodology section. Instagram, Facebook and WhatsApp were the three majorly used social media platform amongst the participants. Average reported time was around 2 hours daily, whereas more than 2 hours on average were spend on Facebook and Instagram applications. The survey was conducted on students aged 18 years and above and this study has gained qualified ethical approval from the department of Computing and Information Science SREP at Anglia Ruskin University. There was no direct involvement of individuals as the feedback form was created on Google forms and supplied to the participants online.

More than 70% reported that they do not face any negative health impacts mentally after using any of their favourite social media platform. Well-being results discovered were able to connect some physical body issues after prolonged use. The Chi-square test was able to draw a graph based on the recorded answers in which the result of mental effects was close to the significance value. Out of the 37 personal feedback received, the emotion detection method was able to divide the emotions in three classifications and found equal results in the state of being happy and sad. The background exploration was able to determine some socially connected concepts of human life along with providing sufficient information on the different processes worked in this theory. It also included longitudinal studies which displayed how and what factors are considered for performing it which can be favourable in future studies. After summarizing all the procedure, the approach was unable to find a direct connectivity in human health and use of social media. Positively, the collected data which was anonymous has the capability of known symptoms of using some of the highly depleted social networking websites.

Introduction

Human lifestyle nowadays is becoming technology oriented as social media is linked and constantly developing its extent by excess mechanisms and applications. One of the studies divided into active and passive usage suggests that extreme use of social media has adverse effects on well-being whilst the other indicates major affirmative outcomes such as social networking and individual well-being. Loneliness and social comparison can be the associative parameters on utilizing these types of technologies (Marttila *et al*., 2021). A well-being psychological state that helps manage stress in life and identifying own capabilities of learning and working is defined to as psychological health. The contribution to mental health is considered as a human right which is crucial to oneself, society and socio-economic enhancement (World Health Organisation, 2022). According to Rao and Kalyani (2022), social media can be called as a publicly available online platform for describing personal or social views and concerns. Students can use this medium to connect with relatives and professionals. Entertainment factor is also derived from this as a creator is gaining business needs through it. A society-based impact of social media in a convinced approach is linked with increased connectivity amongst users, scholastic purpose and updates related to the society.

Whilst getting hacked, cybercrimes and obsession can be some of the major drawbacks of it. Dependency of this technology on human beings as a daily routine has shown consistency along these years. Over 20% of the youth (18-25 years) is subject to lonesomeness constantly as stated by a survey conducted in the U.S., Australia and U.K. which derives a question on confidence required for face-to-face interactions. Belonging on the other side can assist in developing social contacts to overcome factors linked to dignity towards yourself (Smith *et al*., 2021). Concerning the well-being of young people is one of the challenging aspects for society where depression and suicide are considered as the harmful causes linked to it. According to a study done in China which consists of 542 participants (adolescents), it is stated that individuals who were lonely had less social networking as compared to those who weren't friendless. It is necessary to address adolescents about the draw backs and advantages at the start of their pathway towards using the social media network to conserve their well-being (Webster *et al*., 2021). As per Beyens *et al*. (2020), in research done in real-time assessment it was found that out of the 100%, majority of them felt not better to better after using passive social media whilst around 10% felt the other.

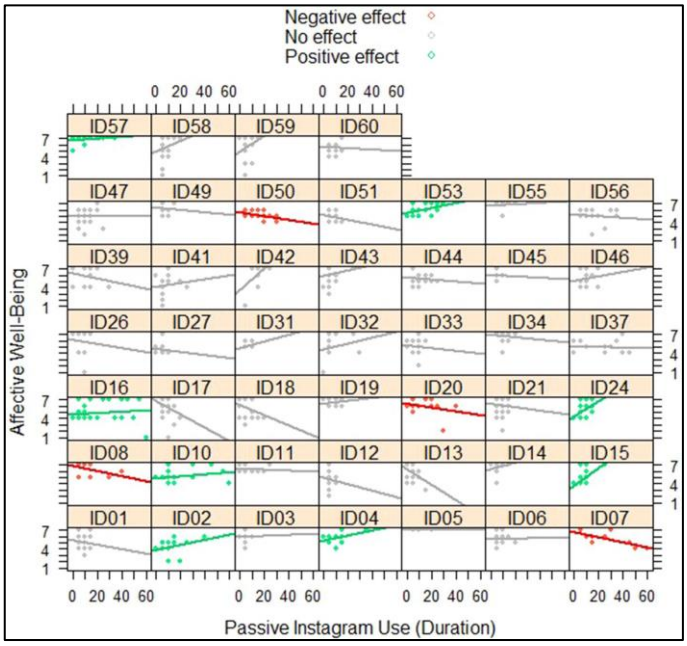


*Figure 1 – (It defines the use of Instagram (one of the leading social media platforms) in a passive way where the colour of the lines differs and has a specific set of interpretation. The green line is for positive effect and the red one is for negative impact. Lastly, the grey line demonstrates irrelevant connections (*Beyens *et al.,* 2020*)*

Fallacious information on these types of networks can create false belief and hate against the chosen criteria. The number of these platform users with mental problems like getting depressed is higher in youngsters as compared to other ages. Smartphone technologies are consisted with various global applications including social networking applications which gives the consumer a hand worth experience in getting connected easily. A general differentiation is mentioned in the below table no. 1 describing advantages and disadvantages of social media platform which can be beneficial for this study.

Advantages	**Disadvantages**
Public connections	Mental stress
Support from facilities through peer network	Harassment and cyberbullying
Services promotions	Routine disturbance

*Table 1 – General advantages and disadvantages of social media (*Naslund *et al.,* 2020*)*

The first important step is to stay conscious from the risks coming from this platform and then nurturing the advantages as stated below (Naslund *et al*., 2020). There are numerous platforms related to social media, however Facebook is the most researched networking site for loneliness or social fear relations. Younger generations majorly use platforms like Instagram and Snapchat which are also designed for many social purposes (O'day and Heimberg, 2021). Communications via social media are flexible and users can freely interact with each other

where the market demands for these types of platforms increase time to time. User data privacy is most of a concern for the consumers as data breach can cause confidentiality issues and the data can be exposed quickly. The premium features of some applications allow users to upgrade their platform in a more convenient manner, however it can prove to be expensive for some of the users (Appel *et al*., 2020).

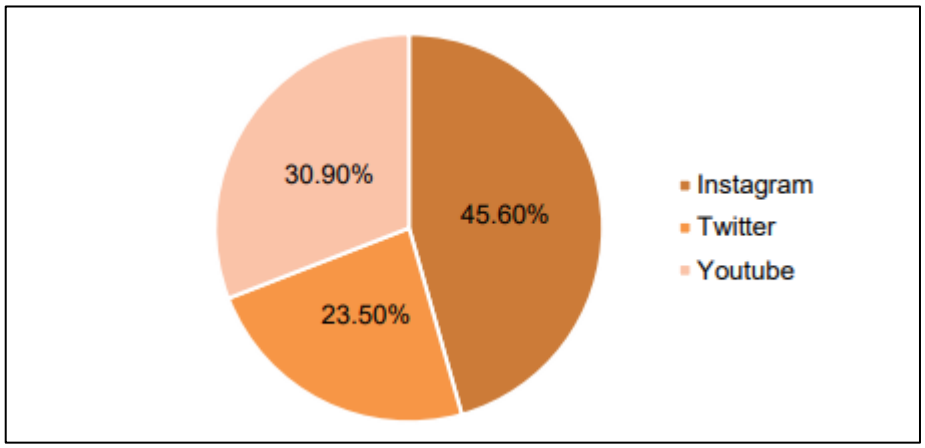


*Figure 2 – Most frequently opened social media platform (*Maharani, 2021*)*

In a google form survey study as per the figure 2., it is indicated that Instagram is the most opened platform on a regular basis in comparison of the other two sites as it consists of information on other people lives who keep it up to date. Influence can easily be drawn from one end if there is more common content associated with one's daily life. As YouTube is also one of the trending networking sites, it is easier for people to connect on Instagram because of its privacy chat and call feature. Lastly if we go through the Twitter site, it has either one of the features from the other two platforms (Maharani, 2021).

An 8-year recorded study specified that the participants were examined based on time spent on these social networking sites linking with depression and anxiety in adolescents (aged 13-20), however the results show that there was no such increase in the above symptoms with heavy time spent on social media. The e-learning platform has taken a jump in terms of education and knowledge providing asset where students of any age can get connected with their study tutors whenever convenient and without hesitation. Thus, the positive approach can be depicted for individual distribution and expert use. Facebook and Twitter have been in consideration for educational advantages since a long period of time. Other than that Instagram and Pinterest have also superior graphical capability in comparison to other sites. The fact that social media is growing, and many such websites are in competition of each other in terms of features and business market (Manca, 2020). With reference to Karim *et al*. (2020), the social media and mental health studies should be conducted in more depth as the information can be more useful for the medical and social science departments when patients are consulted for

any mental illness. The research information is gathered from various sources available and sometimes from personal recorded feedback which can be a necessary factor for such kinds of treatments if the practise has any success rate.

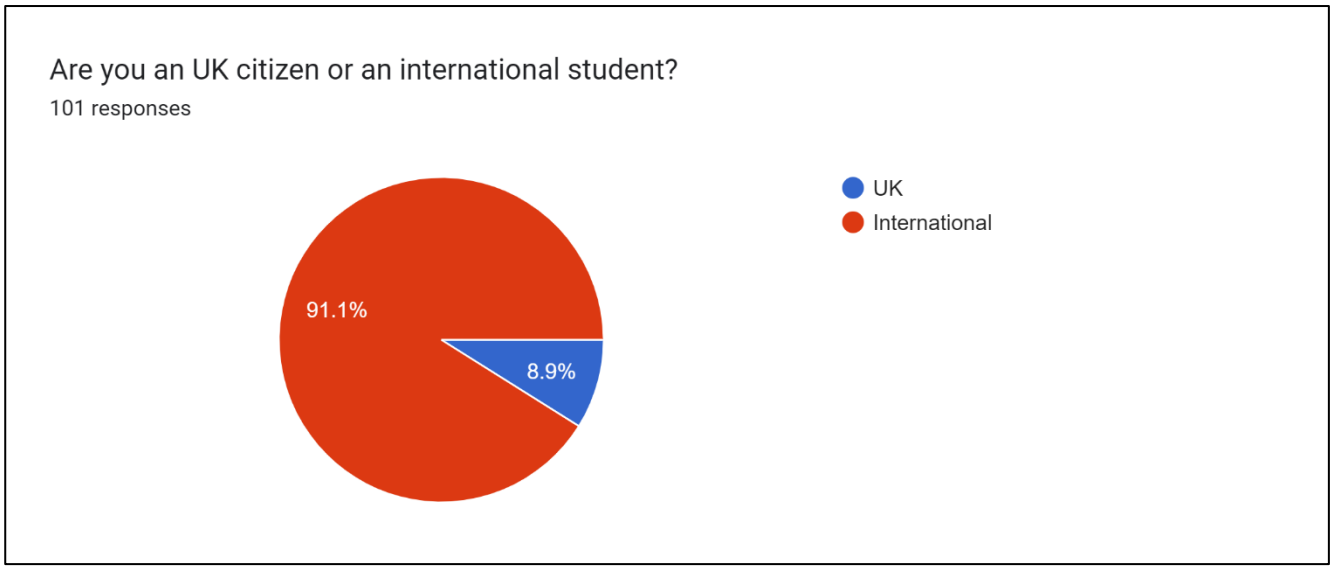


Figure 3 - Percentage count of the participants differentiating International and UK resident students

This development is based on detecting the social media effect on the intellectual healthiness and well-being of students studying in the United Kingdom and modelling the received feedback which can assist in gathering solutions to the mentioned factors if any. Analogy of this study with previous approaches can filter out missing concepts and aim to deliver the anticipated output based on real-time opinions of the individuals. Many international students pursuing their education in the UK away from their home countries manage their daily part of the life where social media also plays a crucial role for information on various aspects including their lifestyle, studies or future endeavours. We ignore the part that many of these platforms are linked with each other, and one gains interest about the other through connectivity of these sites. An example of Facebook and LinkedIn can be related to each other, yet both are created for different purposes (Aichner *et al*., 2021).

This study will try to conclude the use of these social media networks in student's life where specific approach is to understand the usage and effects derived from it. International exposure with other students coming from different parts of the world and from UK can help students to gain knowledge about opportunities in career and other aspects of life. The other side can lead to denial from accepting some facts on these networking sites involving mental stress to take part as a negative impact. Information mislead can lead to belief in certain articles or news without doing proper research. Previous studies have been focused on adolescents and young generations, where the part of staying away from blood relations, friends and earlier connections is ignored at a time. Loneliness can be the cause of getting used to social media

for these students. The representation is based on a questionnaire where the students are asked to fill a survey and answer some questions related to social media use (Al-Rahmi *et al*., 2021). Participants are asked to respond on questions associated with most used platform, along with any mental or physical stress after use. This survey is passed to students studying in the UK to collect personal feedback from them about their views on social media. The responses will be recorded and can be useful for deriving an outcome on any side effects associated with social media. The collected data will be then compared with previous studies and possibilities could be made from it. On an overall basis, this study covers the negative health impacts about social network site use as we all are aware of the technology standard in the current world (Braghieri *et al*., 2022).

Aims

The study aims at the effects of social media on human's mental behaviour and well-being. This study is designed using the sources and evidence collected from previous studies about using social media use and its impacts.

Objectives

- The main purpose of this approach is to comprehend the factors affecting mental and well-being status in humans' and modelling the received reactions from the participants which can assist in future explanations of the problem.
- Aware responses on using social media and the effects on well-being and mental health are collected from UK citizen students and international students on specific basis studying in the university, adapting a new lifestyle than before and managing their overall activities.

Background

- **Historical exploration**

According to Raza *et al*. (2020), the emerging social media has diverse effect on how the people communicate with each other and their behaviour can be tested on this basis. Life satisfaction study has been conducted to understand how students are studying in the university influenced to use social media. Later, the outcomes are categorized based on social information overload and advantages. It is also stated that social benefits like study groups, friends and other professional factors help in building life satisfaction whereas on the other hand, overloading of this type of data can cause emotional stress. The simple conceptual diagram below explains seeing many social terms to link with social media usage and then segregated into gains and problems towards happiness in life.

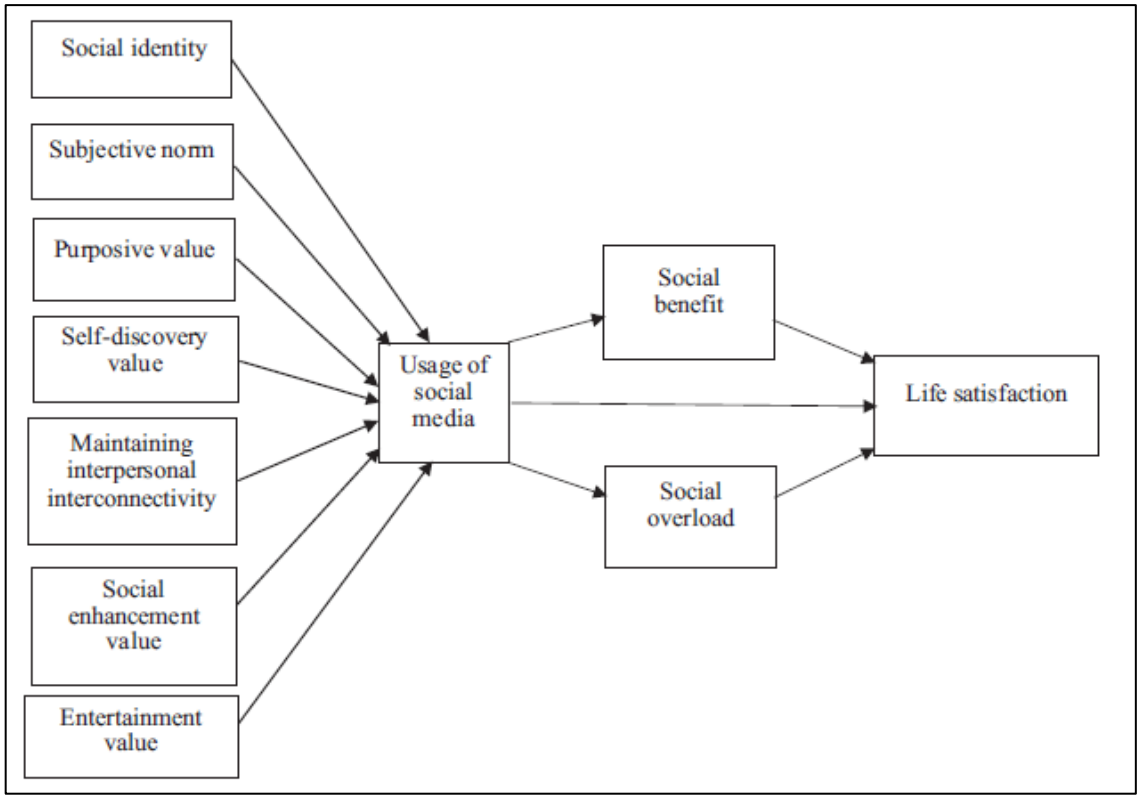


*Figure 4 - Social concepts in human life (*Raza *et al.,* 2020*)*

Social media users generate a huge amount of data which can lead to integrity issues with their personal details and day to day activities. This type of data is referred as ‘Big data’ on which there are much research because of the internet activities. This big data analysis is one of the techniques to a key to human behaviour which are stored by many companies, individuals or government authorities. The approach starts with investigation of this data followed by background review and then the techniques are compared and finally the challenges related to big data analytics are described. One of the major challenges to this may

be the wide amount of data which is loaded on daily basis and keeps on fluctuating. We all are aware that to use any social networking site, there is a need of a device (e.g. mobile, tablet or a screen device) and the time spent for using these websites on the device is referred to as 'screen time'. Therefore, there is a need to focus on the screen time associated with physical activity as well because long use of screens is a worrying circumstance in terms of mental health and physical performance. One of the findings focuses on increase in factors that can decrease the use of another, where involving more in usage of device screen will minimize the use of physical activities which always are positively connected to physical health. To break this chain there should be a promotional awareness to broaden the physical health aspects in daily life (Lavados *et al*., 2023). The COVID-19 pandemic is one of those periods, all of us can hardly forget in these years. It lacked physical behaviour in humans and directed a way towards social media as time spent on social media increased during this pandemic because of less social interactions as it was suggested to stay at home for precautionary safety. The findings about social media use and stress during COVID-19 led increase in users of these social networking sites in a positive manner. Hence, COVID-19 stress was increased and the risk of social media addiction was at an increased rate (Zhao and Zhou, 2021).

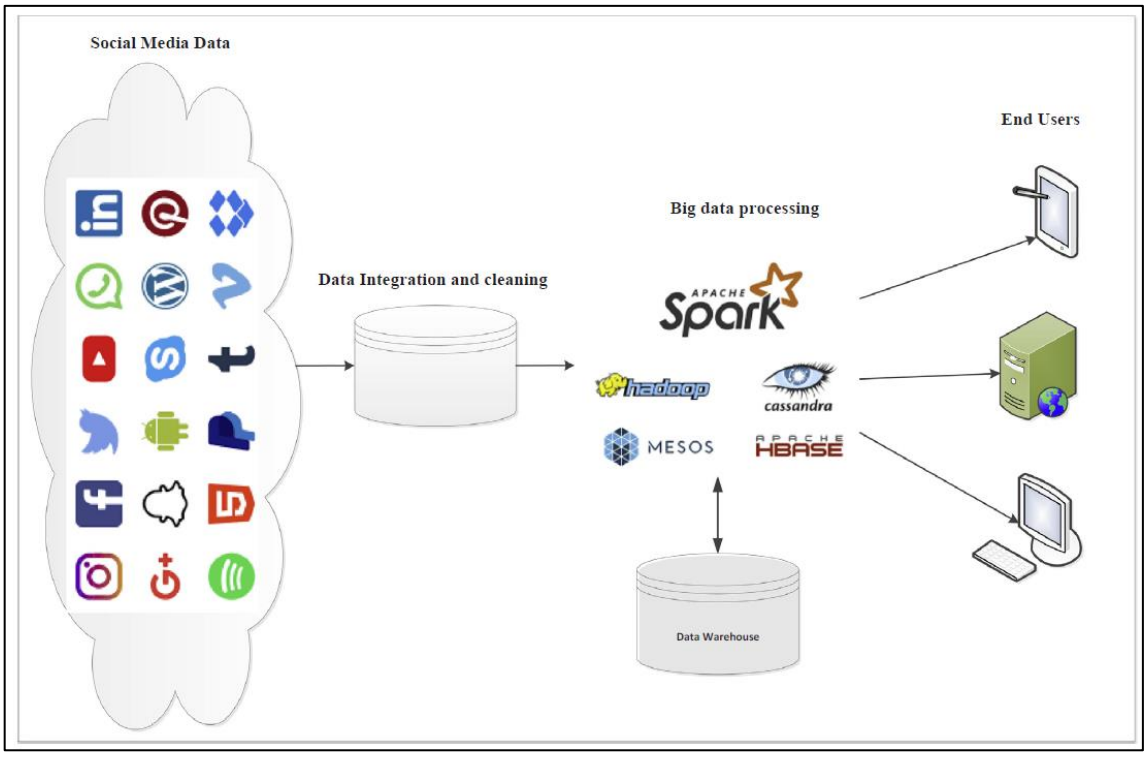


*Figure 5 – Data processing (*Ghani *et al.,* 2019*).*

This figure gives a rough idea on how the data from many social media websites is picked for integration and processed with the help of tools and then sent to the end users. It is one of the types of filtering data before it is sent to the viewers so that any unrelated information is cleared before it is received by the user (Ghani *et al*., 2019). Interpersonal communication has become

an important need for human beings and the rapid growth in social media has changed ways of this communication practice. And there is a need to evaluate the relations between this social media networks and college students. One of those practice on the college students said that it has a negative impact on mental health and academic performances of the participants. However, lowered self-esteem as a part of reduced mental health cannot be ignored in the analysis outcome (Hou *et al*., 2019).

One of the research articles also evaluated the active and reflexive use social media with association of negative and positive impacts on well-being of humans on websites like Facebook, Instagram, Twitter, TikTok during the COVID-19 time. This research then indicated that passive use of Facebook website resulted in lower well-being. On the other hand, active use of Instagram came with neutral effects. The combination of the outcome stated that every social networking site should be used with avoiding the risk attached with it (Masciantonio *et al*., 2021). A mental health examination study in Indonesia about students engaged in social media during COVID-19 pandemic states that more use of these platforms is connected to some amount of depression level. Mental health in students who were more involved in communication with their parents and maintaining healthy relations found out to better as compared to the previous group. The online learning platform opened ways for the students to use more internet and globally used social networking platforms during the COVID – 19 period. The data suggested that age group between 16 to 24 years are engaged in social media around 8 hours per day. There is a need for the universities and assigners of the policies to provide more knowledge about the mental health and harms with social media use (Sujarwoto *et al*., 2023).

- **Procedural analysis**

A statistical learning has stated that one-third of the people living on earth are social media users and it is still counting. Problematic social media use can occur in some users as the preferences in using social media differs from person to person. Accordingly, a survey was conducted on graduate students aged 17 – 32 years on social media motives. The problematic social media use score was found higher in consumers whose social networking preference was Instagram, Facebook and Snapchat. The usage and satisfaction level of these social sites can differ from age, gender and personality. It can also be varied on the type of social network. If we take an example of Instagram which can be used for social connections. And on the other hand, Snapchat is useful for instant environmental communications. To find the relations between age, sex along with current social media sites and the need of fulfilment has also been studied. There was then made a conclusion on social media use problems. The approach used was how the social media is utilized and satisfies the person using it which made predictions

on whether the use is motivational, just for satisfaction, any use for social and mental factors and finally with relational communications. The figure below shows an assumption in table view of social media platforms used daily in hours which is measured in percentage to give a generic ideology associated with the platform use. It was then observed that WhatsApp is the most

Variables	Percent
Whatsapp	92
Instagram	79
Youtube	78
Facebook	55
Snapchat	37
Google	37
Twitter	32
Daily internet use (0–2 h)	14
Daily internet use (2–4 h)	41
Daily internet use (4 h or more)	45
Daily mobile phone use (0–2 h)	13
Daily mobile phone use (2–4 h)	35
Daily mobile phone use (4 h or more)	52

*Figure 6 – Daily use of social media platforms (*Kircaburun *et al.,* 2020*)*

used social site followed by Instagram, YouTube and others. The daily use in the table above also is sub divided into internet and mobile phone use. Out of the 68 students participated in the theory, 44% were seen using internet for more than 4 hours and 51% were involved in mobile phones on a regular basis. If classified gender wise, it is observed that female genders are more into using these platforms for informative and educational purposes, relationships maintenance and managing tasks. Socializing themselves and entertainment use, etc are some detailed usages amongst men. This gender-based study determined that problematic use is more attached with platforms like Facebook, Instagram and Snapchat (Kircaburun *et al*., 2020). One of the major questions i.e. does time spent on social media has effects on mental health? was covered in an article where an 8-year research was done for clarification of impacts in prolonged period. The limitations of other research which are not considered with time factor was an approach of this theme. Youngsters aged between 13 and 20 were involved in the participation in which a questionnaire was asked to fill once in a year time. The questions were measured in three aspects mainly as time spent, depressive syndrome and anxiousness. Below is the classification table derived for understanding of the questionnaire.

Factors	Questions and measures
Time spent on social media	How much time do you spend on social media sites?
Depression	It was considered on loneliness and negative feeling.
Anxiety	This was determined on concerning over things.

*Table 2 – Factors affecting social media use with some research questions (*Coyne *et al.,* 2020*)*

A graphical figure analysis shows how depression plays a role with three participants scale differently. The corelative points are 8 as per the longitudinal study for 8 years. And the dash line indicates that all these years the results have been varied because of depression scale present on the Y – axis. The eight points are not in a straight line as the responses have been different from each other and no straight dash line can be seen. This is how different characteristics plays an important role while obtaining responses from the participants. The

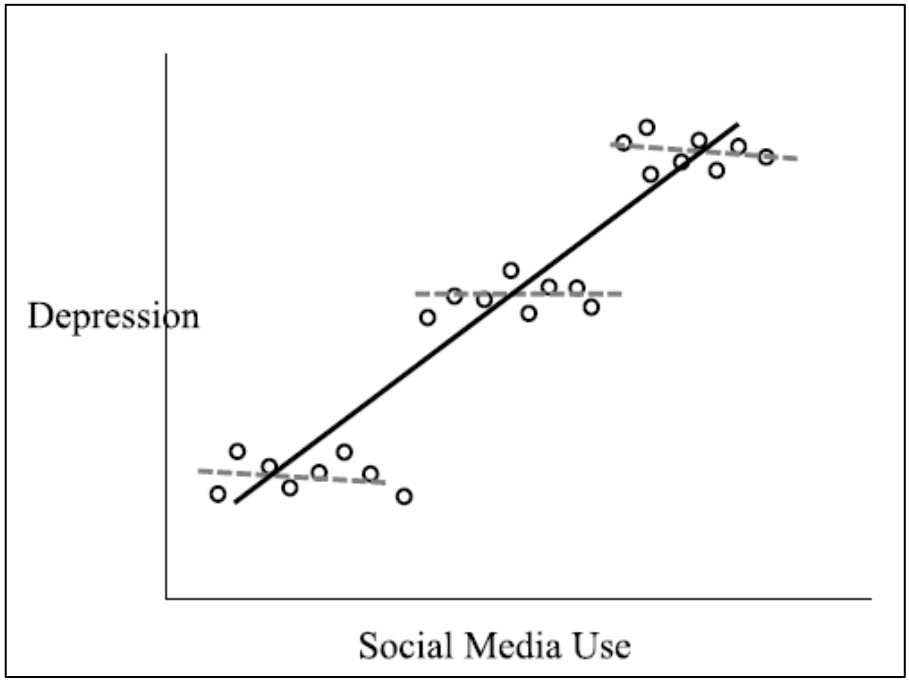


*Figure 7 – Graphical representation of depression symptom with social media usage (*Coyne *et al., 2020)*

way of thinking and state of mind is not only different amongst others, but it changes with time. There was no link found between social media use and mental health as an outcome from this analysis (Coyne *et al*., 2020). Following the research by Plackett *et al*., (2023), an observation was carried out for a period of 10 years. The objective was clear and to check the relationship between using social media and mental problems in youngsters, so participants of aged 10 to 15 years were chosen for this research. It was carried on information wave method from wave

1 to 10 for a period of 10 years. The measurement of social media usage was done by providing answers of two questions which were:

- Does anyone have its place in using social websites like Bebo, Facebook, or Myspace?
- And how many hours does one operates social media in a normal day?

The two other parameters involved in this study were calculated on self-confidence and fellow relations. The findings did not find any connection between social media and adolescents mental health. However, 10 years is a long period of time for a study to calculate all the results and give analysis report.

The table below classifies the parameters grounded on age.

Age group	Classification of study parameters
12 – 13 years	Time spent on social media platforms
13 – 14 years	Self-esteem level
14 – 15 years	Mental health issues

*Table 3 – Age parameters and their classification used in the referred study (*Plackett *et al.,* 2023*)*

Difficulties occurred on outdated platforms like Bebo and Myspace and thus the way of providing feedback can differ in the participants. It also stated that active and passive using can have different impacts on one's mental health. People have diverse mindset when it comes to age, gender or any specificity. Depending on the age and responsibilities, it is also possible that the way of handling the social media can differ from person to person. There is a statistical study in India which includes Chi square test methods which demonstrate that students engaging more time on social networking sites have reduced record in academic terms. The survey was conducted with 277 participants on gender basis. It covered positive and negative effects of using social websites on academics. The findings involved students spend more time on these websites than book reading activities and physical activities along with sleep disturbance because of the use which can lead to an inactive daily life. It also helped in understanding the website which students are involved of use and their spent time with known purpose. Hypothesis testing was conducted, and an overview is represented below which shows the testing analysis as per figure no. 8. The values found were less and equivalent with the significance level and the results showed some amount of evidence as per the studied responses (Behera *et al*., 2023).

Attributes	Hypotheses	x^2-value	Sig. ($\alpha = 0.05$)	Decision
Academic Achievements & Total Hours SNSs	H_{50} = Academic achievements are independent of total hours spend on SNS. H_{51} = Academic achievements depend on total hours spent on SNS.	35.038	0.241	H_{50} : Accepted
Academic Achievements & Gender	H_{60} = Academic achievements are independent of Gender. H_{61} = Academic achievement depends on Gender.	12.139	0.059	H_{60}: Accepted
Academic Achievements & Total Hours of Sleep at Night	H_{70} = Academic achievements are independent of total hours of sleep at night. H_{71} = academic achievements depend on total hours of sleep at night.	30.034	0.184	H_{70}: Accepted

('*' implies chi-square is significant at $\alpha = 0.05$)

*Figure 8 – Chi square test analysis referred in a study (*Behera *et al.,* 2023*).*

Social media can come with both difficulties and opportunities for students, and it depends on them on how they can make a positive outcome in the real world. The distraction levels in these sources are typically high and can easily affect a person's state of mind towards it. Positively, online features like sharing of files can be beneficial for student participation in important aspects. Sharing information related to study can be a key feature of social media for students which can contribute to knowledge sharing. This networking platform can increase students' contribution level and assist in appropriate motivation and support towards their career (Sivakumar *et al*., 2023).

- **Primary findings**

False information on the social networking platforms can be spread further easily from person who has received it. This type of information can cause misleading in human behaviour. Usually, people who get the wrong information are likely to think it as true or having same materialistic thinking. Some of the negative effects can lead to political issues, trust problems, and can affect the thoughts of a society too. Changes in attitude can also occur in an individual by false information. Further, if this data is distributed on social media, it can spread quickly and continuously and hence these sites are the top target for the person. These can be done for personal comfort, or with wrong thoughts towards someone and making things viral even when they are not real. Sometimes the information is properly investigated and forwarded whilst another time it is shared quickly based on personal experience. There is a lesser number of people spreading this fake news and it is just 10% of the total individuals who share overall information. It is also stated that mostly younger and minimal educated individuals are likely to share such type of data (Buchanan, 2020).

One of current research done in the U.S. done by (Olorunsogo *et al*., 2024), has been determined based on positive and destructive factors related to societal broadcasting consumption. It implies that mental problems from social media can vary in people with different ages. There is a need to find mitigations related to mental health which come from using either of the online social media platforms. Cyber-bullying is a risk of using this platform without any

data security and can relate to mental stress or depression in individuals and can put professional sensitive information at risk if occurred at an organization level. More set of rules should be placed with respect to these platforms after seeing the rapid increase in current world technologies. Ruling bodies should set some proper criteria of using these platforms along with support for frauds and other risks that come it. Apart from that, connectivity around the world and staying updated with the current situations are some positive outcomes which should be enlightened and can change user level mindset. Addiction, privacy issues and information mislead are the challenges of using social media and more focus should be provided to these types of encounters. Some studies say that young age is the period where challenges related to mental health can often occur and these heavy use of social media platforms can also lead to problems such as inadequate body posture.

The life of adolescents is quite challenging in today's world if it has no social media knowing the fact that these young generations use their mobile phone and the other devices linked with social media in a large manner which can create misdirection in their school routine activities and communal connectivity functions. Harm to self and suicidal causes are also at urge by the usage of these social platforms. It is also asserted that less social connections can lead to psychological stress and can impact mental health of a person. According to the information available on the internet, the figure no.9 indicating a table below can predefine the social media platforms and their potential usage applications which provide a general overview of the application process in the current world.

Social media applications	Examples
Social networks	Facebook, Twitter, Instagram, Snapchat
Media sharing	WhatsApp, Instagram, YouTube, Snapchat, TikTok
Messengers	Facebook Messenger, WhatsApp, Telegram, Viber, iMessage
Blogging platforms	WordPress, Wikipedia
Discussion forums	Reddit, Twitter
Fitness & lifestyle	Fitbit

*Figure 9 – Tabular representation of applications and their resources (*Khalaf *et al.,* 2023*)*

The different purposes that are served by these platforms are depicted above and based on their use the networking websites are listed in parallel. The community relationships are directly or indirectly linked to a person's mental and physical health and can play an important role in preserving the health factors. Connecting with people socially at various places like home, study place or any other gathering places has become a normal agenda in adolescents where

these sites offer them to broadcast their own photographs and videos on their own. They also get a chance to express themselves by showing and gathering interests which they typically like along with sharing or posting contents available without any limitations. Trolling is also an attached term to these platforms where lifestyle, speech and body language are at risk which is mostly unknown by some of these users. It can arise concerning issues like thinking of suicide, feel to stay alone, and other depression causing syndromes. All the problems should be provided solution by understanding the cause and well-being should be graded in terms of this social media use. The digital world has been less harming to one on an average, but it cannot comply with every circumstances. The debate on using these socio-platforms has been always equal as the benefits and drawbacks are always changing and evolving (Khalaf *et al*., 2023).

The aged between 17 – 25 years can be crucial which can increase the effect of mental, physical and other body related problems. This can also include behavioural and social changes in an individual. Social media comes with better relationships and innovations whereas it can also result in body image issues with people tending to use it frequently. There has been lesser evidence that can be able to conclude that the social media use is associated with body posture problems as it has been only based on theories. A 4-week study was conducted on undergraduate students through an online portal whose purpose was to evaluate the impacts of reducing social media use on mentally related problems and it was kept undisclosed from the applicants. There was a division in the participants in two groups where the usage was reduced up to 1 hour in one and the other was given permission to use it with no limitations. The purpose of the study was then made clear virtually when the participation process was completed successfully. It was then presumed that reduction in hours on social media can benefit in increase in one's body image value (Thai *et al*., 2023).

Nowadays, people are engaged in text-based contents, photos, audio sounds and videographic visuals to share their views on social media. It means that social media platforms have been providing these features to satisfy user needs. Textual analysis is a large amount of information to be read by the system to give user the response received from another user. The count of users texting another is massive, and this data moves so quickly that the systems are configured so drastically for understanding this natural language process. A text can show many emotions, and which is why there is a need for sentiment and upgraded emotional analysis which can read and provide meaningful service to the consumer. Feedback collection is a method where one can give their opinion about a service or a person working on behalf of the service which can help the organization involved in providing services to learn and understand how a user reacts to their used product. Emotional testing can play a vital role in

learning reaction received from a person on a specific product. For example, a student can provide feedback about the teacher who delivers knowledge on a specific subject, and it is important for the contributor which can assist in future work and previous received opinions on the other hand can be helpful when creating a study plan. A model is referred below which depicts the dimensions of human emotions.

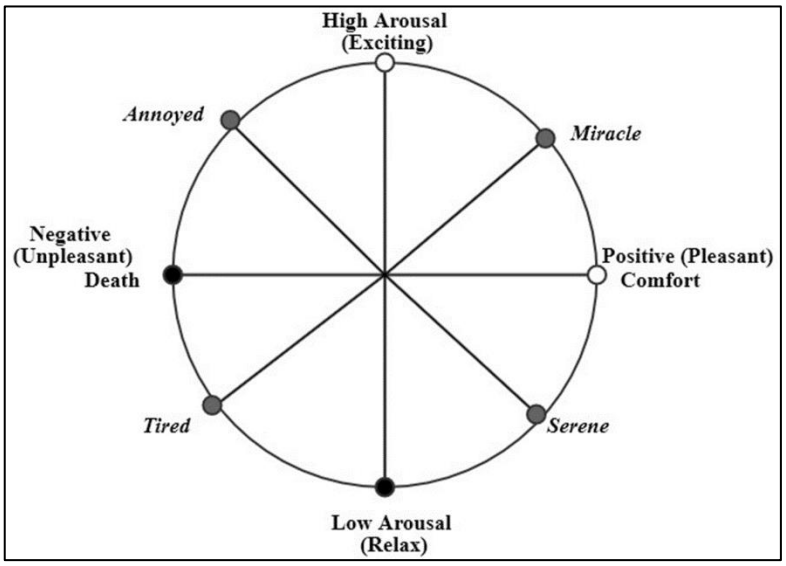


*Figure 10 – Reference model on human emotions (*Nandwani and Verma, 2021*)*

A closer look at the above figure states that emotions are further divided into positive and negative types depending on the joy or sorrow factors. Anger, happy, sad and fear are some of the emotions which can be understood by a human nature. To perform the emotional analysis, there are basically two methods such as Lexicon-based approach and machine learning method. The lexicon-based approach method generally looks out for keywords in a text dealing with the mental state and delivers the results. WordNet-Affect is one of the types used to perform this test based on emotion descriptions. The second approach has various modelling techniques involved in the process namely Naïve Bayes, support vector machine etc. Selection of the right model for the type of given dataset can lead to successful outcomes in a report. On the contrary, testing of the model before real time use is mandatory and can define any misconceptions or errors before the model is ready to deliver its outcome. There is a dependency of data set size and how the information is pre-processed while performing the machine learning methodology (Nandwani and Verma, 2021).

Methodology

- **Step 1: Collection of survey responses**

The first step in this approach began with collecting responses for the survey questionnaire from the participants. Thus, with the help of Google forms which is an online software for creation of surveys by Google which is one of the popular search engines, a survey was prepared based on the background study used in this research. The survey was live between 3rd August 2024 to 13th August 2024 where the respondents were allowed to provide their valuable feedback after the ethics approval obtained by Anglia Ruskin University. The survey was sent to the students studying in the UK through an online generated link. Mostly students who took participation were studying at Anglia Ruskin University, Cambridge. Considering the ethics, the survey was only sent to age above 18 years as the first question confirms the age by selecting **Yes** or **No** for the participant who must be or have completed 18 years of age. It was then asked to inform about the living status in the UK as a citizen or an international student. The major part was confirming about the most used social media platform to calculate the usage consequently along with the time spent on the selected networking site (Hruska and Maresova., 2020). Physical and mental impact was confirmed in the last two questions where an approach of personal feedback was also asked to know any certain information in depth. Participants were given choice to add their personal experienced social media website, or any impact associated with using social media if not available in the default list. Charts of the responses which are defined in the discussion part are derived from this website after completing the survey. However, more features can be only used after getting a paid membership plan and thus it lacks crucial modifications.

- **Step 2: Creating a process flow diagram**

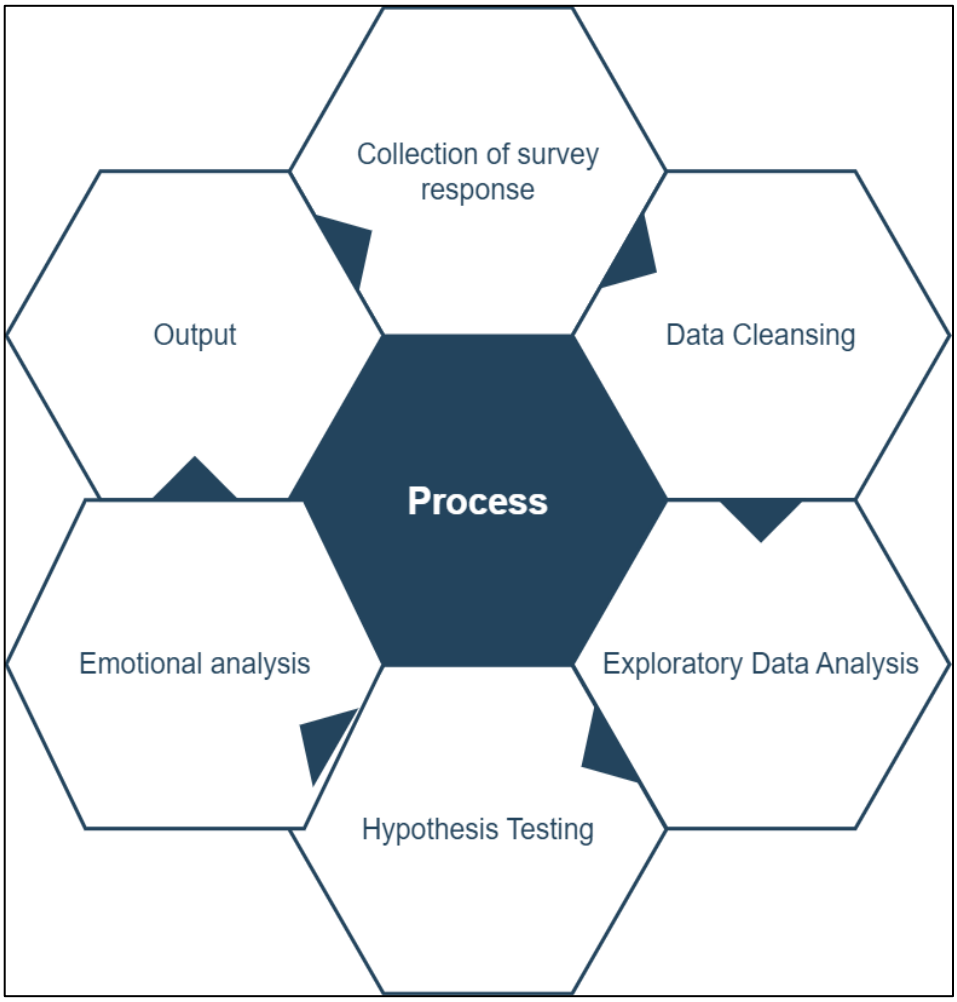


*Figure 11 - Process flow diagram of the methodology (*Draw.io, 2024*)*

The primary focus is on understanding any physical or negative impacts after social media use and hence the purpose of the social network sites and positive effects of using these sites are excluded for the same. The process diagram gives a brief analysis of the important methods involved in the study and below figures are the representation of the survey used in this research. The preparation of the survey questionnaire was kept with minimal technicalities so that the user responses can be recorded in a smaller amount of time. The participants were asked to fill it in their free time so that it does not affect their personal and academic purpose. The personal details were not displayed and are not linked with the user to avoid any disclosure regarding sensitive information. The minimal age was set to 18 years to understand the development period and there was no maximum limit. This ideology gives freedom to students of any age to participate and provide their valuable feedback (Google, 2019).

Social media impacting mental health and well-being.

As a student your response can help other students from choosing a wrong direction. This is an anonymous survey and our aim is to understand real time response analysis.

* Indicates required question

1. Are you aged 18 or above? *
   This is a mandatory question under Ethical risk consideration.

   Mark only one oval.

   - ◯ Yes
   - ◯ No

2. Are you an UK citizen or an international student? *
   This question is required for categorizing data.

   Mark only one oval.

   - ◯ UK
   - ◯ International

3. Do you like spending time on any of these popular social networking sites? *
   Choose the most relevant option from below.

   Mark only one oval.

   - ◯ Instagram
   - ◯ Facebook
   - ◯ Whatsapp
   - ◯ Snapchat
   - ◯ Twitter
   - ◯ Other: _____

Figure 12 – First page of our executed survey form

4. How much time do you spend on the above mentioned platform daily? *

Mark only one oval.

- ◯ Less than an hour
- ◯ 1 - 3 hours
- ◯ 3 - 5 hours
- ◯ Other: _____

5. Any negative impact on mental health from the following options? *

Check all that apply.

- ☐ Depression
- ☐ Stress
- ☐ No such effect
- ☐ Other: _____

6. Any negative impact on physical health from the following options after prolonged use? *

Check all that apply.

- ☐ Body pain
- ☐ Eye strain
- ☐ Headache
- ☐ No such effect
- ☐ Other: _____

7. Would you like to give some personal feedback about social media use?
This may help us to gather some more information on our analysis.

_____

_____

_____

_____

_____

Figure 13 – Second page of our survey

- **Step 3: Data pre-processing**

The back-end data was then transferred in a Microsoft excel workbook and filtrations were made for a better overview of the analysis like filling the blank spaces and alphabetical issues were resolved. All the auto-created chart reports are referred in the result section for producing any theoretical assumption. The report is cleansed in a standard manner as some of the questions have personal opinions which are unable to alter as it can affect the quality of the results. The personal feedback notes are not linked with the outcomes considering data privacy. The cleansing of the information is performed below in the hypothesis testing and emotional analysis sections.

- **Step 4: Implementing a social media impact model**

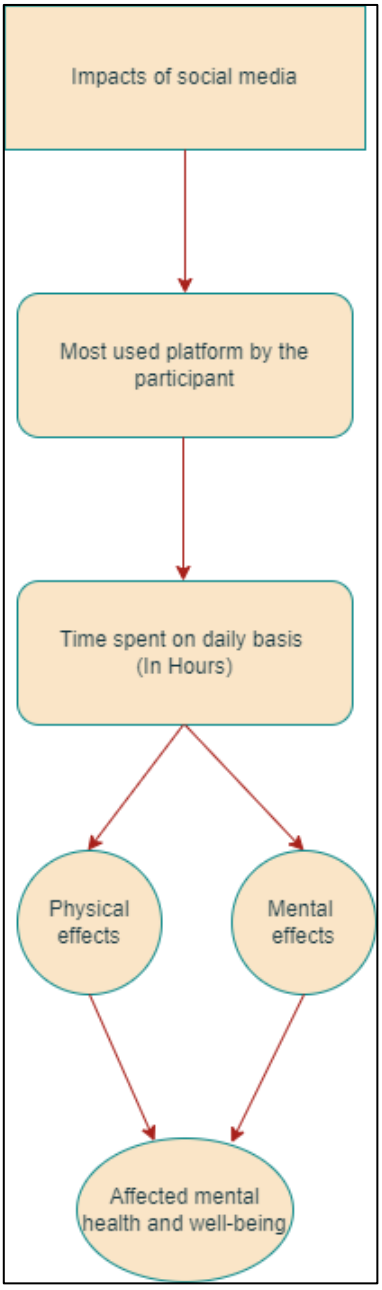


*Figure 14 - Social media impact model (*Draw.io, 2024*)*

The current model is depicted in a flow chart manner where all the predictions are made by analysing the social media impacts having adverse effects on mental health and well-being. The Social media impact model is designed for understanding the theoretical reasons of affected human mental health and well-being. It is linked with the survey questionnaire used for gathering feedback from the students. It tries to collect data about the most relevant website and how much time is invested to have any drawbacks as on physical and mental health. At the final point, the impact is measured on the points selected by the contributors and an effort is made to know the consequences associated with it (Draw.io, 2024).

- **Step 5: Statistical computation**

For e.g., if a user spends around 2 hours on Snapchat by seeing graphical videos or posting any content or watching other's post and goes through anxiety by seeing a specific video, an attempt to draw a conclusion can be yielded further. On an overall basis, the purpose of this attempt is to relate the platform where most of the time by a user is spent and the negative mental and physical effects for future analysis on these assumptions which can aid in knowing the reasons associated with the use for medical purposes. The figures., 15 and 16 depicts a pie chart of the most popular websites indicating the mostly used platform suggested by the respondents. More than 50% of the participants like spending time on Instagram which is a content sharing application where people are allowed to post their own photos or videos with their connections on the application. It is an open platform which makes users to comment or like on each other's content and allows viewing it for other people (media, 2024).

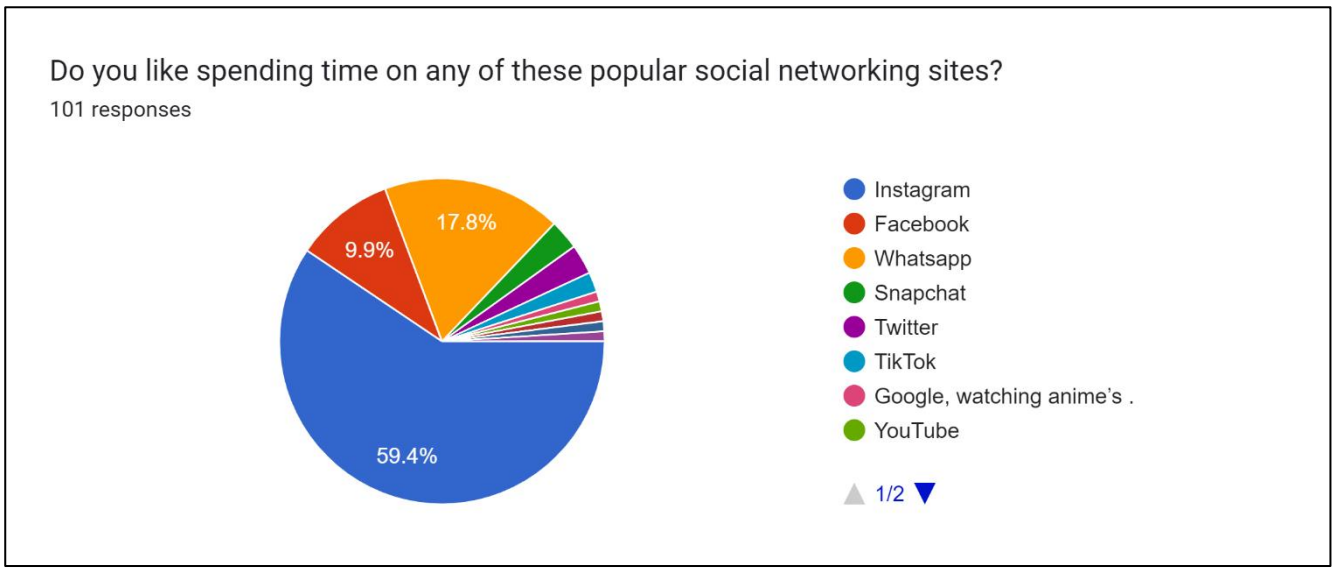


Figure 15 – Most selected social media platform of use in our survey

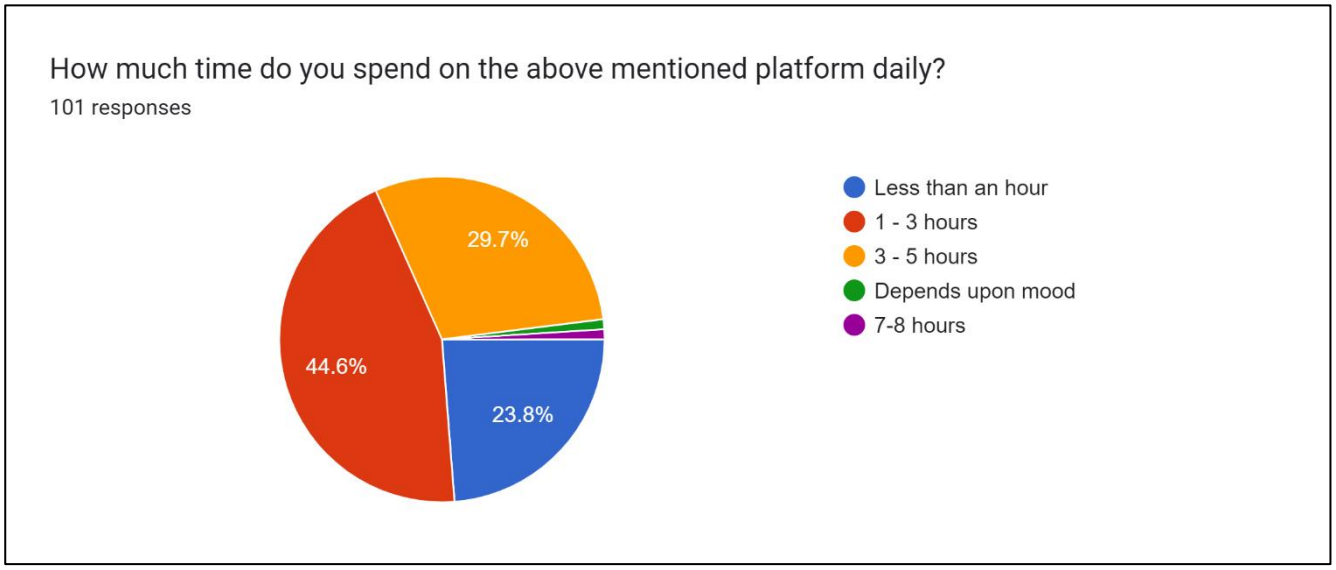


Figure 16 – Time spent on social media daily as per our survey report

using the platform. Secondly, around 46 of the 101 participants like to spend around 1 to 3 hours on daily basis on any of their favourite platform whereas 29 of them are involved in more than 3 hours up to 5 hours and 24 of them likely spend less than an hour. This is illustrated in the figure no 16. The table 4 below states the average spent time on daily basis is 1.9 hours by the respondents. And most of the participants are engaged in Facebook, Instagram and WhatsApp platforms for daily use.

Social media platforms	**Mean time (Hours)**	**Overall average time spent on social media (Hours)**	**Most median time (Hours)**	**Most frequent value observed (Hours)**
Daily communication	2	**1.9**	**2**	**2**
Facebook	2.4			
Google, watching anime's.	2			
Instagram	2.4			
Not interested but not using any networking website	0.5			
Snapchat	1			
TikTok	2			
Twitter	2			
WhatsApp	2.2			
YouTube	3			

Table 4 - Average time spent by the participants on their most used social networking website.

*****NOTE – All the back-end code are practiced on W3Schools and performed on Google Colab for Hypothesis testing and Emotional analysis (**W3Schools, 2019**). *****

▪ Step 6: Hypothesis analysis with Chi-Square Test of Independence

Based on the social media impact model, this study is particularly classified into two aspects such as dealing with the mental and physical health impacts. To start with this testing, there is a need to evaluate the research question which fits perfect to carry with the assessment. The testing approach related to mental effects is linked with depression and stress parameters. On the other hand, body pain, eye strain and headache are the three considerations for physical effects. In the table 5 below, there are two rows for hypothesis testing. The analysis questions are mentioned along with the effects which are taken for evaluation.

Questions for test analysis	**Does social media really have impacts on psychological health?**	**Are physical impacts connected with social media usage?**
Social media use	**Mental Effects**	**Physical Effects**
	Depression	Body pain
	Stress	Eye strain
	-	Headache

Table 5 – Chi-Square test of independence primary questions

The major key points to have a look at before starting with the hypothesis testing include but not limited to developing concepts, preparation of authorized methods, linking the conceptual learning and the functioning methodology and finally, realizing the border limit of the testing part and expectations as a result (Scheel *et al*., 2021). The first step to begin in testing hypothesis should start with defining the null and alternate hypothesis value. The null hypothesis (MH_0) will be used for no such effects from using these social networking platforms. MH_1 which represents that depression and stress are linked with social media usage. The testing is then forwarded to physical impacts where PH_0 is associated with negative physical effects and PH_1 with affirmative chances of caused effects. Chi-Square Test of Independence method is used as the data is categorial into hours spent on social media and any regulated mental or physical effects on the participants. It determines the relation between the two categorial variables used in the report. The frequency table count will also be calculated for interpretation of the results (Chamidah *et al*., 2023). Predicting the significance value thereafter and determining the correct statistical testing methodology is an important step before carrying out the analysis.

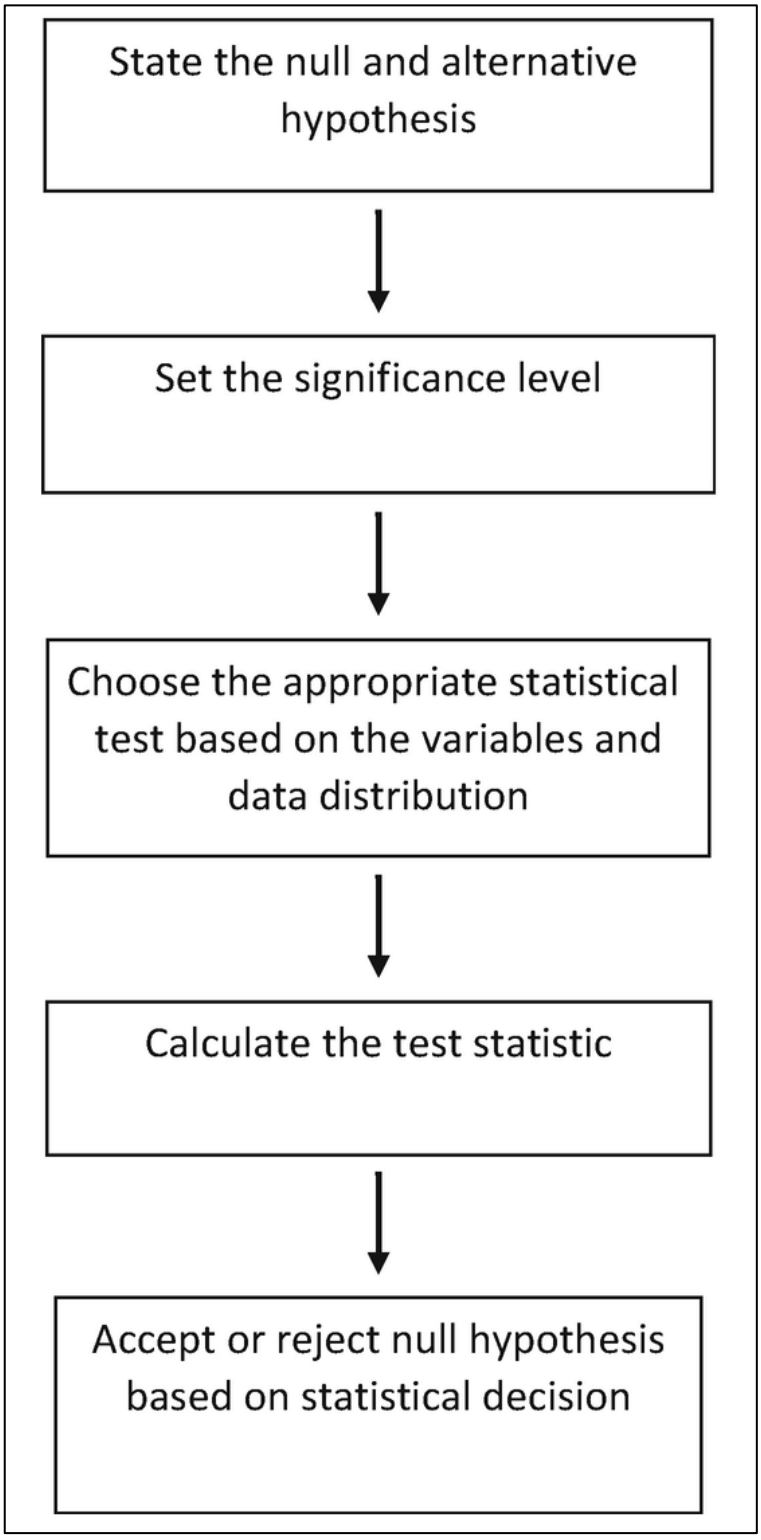


*Figure 17 – Process flow diagram of hypothesis test (*Balasundaram, 2022*)*

The primary response file was divided into two separate files which categorized into mental effects and physical effects. This was made to clearly see two different perspectives for the research questions. Accordingly, the significance value was set to default as 0.05 as the results are predicted to be MH_0 and PH_0 if the value is less than 0.05. If the value is above the significance level, then the association can be true determining the effects on use. Choosing the right platform for performing the backend activity and then providing the calculations to

justify the result is a must note step. Once the decision is made and results are displayed, the comparison can clarify if the decision is to accept or reject the null hypothesis value. The testing is performed on an online platform named ‘Google Colab’ which allows users to perform coding and other technical activities online without downloading the platform reference. The process started by filtering and sorting of the data in Microsoft Excel where only the number of hours and associated mental effects were made available to read (Balasundaram, 2022).

The data file is divided into two aspects naming hypothesis testing for mental effects and hypothesis testing for physical effects (well-being). This helps in reading the data and results differently and with no difficulties because of the process split practice used. Further the data was uploaded on the Pandas platform where actual coding was executed for getting the observations. Followed by loading the file data and creating a contingency table, the final Chi-square test was implied, and the required values were allowed to print which can be observed in the Appendix section at the end. The contingency table shows the frequent values and the summary of the responses from the participants. The Chi-square value which is the differentiated value between the observations and expected frequencies is 29.84 and the degrees of freedom used for calculating the analysis is 20. The final table gives a brief of the frequent values which work with no correlation between the variables. As we can see the p-value is approximately 0.072 > 0.05 (signified value), the relation between the social networking sites and psychological problems is not proved.

The same process was done for interpreting the physical effects where the experimental value was 0.40 which came way higher than 0.05 and shows no direct connectivity with the physical effects. The Degree of freedom value calculated was 28 and the statistical value occurred here is 29.18. A summary of all the process states that the alternative hypothesis is rejected, and the null hypothesis value is to be considered as the test outcome (Google.com, 2019). The coding section is presented in the form of images in the appendix segment with the figure nos. 24, 25, 26 & 27.

- **Step 7: Emotional analysis**

Systems that can detect human states of mind which use artificial intelligence algorithms are spread world-wide in the recent years. The served purpose is to identify decode and reply with a response considering the emotional level of the user. This analysis is based upon the emotive factors of the participants which include stress levels, depression and other aspects related to mental health. Identification of the responses received as feedback and using the knowledge skills can provide a better output criterion. This analysis is carried with the help of machine learning technologies to define the emotion related to the context. Emotions can be divided into positive and negative level where for e.g., being happy about a think expresses positivity

and otherwise being unhappy because of some thing can be derived into negative outcome of emotion. AI models use below types of techniques for emotion detection which include:

Text analysis: This technique is used to understand the level of emotions with the help of transcribed or communicated languages.

Visual analysis: It can be analysed with the help of photos, videos and real-time facial expressions.

Audio analysis: This contains data through voice inputs (Plaza *et al*., 2022).

The survey in this approach has textual responses received from the applicants and the responses are not classified directly with emotions. Hence, the filtration is done in the file named 'Emotional analysis.csv' where the feedback is tested with adding a manual column describing emotions. This can be cleared with the help of the emotional testing analysis which is carried out to understand the emotive level of the participants for social networking use. The evaluation will include two attempts, where one to detect the 'happy' and 'unhappy' measurements from mental health effects and other from the well-being impacts. The results of this analysis help people to draw an outline of the given data very quickly which saves time to read and understand the given context of information. Also, it assists in restricting the adverse effects of some displays on the social media platforms. Using these websites which connect you socially can affect one's mental health of the people of this generation as it serves as an interest to many individuals. People using these websites are unknown of the harmful effects that can happen as they are unaware of it because of lack of awareness programs on these websites. Detecting the textual emotions is one of the challenges as sometimes even human beings find it difficult to understand the sentiments too. The method used in this testing are probably referred from Benrouba and Boudour *et al*. (2023), where the objective was to extract the content which can prove dangerous to a person's emotion. The target website was Twitter API where the content was chosen with the help of machine learning technology. The five major emotions classified for this research were happiness, sadness, rage, terror and shock. The use of Artificial Intelligence in human life is equal in terms of gains and losses. If an unsecured system is used there is always a concern of misplacing sensitive information. Depending on these machineries can save time and efforts done by a human-being and can also bypass the human-error terminology.

- **Step 8: Development of the machine learning model**:

The first step is to filter all the blank and unavailable responses so that the data is filled with opinions. Then the column named 'Would you like to give some personal feedback about social media use?' is updated to 'Feedback' for a short name and a manual column as mentioned above is added with the name 'Emotions'. There are a total of 37 responses which will now be

given a specific emotion based on the logical level. The read responses are further distributed into three emotive sections named happy, sad and doubt and below table 6 explains how the emotions will work throughout the evaluation process. The file is then manually uploaded to the Google Colab online platform where the backend server used is python to run the tests. The libraries important for the process are 'pandas', 'NumPy', 'Scikit learn', 'NLKT, 'Matplotlib' and 'seaborn' are made sure to be installed before the analysis. 'Pandas' and 'NumPy' are used to determine the alteration of information available. 'Scikit learn' will be useful to prepare any models related to machine learning. To understand the human text data, 'NLKT' which is the Natural Language Toolkit is made available to the system. Matplotlib is a library which will assist in designing any graphical interface diagrams or charts if necessary. To define the complexity of the graphs, there is a need to utilize the Seaborn library (Janiesch *et al*., 2021).

Emotions	Explanation
Happy	This emotion is used when the user is happy using any social media platform.
Sad	This emotion is used when the user is not happy after using any social media platform.
Doubt	A combined response and can be used to find what lays the question on using these networking platforms

Table 6 – Human emotions selected for the analysis with their brief explanation.

- Uploading the .csv file in Google Colab:

To simply upload the file, we must click on the Folder option and then select the menu list on the right-hand corner and click upload which will ask to load the file and once clicked OK the file will be added automatically and is visible in the folders section. See the figure below for explanation where the file upload is completed successfully. Once the file has been uploaded, we can start the back-end coding where firstly we will install the required libraries and then the libraries will be imported in the code along with loading the file data. The phases are described in the figures below. The final print command is used to display some of the rows and columns to check if the file is loaded correctly.

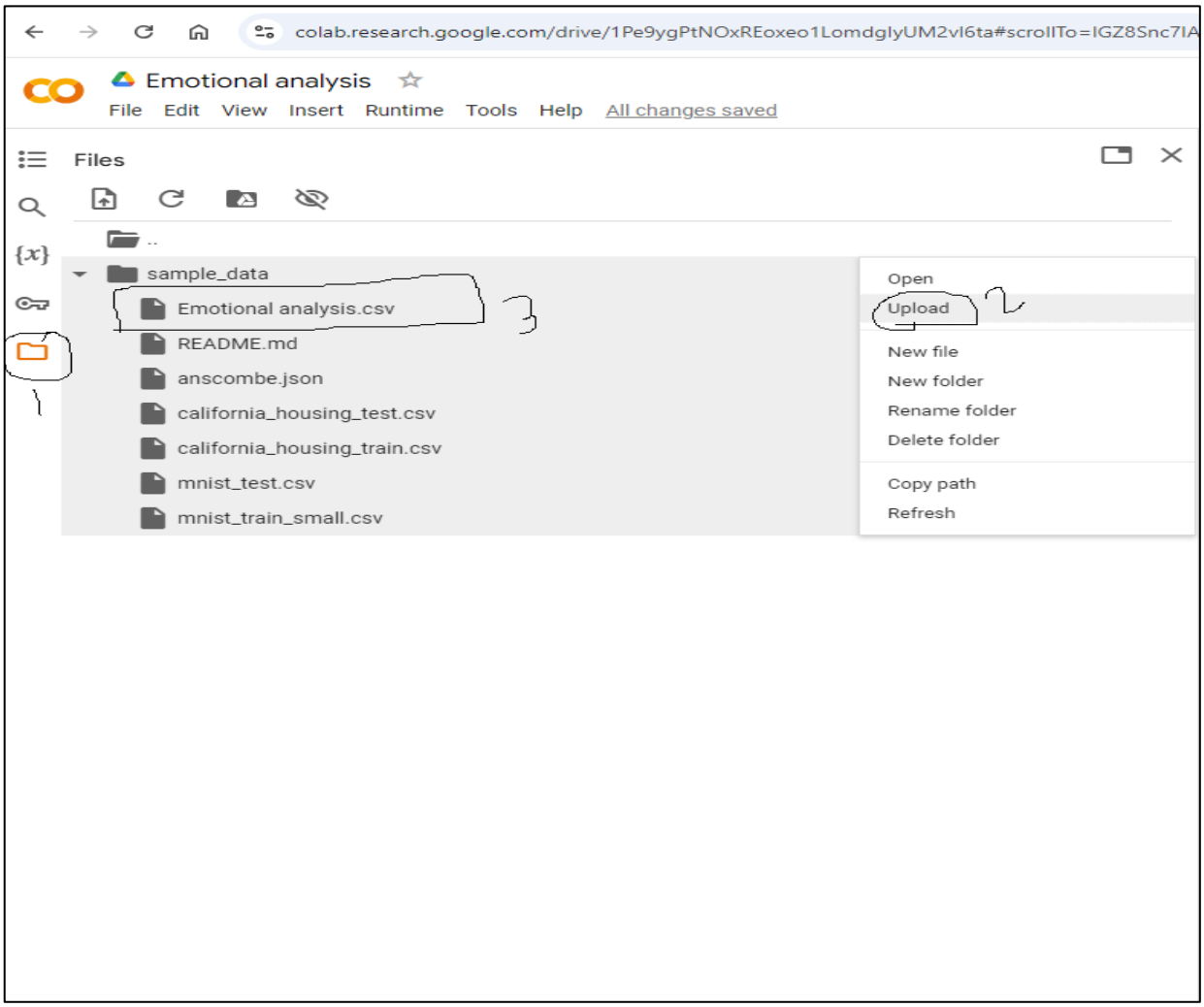


Figure 18 - Loading the .csv file in the python platform on Google Colab (Google.com, 2019)

As described above, the process starts with installation of the libraries to run in the program. The necessary libraries are then imported into the system as shown in the figure no. 22. The manually uploaded .csv file will now be executed in the code and a general print command for making sure that the correct file is uploaded for the process. Hence, we will print some content of the file.

The data needs to be cleaned before performing the testing analysis and hence the required NLTK data Stopwords, WordNet and Punkt are downloaded to ensure that the redundant words are skipped, database is loaded which will be helpful for the machine to understand general English methodologies and the sentence list is available. After that the tools are initialized in the system which can now run as per the need of the user. Data cleansing is a primary step thereafter to avoid any incorrect words or letters. The text is converted to lower syntax and spacing options are organized in the code.

The data is then divided into two columns as we have made it available in the .csv response file and the text is converted to numeric format for matrix analysis and modelling. Naive Bayes classifier model is one of the machine learning models which can perform algorithms on texts.

This model will assist in generating probabilities based on the textual content of the questionnaire responses. After that a prediction is made on the model for testing the model before use where accuracy of the model is calculated along with a predictable short display of the file contents. If you see the figure 29 in the appendix, the accuracy and output is printed for reference.

Tuning the hyperparameter is an important step which cannot be missed while testing a machine learning model. This approach is used for checking if the model can analyse all the variables which will be used for demonstrating a model (Reddy *et al*., 2023). The grid search method is used because of its availability in the python database, and it is beneficial as this technique will try to fit through all the combinations of a matrix so that the best fit result is provided at the end (Amazon Web Services, Inc., n.d.). The model is further checked on the performance parameter by cross validation method as referred in the figure 30, to check if it can deliver quality results. The code will not run a confusion matrix to check if there are any possible errors and the matrix as an output is demonstrated in the figure 31. Finally, once the model is ready after testing, we can try to execute real-time results (Google.com, 2019). All the presentation of the back-end code executed can be found in the appendix part at the end.

Results & Discussion

The survey based on using social media website whether impacting mental and well-being health was configured to know the negative outcomes on mental health and well-being of humans. The questions were answered by the age group which is more than 18 years old and furtherly divided into home or international students. The age criteria were defined and made available for only over 18 years keeping Ethical considerations in proportion. Major responses were from the international students as the main purpose was to check if there is any negative outcome from using social media because of the distance from their family relations and friends. Instagram was found as the most used platform among all the listed and unlisted social media websites. And around 45% of the total contributors were using any of the most used website between 1 to 3 hours daily.

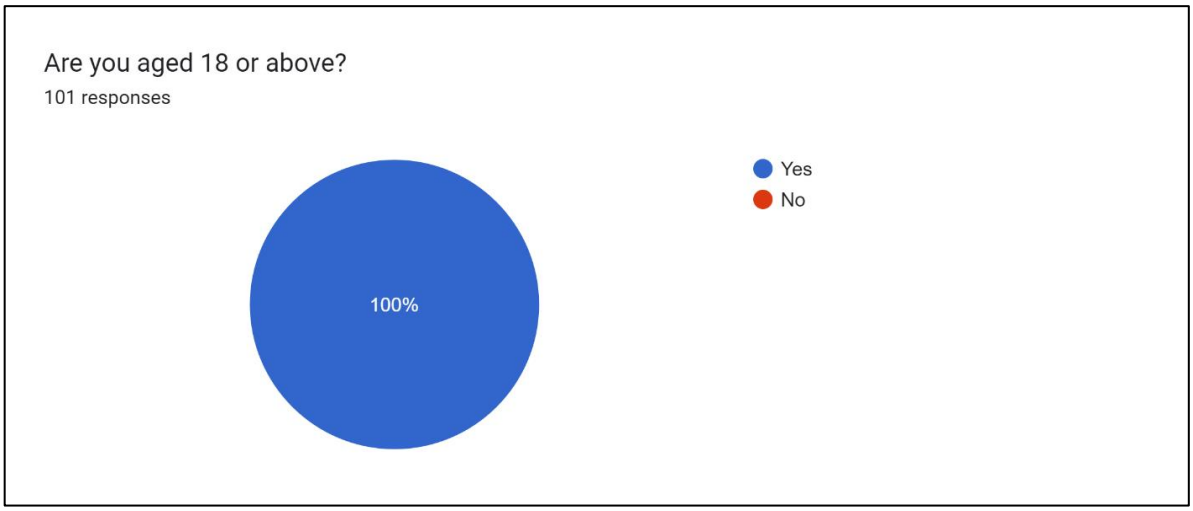


Figure 19 - Response rate on the age group question of our survey

The outcome for negative effects from using social media was majorly reported as with no effect where 73 of the 101 members selected it. However, the part that 25% felt stress and around 6% went through depression as mentioned by them. The outcomes are always not effective due to different body nature of everyone. The results for physical effects showed some variations as compared to the mental impacts in which only 48 reported having no impact physically after sustained use. More than 50 participants had body pain, eye stress and headache associated with use. In fact, it was a draw situation where results were still imbalanced about the negative body effects due to these websites use. The chart representation of mental and physical effects is depicted in the figures below where all the outcomes are stated graphically.

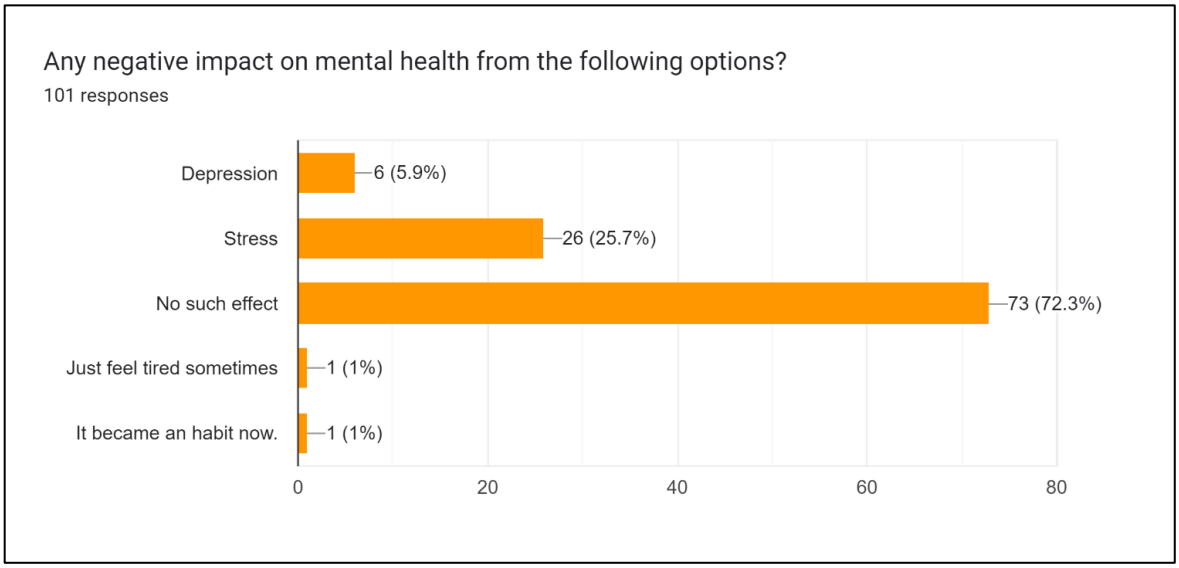


Figure 20 - Negative impact graph result on mental health due to social media use as per the survey

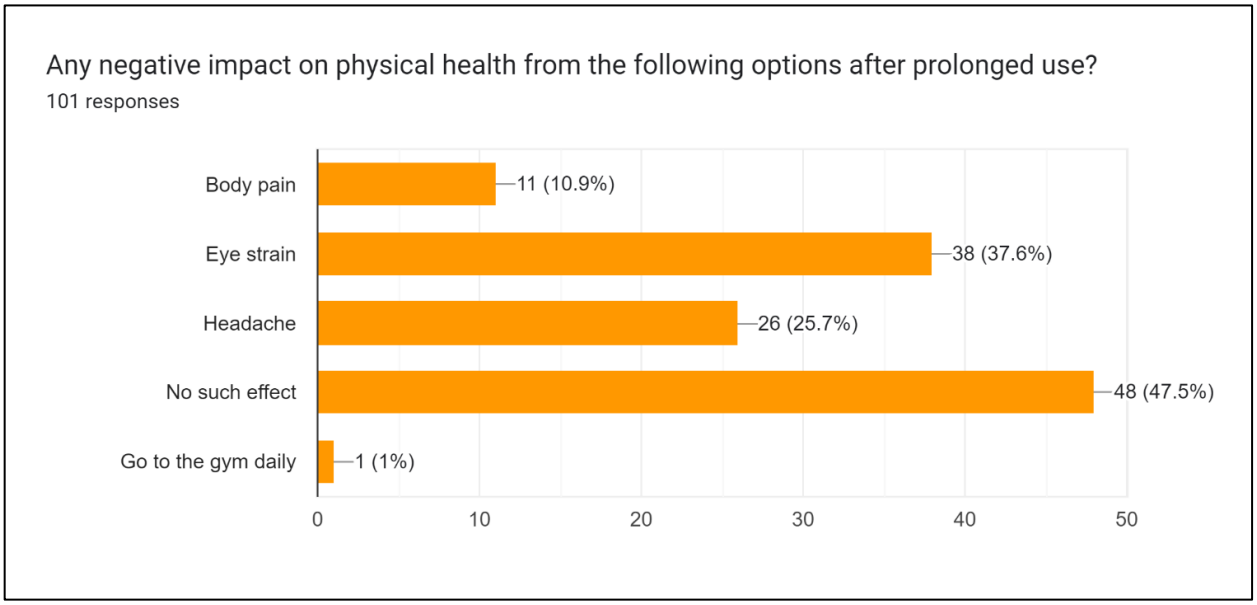


Figure 21 - Negative impact graph on physical health due to social media use as per the survey

As demonstrated by Kircaburun *et al*., (2020), the statistical report almost presented 92% of the individuals using the WhatsApp platform and the survey made in this report has claimed Instagram as the most used platform in social networking on daily basis. Whereas Instagram was the second highest with 79% of use. The increase in data can be with the increase in number of users and young generations as well. If we trend 2020 and 2022, more application features and updates have been noticed and thus can be the reason of variations. Previous studies have also stated anxiety and depression as common mental impacts compared with usage of social media and they are still constant based on current study data. The daily use time is still a question for many researchers as there are different results when performing analysis and the time factor even though remains constant, however the fluctuation in spent time varies every time. Thus, it is still difficult to draw a conclusion on how much time spent

can affect one's mental health in terms of social media in longitudinal and short-term studies. (Plackett *et al*., 2023).

The predictions and the responses were then treated hypothetically to find if there is any direct correlation between the social media usage and human health. The hypothesis showed that the predictable connection was unjustified. There were two predictable parameters in which one was related to mental health and the other physically. The responses received from the applicants stated that physical health has some link with this use but unfortunately the hypothesis testing analysis was unable to prove the exact findings. In a research report done for checking if there is a link with academic scores of students with social media, there was a little proof associated with it which downs a student's academic performance as it is known that some of the social platforms are used by students to attend meeting, share study notes, etc. The hypothesis testing involved chi-square testing method which was unable to collect evidence about the fact that links social media with academic records. The hours spent on these platforms and sleeping hours calculations showed no connectivity whilst the academic performance on gender as compared to the significant value (p = 0.05) was found to be 0.059 which is still more than the considerable value (Behera *et al*., 2023).

Hypothesis testing	Defined value	Significance value
Mental effects	0.07	0.05
Physical effects	0.40	0.05

Table 7 – Hypothesis testing analysis and result table

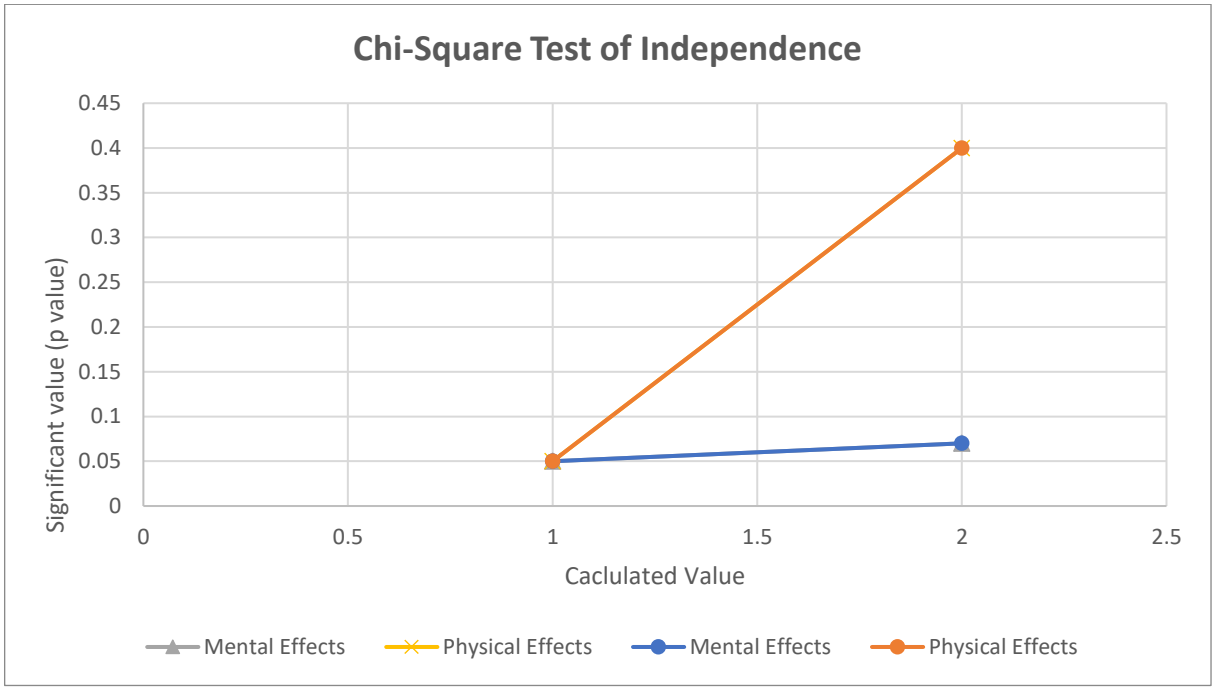


Figure 22 - Graphical representation of Chi-Square Test of Independence outcome as per our survey

As described in the figure above, the p value on the Y-axis shows the significance of theory whereas on the X-axis the two plot points are of the calculation methodology used in the procedure. The graph clearly shows that the points are not in a straight line with the p-value which makes it difficult to state a relation of the measurements.

The Emotional analysis was built on the received personal feedback from the participants. The data was converted into readable files and manually updates were performed to correspond with the machine learning requirements. The emotional prediction parameters determined were ‘happy’, ‘sad’ & ‘doubt’. The ‘doubt’ emotion was used for neutral opinion nor ‘happy’ nor ‘sad’. Based on the results, the emotions were then adjusted in the table below where the overview of the total percentage of emotions is displayed.

Emotions	Count of Feedback
Doubt	15
Happy	11
Sad	11

Table 8 – Response count on emotions.

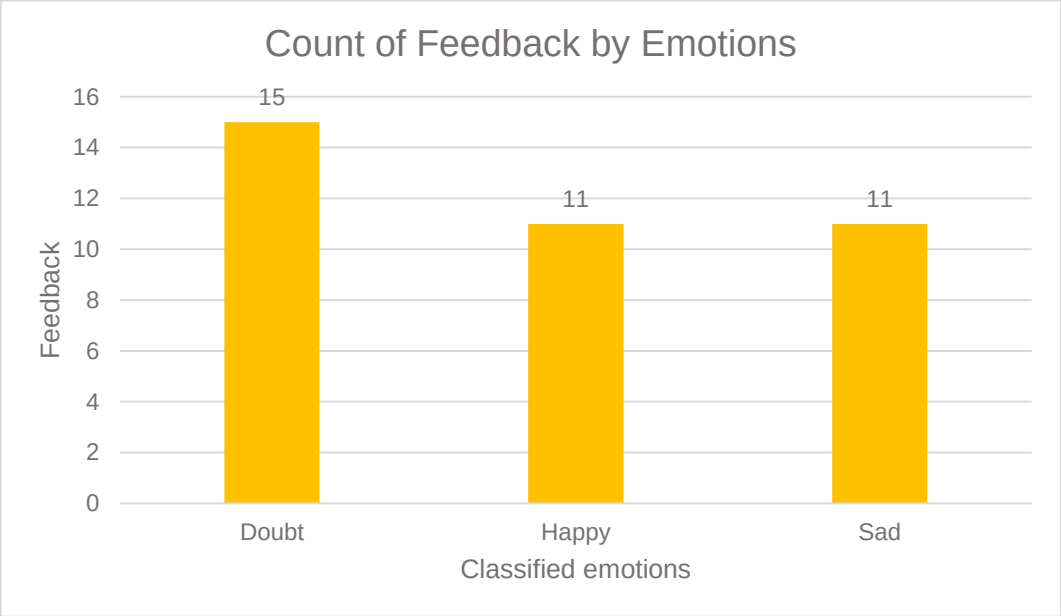


Figure 23 - Chart of emotional analysis test done on our survey feedback

A total of 37 personal opinions were collected out of which 15 of the contestants were neutral with their reviews. 11 opinions were directed towards happiness and 11 exhibited sadness with their response on social media. The outline of these findings is still unable to provide the best fit results and the social networking platforms and human relation remains unproved. The results can fluctuate as personal opinions are derived from experiences and there is a need for stable and long - term study in this research which can be more accurate than the short-time survey. Extraction of available attributes and pre-dispensation of the data can have a

crucial advantage on the execution and the evolution can be of more advantage for these testing approaches (Nandwani and Verma., 2021). The emotional analysis is a wide area of study which can be densely linked with sentimental testing. Various fields like marketing, artificial intelligence, psychological departments are in lined with these effective applications. Sharing content and emotions on social websites like Facebook, Instagram and Google, etc is a common practice in day-to-day life of humans (Sharma *et al*., 2021).

Limitations

This research is not focusing on the purpose of the social media use as the primary objective is to understand the negative impacts of using these social networking applications. There is also a smaller amount of light on the positive effects of using these platforms. More questions can be added in the list after reviewing the recent scenarios on social media. Different methods can also be executed like long term studies on this research to know the late outcomes of the analysis. The social media impact model is designed thoroughly to understand the negative impacts; however, the model can be designed in more detail with reference to the findings and the results. Different convenient methods and more data extraction is a lookup point in terms of hypothesis testing. Text analysis done based on the feedback can be more in detail and more emotions can also be listed to make the analysis have more positive outcome. The Google forms survey has some drawbacks like it always requires an active internet connectivity to prepare the forms, generate online and record responses. The system does not follow conditional logic as there is no availability of showing or hiding opinions which can control the question formats. There are also low options for designing in Google forms like for accessing premium features there is a need for payment and is not free ((Jotform, 2024). Chi-Square test is limited to size of the sample as it is difficult to handle large amount of dataset with it. Relationships of variable remains undefined with one another (Turhan, 2020). Certain drawbacks of emotional analysis can include low accuracy compared to other techniques. Also, the evolvement in emotional and sentiment tests can also restrict the technical aspects of these tools. People express their emotions in various ways, and which is why it is complicated to achieve good results using one model for all the functions (Pietikäinen and Silven, 2022).

Conclusions

The study was conducted to demonstrate the mental and well-being of humans in associated with using social media. As the survey was made live and responses were collected from the participants, many major findings were noted as a part of the result of the questionnaire. Majority of the individual who participated in this approach were international students living in the U.K. for studies and developing their career away from their home countries. The feedback received were very helpful in understanding the factors involved in social media usage. Around 31% of the applicants felt they were going through some depression and stress after using one of their daily used social media platform. 50% of the candidates stated having some sort of body pain, stress in eyes and headache which were the points covered in the negative impacts on physical health. Instagram was ranked first which is most used by many of the participants along with WhatsApp as the second highest recorded platform of use. The question on time spent on these platforms also defined that more consumption of time can be somewhere linked with social media. The least time calculated was reported to as 0 to 1 hour which in terms was not selected by major participants. The hypothetic questions chosen for testing were directly advanced from the title of this research. The statistical decision was not in the favour of the report as both the mental and physical analysis were not directly related with their significance level of value and hence there was little evidence proving the same and hence null hypothesis value was at acceptance. The alternative hypothesis value for measuring mental impacts from using social media was 0.07 and the other for physical effects was around 0.40 as a part of statistical data. Machine learning approach was utilized for checking emotional level attached with the feedback responses. As the survey received its answers in text formats, the textual procedure was the best fit for calculating the emotions. Naïve Bayers classifier model was attempted to make predictions based on the three selected emotions such as the state of being happy, sad or neutral with answers. The accuracy level was 1.0 which states that there are less chances of positive and negative errors. The random matrix was properly designed and gave judgemental outcome when the model was tested with confusion matrix. When the emotions were classified, it was revealed that 15 people were in doubt whether social media is good for use or no. On the other hand, 11 people were happy and 11 people reported sadness after using these platforms.

These findings can help build a strong model and study to recognize the impacts of this topic which can prove beneficial in terms of medical and treatments. The level of addiction can be reduced when knowing the points that affect health psychologically and physically of a person. These records can help medical practitioners to achieve positive treatment results for their patient’s problems. Different techniques were used to prove the relation between the usage of

social platforms with mental and well-being in human-beings, however the collected information and results are less likely to draw an assumption which can generate a correlation. However, the review report helped in knowing that people have problems when they use such social media websites. These issues are ignored as the platforms serves to focus on the positive sides and there is less information available on the drawbacks of the same. Social media is a part of human life and the fact that disconnecting it will not be of assistance as it can impact in more depressive symptoms, however limiting the use is one of the try to get positive results. Engaging in physical activities, developing good mental habits by reading books, trusting the information only when appropriate proof is present, generating a schedule of the day and many more can be of more use. Doing research on any news before forwarding it, checking if the links which are forwarded to us are not critical which can steal personal information or can affect financial status and ensuring to capture all the advantages that are offered from any social network platform keeping the drawbacks in mind can lead to a good social life for an individual. The objective of this research was to determine whether social media affects human health. Throughout the study, many responses were explained which showed that social media is linked to mental and physical well-being of an individual. Personal experiences on real-time basis are recorded as the survey outcomes and this can be kept at focus when planning any future studies on social media usage. Other than this, emotion detection model was successfully built in the online system which needs some work with technical aspects for using on real-time basis. The results and testing of the model provided efficient responses as an outcome.

Future work

Studies like these which are focusing on impacts related to social network should be used widely on every age group using these platforms whether they have a positive, negative or general mindset about these sites to understand the opinions and to dig deeper into this research. The studies should be conducted on long-term basis to analyse the periodical outcome. There are several online websites for online form creations which can be explored to gain any benefits which are preinstalled in the application. All the work done on this design is done online and it shows advantages of using internet for the right purpose which can be used as an example of positive use of the available resources. In future, the researchers can modify the set of questions in the questionnaire depending on the consequences leading to mental and physical stress in humans. Different techniques can be experimented for hypothesis testing analysis and the research questions can also vary in upcoming time. Emotional analysis can also be done with other testing models to generate some positive outcomes for the

evaluation. The future of social media is bright, and it can also affect the person's overall mindset if no precautions are taken which can help promote the advantage of using the social resources. There is also a requirement of risk assessments which can calculate the risks and their measures can be drawn which can benefit safety from wrong use of these social media sites. Detection of emotions with the help of artificial intelligence can be made more sufficient by adding some more research and features. Additionally, the addiction of the social media platforms in harmful conduct should be lowered and the initial spotlight should be on digging more data on the more dangerous negative impacts and threats from such effects.

References

1. Amazon Web Services, Inc. (n.d.). *What is Hyperparameter Tuning? - Hyperparameter Tuning Methods Explained - AWS*. [online] Available at: https://aws.amazon.com/what-is/hyperparameter-tuning/. (Last accessed on: 31st August, 2024).

2. Aichner, T., Grünfelder, M., Maurer, O. and Jegeni, D. (2021). Twenty-five years of social media: a review of social media applications and definitions from 1994 to 2019. *Cyberpsychology, behavior, and social networking*, 24(4), pp.215-222.

3. Al-Rahmi, A.M., Shamsuddin, A., Alturki, U., Aldraiweesh, A., Yusof, F.M., Al-Rahmi, W.M. and Aljeraiwi, A.A. (2021). The influence of information system success and technology acceptance model on social media factors in education. *Sustainability*, 13(14), p.7770.

4. Appel, G., Grewal, L., Hadi, R. and Stephen, A.T. (2020). The future of social media in marketing*. Journal of the Academy of Marketing science,* 48(1), pp.79-95.

5. Balasundaram, L. (2022). Hypothesis Testing. In *Introduction to Basics of Pharmacology and Toxicology: Volume 3: Experimental Pharmacology: Research Methodology and Biostatistics* (pp. 865-875). Singapore: Springer Nature Singapore.

6. Behera, D.K., Gartia, D.R., Sahoo, D.R., Munda, S. and Sahu, P. (2023). A statistical analysis of the effect of social networking sites on students' academic performance in the western Odisha universities. *Journal of Mathematical Problems, Equations and Statistics,* 4(1), pp.05-09.

7. Benrouba, F. and Boudour, R. (2023). Emotional sentiment analysis of social media content for mental health safety. *Social Network Analysis and Mining,* 13(1), p.17.

8. Beyens, I., Pouwels, J.L., van Driel, I.I., Keijsers, L. and Valkenburg, P.M. (2020). The effect of social media on well-being differs from adolescent to adolescent. *Scientific reports,* 10(1), p.10763.

9. Braghieri, L., Levy, R.E. and Makarin, A. (2022). Social media and mental health. *American Economic Review*, 112(11), pp.3660-3693.

10. Buchanan, T. (2020). Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *Plos one*, 15(10), p.e0239666.

11. Chamidah, N., Siregar, N.R.A.A., Al Farizi, M.F., Pratama, B.S., Faiza, A. and Fibryan, M.H. (2023). Categorical Analysis To Perceptions of Government Policy in Electricity Fuel

Management As Alternative To Substitute Oil Fuel Using Chi-Square Test. *BAREKENG: Jurnal Ilmu Matematika dan Terapan,* 17(3), pp.1287-1300.

12. Coyne, S.M., Rogers, A.A., Zurcher, J.D., Stockdale, L. and Booth, M. (2020). Does time spent using social media impact mental health?: An eight year longitudinal study. *Computers in human behavior*, 104, p.106160.

13. Firth, J., Torous, J., López-Gil, J.F., Linardon, J., Milton, A., Lambert, J., Smith, L., Jarić, I., Fabian, H., Vancampfort, D. and Onyeaka, H. (2024). From “online brains” to “online lives”: understanding the individualized impacts of Internet use across psychological, cognitive and social dimensions. *World Psychiatry*, 23(2), pp.176-190.

14. Ghani, N.A., Hamid, S., Hashem, I.A.T. and Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human behavior*, 101, pp.417-428.

15. Google. (2019). *Google Forms: Free Online Surveys for Personal Use.* [online] Google.co.uk. Available at: https://www.google.co.uk/forms/about/. (Last accessed on: 30th August, 2024).

16. Google.com. (2019). *Google Colab.* [online] Available at: https://colab.research.google.com/drive/1vqAa-tB5biUZ1I99bMN9pe7ck0Si9kw6?authuser=0 (Last accessed on: 30th August, 2024).

17. Hou, Y., Xiong, D., Jiang, T., Song, L. and Wang, Q. (2019). Social media addiction: Its impact, mediation, and intervention. *Cyberpsychology: Journal of psychosocial research on cyberspace,* 13(1).

18. Hruska, J. and Maresova, P. (2020). Use of social media platforms among adults in the United States—behavior on social media. *Societies*, 10(1), p.27.

19. Huang, P.C., Latner, J.D., O’Brien, K.S., Chang, Y.L., Hung, C.H., Chen, J.S., Lee, K.H. and Lin, C.Y. (2023). Associations between social media addiction, psychological distress, and food addiction among Taiwanese university students. *Journal of Eating Disorders,* 11(1), p.43.

20. Janiesch, C., Zschech, P. and Heinrich, K. (2021). Machine learning and deep learning. Electronic Markets, 31(3), pp.685-695.

21. Jotform, (2024). *The pros and cons of Google Forms*. [online] The Jotform Blog. Available at: https://www.jotform.com/google-forms/pros-and-cons-of-google-forms/. (Last accessed on: 10th August, 2024).

22. Karim, F., Oyewande, A.A., Abdalla, L.F., Ehsanullah, R.C. and Khan, S. (2020). Social media use and its connection to mental health: a systematic review. *Cureus*, 12(6).

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

24. Kircaburun, K., Alhabash, S., Tosuntaş, Ş.B. and Griffiths, M.D. (2020). Uses and gratifications of problematic social media use among university students: A simultaneous examination of the Big Five of personality traits, social media platforms, and social media use motives. *International journal of mental health and addiction*, 18, pp.525-547.

25. Lavados-Romo, P., Andrade-Mayorga, O., Morales, G., Muñoz, S. and Balboa-Castillo, T. (2023). Association of screen time and physical activity with health-related quality of life in college students. *Journal of American College Health,* 71(5), pp.1504-1509.

26. Maharani, A.C. (2021). The influence of excessive use of social media*. Indonesian Journal of Social Sciences*, 13(1), p.11.

27. Manca, S. (2020). Snapping, pinning, liking or texting: Investigating social media in higher education beyond Facebook. *The Internet and Higher Education*, 44, p.100707.

28. Marttila, E., Koivula, A. and Räsänen, P. (2021). Does excessive social media use decrease subjective well-being? A longitudinal analysis of the relationship between problematic use, loneliness and life satisfaction. *Telematics and Informatics*, 59, p.101556.

29. Masciantonio, A., Bourguignon, D., Bouchat, P., Balty, M. and Rimé, B. (2021). Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. *PloS one,* 16(3), p.e0248384.

30. Masciantonio, A., Bourguignon, D., Bouchat, P., Balty, M. and Rimé, B. (2021). Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. *PloS one*, 16(3), p.e0248384.

31. media, S. (2024). *Social media impacting mental health and well-being.* [online] Google Docs. Available at: https://docs.google.com/forms/d/e/1FAIpQLSftroIEzFcuGeS9zArahopQaBuw-j0yyp3L0Y9MNqp0Y93twA/closedform (Last accessed on: 25th August, 2024).

32. Nandwani, P. and Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social network analysis and mining*, 11(1), p.81.

33. Naslund, J.A., Bondre, A., Torous, J. and Aschbrenner, K.A. (2020). Social media and mental health: benefits, risks, and opportunities for research and practice. *Journal of technology in behavioral science*, 5, pp.245-257.

34. Nurudeen, M., Abdul-Samad, S., Owusu-Oware, E., Koi-Akrofi, G.Y. and Tanye, H.A. (2023). Measuring the effect of social media on student academic performance using a social media influence factor model. *Education and Information Technologies*, 28(1), pp.1165-1188.

35. O'Day, E.B. and Heimberg, R.G. (2021). Social media use, social anxiety, and loneliness: A systematic review. *Computers in Human Behavior Reports*, 3, p.100070.

36. Olorunsogo, T.O., Balogun, O.D., Ayo-Farai, O., Ogundairo, O., Maduka, C.P., Okongwu, C.C. and Onwumere, C. (2024). Mental health and social media in the US: A review: Investigating the potential links between online platforms and mental well-being among different age groups. *World Journal of Advanced Research and Reviews*, 21(1), pp.321-334.

37. Paiman, N. and Fauzi, M.A. (2023). Exploring determinants of social media addiction in higher education through the integrated lenses of technology acceptance model (TAM) and usage habit. *Journal of Applied Research in Higher Education*, 16(3), pp.776-803.

38. Pietikäinen, M. and Silven, O. (2022). Challenges of Artificial Intelligence--From Machine Learning and Computer Vision to Emotional Intelligence. *arXiv preprint arXiv:2201.01466*.

39. Plackett, R., Sheringham, J. and Dykxhoorn, J. (2023). Correction: The Longitudinal Impact of Social Media Use on UK Adolescents' Mental Health: Longitudinal Observational Study. *Journal of Medical Internet Research*, 25, p.e47678.

40. Płaza, M., Trusz, S., Kęczkowska, J., Boksa, E., Sadowski, S. and Koruba, Z. (2022). Machine learning algorithms for detection and classifications of emotions in contact center applications. *Sensors*, 22(14), p.5311.

41. Purnama, Y. and Asdlori, A. (2023). The Role of Social Media in Students' Social Perception and Interaction: Implications for Learning and Education. *Technology and Society Perspectives (TACIT)*, 1(2), pp.45-55.

42. Rao, B.N. and Kalyani, V. (2022). A study on positive and negative effects of social media on society. *Journal of Science & Technology (JST),* 7(10), pp.46-54

43. Raza, S.A., Qazi, W., Umer, B. and Khan, K.A. (2020). Influence of social networking sites on life satisfaction among university students: a mediating role of social benefit and social overload. *Health Education*, 120(2), pp.141-164.

44. Reddy, E.M.K., Gurrala, A., Hasitha, V.B. and Kumar, K.V.R. (2022). Introduction to Naive Bayes and a review on its subtypes with applications*. Bayesian reasoning and gaussian processes for machine learning applications,* pp.1-14.

45. Scheel, A.M., Tiokhin, L., Isager, P.M. and Lakens, D. (2021). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science*, 16(4), pp.744-755.

46. Sharma, T., Diwakar, M., Singh, P., Lamba, S., Kumar, P. and Joshi, K. (2021). Emotion Analysis for predicting the emotion labels using Machine Learning approaches. In *2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)* (pp. 1-6). IEEE.

47. Sivakumar, A., Jayasingh, S. and Shaik, S. (2023). Social media influence on students' knowledge sharing and learning: *An empirical study. Education Sciences*, 13(7), p.745.

48. Smith, D., Leonis, T. and Anandavalli, S. (2021). Belonging and loneliness in cyberspace: impacts of social media on adolescents' well-being. *Australian Journal of Psychology,* 73(1), pp.12-23.

49. Sujarwoto, Saputri, R.A.M. and Yumarni, T. (2023). Social media addiction and mental health among university students during the COVID-19 pandemic in Indonesia. *International journal of mental health and addiction*, 21(1), pp.96-110.

50. Thai, H., Davis, C.G., Mahboob, W., Perry, S., Adams, A. and Goldfield, G.S. (2023). Reducing social media use improves appearance and weight esteem in youth with emotional distress. *Psychology of Popular Media*.

51. Turhan, N.S. (2020). Karl Pearson's Chi-Square Tests. *Educational Research and Reviews*, 16(9), pp.575-580.

52. Tsao, S.F., Chen, H., Tisseverasinghe, T., Yang, Y., Li, L. and Butt, Z.A. (2021). What social media told us in the time of COVID-19: a scoping review. *The Lancet Digital Health*, 3(3), pp.e175-e194.

53. Webster, D., Dunne, L. and Hunter, R. (2021). Association between social networks and subjective well-being in adolescents: *A systematic review. Youth & Society,* 53(2), pp.175-210.

54. World Health Organisation. (2022). *Mental health.* [online] World Health Organization. Available at: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response. (Last accessed on: 14th August, 2024).

55. Zhao, N. and Zhou, G. (2021). COVID-19 stress and addictive social media use (SMU): Mediating role of active use and social media flow. *Frontiers in psychiatry*, 12, p.635546.

Bibliography

1. Draw.io. (2024). *Flowchart Maker & Online Diagram Software*. [online] app.diagrams.net. Available at: https://app.diagrams.net/. (Last accessed on: 20th August, 2024).

2. W3Schools. (1998). *W3Schools Online Web Tutorials*. [online] W3schools.com. Available at: https://www.w3schools.com/. (Last accessed on: 16th August 2024).

Appendices

16/08/2024, 18:15 Copy of Untitled6.ipynb - Colab

```
import pandas as pd

from scipy.stats import chi2_contingency

survey_data = pd.read_csv('/content/sample_data/Hypothesis testing - mental effects.csv')

contingency_table = pd.crosstab(survey_data['How much time do you spend on the above mentioned platform daily?'], survey_data['Any negati

chi2, p, dof, expected = chi2_contingency(contingency_table)

print("Contingency Table:")
print(contingency_table)
```

```
Contingency Table:
Any negative impact on mental health from the following options?  Depression  \
How much time do you spend on the above mention...
1 - 3 hours                                                                  0
3 - 5 hours                                                                  1
7-8 hours                                                                    0
Depends upon mood                                                            0
Less than an hour                                                            0

Any negative impact on mental health from the following options?  Depression, Stress  \
How much time do you spend on the above mention...
1 - 3 hours                                                                          1
3 - 5 hours                                                                          3
7-8 hours                                                                            1
Depends upon mood                                                                    0
Less than an hour                                                                    0

Any negative impact on mental health from the following options?  Just feel tired sometimes  \
How much time do you spend on the above mention...
1 - 3 hours                                                                           1
3 - 5 hours                                                                           0
7-8 hours                                                                             0
Depends upon mood                                                                     0
Less than an hour                                                                     0

Any negative impact on mental health from the following options?  No such effect  \
How much time do you spend on the above mention...
1 - 3 hours                                                                  34
3 - 5 hours                                                                  18
7-8 hours                                                                     0
Depends upon mood                                                             1
Less than an hour                                                            20

Any negative impact on mental health from the following options?  No such effect, It became an habit now.   \
How much time do you spend on the above mention...
1 - 3 hours                                                                                              1
3 - 5 hours                                                                                              0
7-8 hours                                                                                                0
Depends upon mood                                                                                        0
Less than an hour                                                                                        0

Any negative impact on mental health from the following options?  Stress
How much time do you spend on the above mention...
1 - 3 hours                                                          9
3 - 5 hours                                                          8
7-8 hours                                                            0
Depends upon mood                                                    0
Less than an hour                                                    4
```

```
print("\nChi-Square Statistic:", chi2)
```

```
Chi-Square Statistic: 29.84015655577299
```

```
print("p-value:", p)
```

```
p-value: 0.07248542879792748
```

```
print("Degrees of Freedom:", dof)
```

https://colab.research.google.com/drive/1VIdlocuB3vA6qF7zX9VIGIAYXHYUo4j4#printMode=true 1/2

*Figure 24 - Back-end python language coding on Google Colab as part of Hypothesis testing part – I (*Google.com, 2019*)*

```
Degrees of Freedom: 20


print("Expected Frequencies Table:")
print(expected)

Expected Frequencies Table:
[[4.50980392e-01 2.25490196e+00 4.50980392e-01 3.29215686e+01
  4.50980392e-01 9.47058824e+00]
 [2.94117647e-01 1.47058824e+00 2.94117647e-01 2.14705882e+01
  2.94117647e-01 6.17647059e+00]
 [9.80392157e-03 4.90196078e-02 9.80392157e-03 7.15686275e-01
  9.80392157e-03 2.05882353e-01]
 [9.80392157e-03 4.90196078e-02 9.80392157e-03 7.15686275e-01
  9.80392157e-03 2.05882353e-01]
 [2.35294118e-01 1.17647059e+00 2.35294118e-01 1.71764706e+01
  2.35294118e-01 4.94117647e+00]]
```


*Figure 25 - Back-end python language coding on Google Colab as part of Hypothesis testing part – II (*Google.com, 2019*)*

16/08/2024, 18:22 Untitled7.ipynb - Colab

```
import pandas as pd

from scipy.stats import chi2_contingency

survey_data = pd.read_csv('/content/sample_data/Hypothesis testing - physical effects.csv')

contingency_table = pd.crosstab(survey_data['How much time do you spend on the above mentioned platform daily?'], survey_data['Any negat:

chi2, p, dof, expected = chi2_contingency(contingency_table)

print("Contingency Table:")
print(contingency_table)
```

```
Contingency Table:
Any negative impact on physical health from the following options after prolonged use?  Body pain  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                    3
3 - 5 hours                                                                                    2
7-8 hours                                                                                      0
Depends upon mood                                                                              0
Less than an hour                                                                              0

Any negative impact on physical health from the following options after prolonged use?  Body pain, Eye strain  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                                 2
3 - 5 hours                                                                                                 2
7-8 hours                                                                                                   0
Depends upon mood                                                                                           0
Less than an hour                                                                                           0

Any negative impact on physical health from the following options after prolonged use?  Body pain, Eye strain, Headache  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                                              2
3 - 5 hours                                                                                                              0
7-8 hours                                                                                                                0
Depends upon mood                                                                                                        0
Less than an hour                                                                                                        0

Any negative impact on physical health from the following options after prolonged use?  Eye strain  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                 8
3 - 5 hours                                                                                 4
7-8 hours                                                                                   0
Depends upon mood                                                                           0
Less than an hour                                                                           7

Any negative impact on physical health from the following options after prolonged use?  Eye strain, Headache  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                              2
3 - 5 hours                                                                                              9
7-8 hours                                                                                                1
Depends upon mood                                                                                        0
Less than an hour                                                                                        2

Any negative impact on physical health from the following options after prolonged use?  Headache  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                 7
3 - 5 hours                                                                                 2
7-8 hours                                                                                   0
Depends upon mood                                                                           0
Less than an hour                                                                           1

Any negative impact on physical health from the following options after prolonged use?  No such effect  \
How much time do you spend on the above mention...
1 - 3 hours                                                                                    21
3 - 5 hours                                                                                    11
7-8 hours                                                                                       0
Depends upon mood                                                                               1
Less than an hour                                                                              14

Any negative impact on physical health from the following options after prolonged use?  No such effect, Go to the gym daily
```

```
print("\nChi-Square Statistic:", chi2)
```

```
Chi-Square Statistic: 29.181391603430413
```

https://colab.research.google.com/drive/1DjkuZYfKAhQV67RdSgLAM64cCJWc_yJy#scrollTo=n_vRqcmVNWfq&printMode=true 1/2

*Figure 26 - Back-end python language coding on Google Colab as part of Hypothesis testing part – III (*Google.com, 2019*)*

```
print("p-value:", p)
```

```
p-value: 0.4033710158497259
```

```
print("Degrees of Freedom:", dof)
```

```
Degrees of Freedom: 28
```

```
print("Expected Frequencies Table:")
print(expected)
```

```
Expected Frequencies Table:
[[2.25490196e+00 1.80392157e+00 9.01960784e-01 8.56862745e+00
  6.31372549e+00 4.50980392e+00 2.11960784e+01 4.50980392e-01]
 [1.47058824e+00 1.17647059e+00 5.88235294e-01 5.58823529e+00
  4.11764706e+00 2.94117647e+00 1.38235294e+01 2.94117647e-01]
 [4.90196078e-02 3.92156863e-02 1.96078431e-02 1.86274510e-01
  1.37254902e-01 9.80392157e-02 4.60784314e-01 9.80392157e-03]
 [4.90196078e-02 3.92156863e-02 1.96078431e-02 1.86274510e-01
  1.37254902e-01 9.80392157e-02 4.60784314e-01 9.80392157e-03]
 [1.17647059e+00 9.41176471e-01 4.70588235e-01 4.47058824e+00
  3.29411765e+00 2.35294118e+00 1.10588235e+01 2.35294118e-01]]
```


*Figure 27 - Back-end python language coding on Google Colab as part of Hypothesis testing part – IV (*Google.com, 2019*)*

28/08/2024, 17:54 Emotional analysis_1 - Colab

Installing necessary libraries

```
pip install pandas numpy scikit-learn nltk matplotlib seaborn
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (2.1.4)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (1.26.4)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.10/dist-packages (1.3.2)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: seaborn in /usr/local/lib/python3.10/dist-packages (0.13.1)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.1)
Requirement already satisfied: scipy>=1.5.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.13.1)
Requirement already satisfied: joblib>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.4.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (3.5.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2024.5.15)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.5)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.2.1)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (4.53.1)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.5)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (24.1)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (9.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (3.1.4)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
```

Importing installed libraries

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.pipeline import make_pipeline
from sklearn.metrics import classification_report, accuracy_score
import nltk
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
import re
```

Loading the data file & printing file content to ensure the right file is uploaded.

```
df = pd.read_csv('/content/sample_data/Emotional analysis.csv')
print(df.head())
```

```
                                      Feedback Emotions
0  It's good to get the general awareness about t...    Happy
1  Everything in moderation is good. Even water i...    Happy
2                       Sometimes it looks boring    Doubt
3  It reduces ability to focus due to the fact th...      Sad
4  Social media use has become an integral part o...    Doubt
```

Downloading tools for removing unused words and database for machine learning program

```
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('punkt')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
True
```

Starting the downloaded NLKT tools

```
lemmatizer = WordNetLemmatizer()
stop_words = set(stopwords.words('english'))
```

Cleansing the data with defining the functions used to clean it

https://colab.research.google.com/drive/1pQyunaRcpfaxRGhDTFlfbO1c-5nFCgTq#scrollTo=x_KfYGd-4iLy&printMode=true 1/4

Figure 28 - Emotional analysis back-end python language coding on Google Colab, Part – I (Google.com, 2019)

```
def clean_text(text):
    text = re.sub(r'\W', ' ', text)
    text = text.lower()
    tokens = nltk.word_tokenize(text)
    tokens = [lemmatizer.lemmatize(word) for word in tokens if word not in stop_words]
    return ' '.join(tokens)


df['Feedback'] = df['Emotions'].apply(clean_text)
```

Splitting the data

```
x = df['Feedback']
y = df['Emotions']


x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
```

Conversion of text to numerical format

```
tfidf = TfidfVectorizer(max_features=5000)
x_train_tfidf = tfidf.fit_transform(x_train)
x_test_tfidf = tfidf.transform(x_test)
```

Selecting the machine learning model - Naive Bayes classifier

```
model = MultinomialNB()
model.fit(x_train_tfidf, y_train)
```

```
▾ MultinomialNB
MultinomialNB()
```

Making predictions

```
y_pred = model.predict(x_test_tfidf)
```

Testing the model

```
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy}")
print(classification_report(y_test, y_pred))
```

```
Accuracy: 1.0
              precision    recall  f1-score   support

       Doubt       1.00      1.00      1.00         4
       Happy       1.00      1.00      1.00         2
         Sad       1.00      1.00      1.00         2

    accuracy                           1.00         8
   macro avg       1.00      1.00      1.00         8
weighted avg       1.00      1.00      1.00         8
```

Using grid search to locate the hyperparameters for the used model

```
from sklearn.model_selection import GridSearchCV
param_grid = {
    'alpha': [0.1, 0.5, 1.0, 2.0]
}
grid_search = GridSearchCV(MultinomialNB(), param_grid, cv=5, scoring='accuracy')
grid_search.fit(x_train_tfidf, y_train)
print(f"Best Parameters: {grid_search.best_params_}")
print(f"Best Score: {grid_search.best_score_}")
```


*Figure 29 - Emotional analysis back-end python language coding on Google Colab, Part – II (*Google.com, 2019*)*

```
Best Parameters: {'alpha': 0.1}
Best Score: 1.0
```

Cross validation

```
from sklearn.model_selection import cross_val_score
cv_scores = cross_val_score(model, x_train_tfidf, y_train, cv=5)
print(f"Mean Cross-Validation Score: {np.mean(cv_scores)}")
```

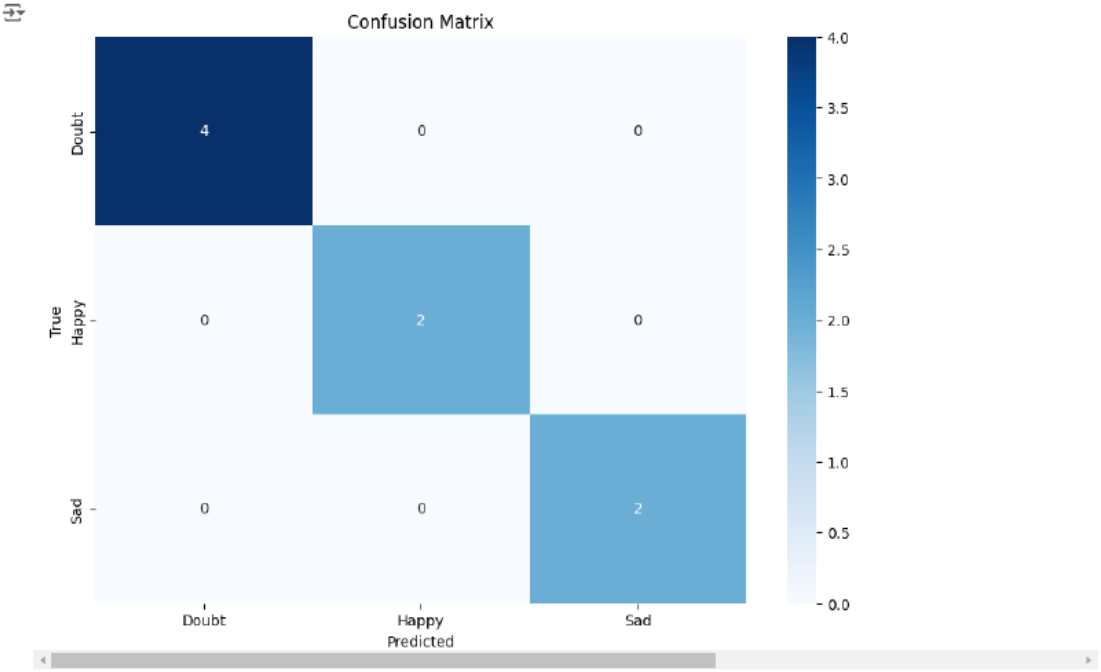
```
Mean Cross-Validation Score: 1.0
```

Reading the obtained results

```
from sklearn.metrics import confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt

cm = confusion_matrix(y_test, y_pred, labels=model.classes_)

plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=model.classes_, yticklabels=model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()
```



Testing the model on real-time examples

```
new_text = ["I am doubtful."]
new_text_cleaned = [clean_text(text) for text in new_text]
new_text_tfidf = tfidf.transform(new_text_cleaned)
predictions = model.predict(new_text_tfidf)
print(predictions)
```

```
['Doubt']
```

```
new_text = ["I am happy."]
new_text_cleaned = [clean_text(text) for text in new_text]
new_text_tfidf = tfidf.transform(new_text_cleaned)
predictions = model.predict(new_text_tfidf)
print(predictions)
```

```
['Happy']
```

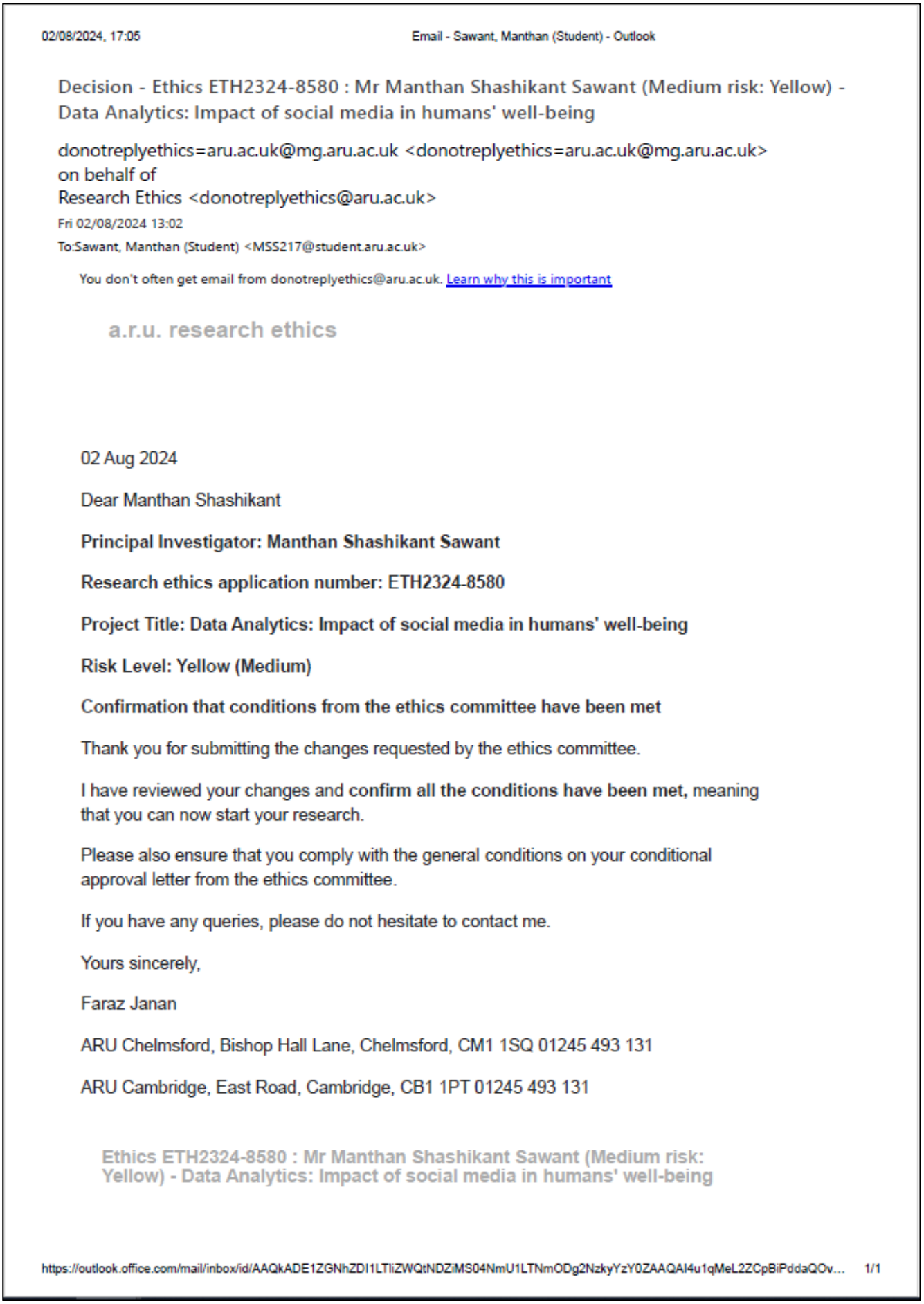

*Figure 30 - Emotional analysis back-end python language coding on Google Colab, Part – III (*Google.com, 2019*)*

```
new_text = ["I am sad."]
new_text_cleaned = [clean_text(text) for text in new_text]
new_text_tfidf = tfidf.transform(new_text_cleaned)
predictions = model.predict(new_text_tfidf)
print(predictions)
```

```
['Sad']
```


Figure 31 - Emotional analysis back-end python language coding on Google Colab, Part – IV (Google.com, 2019)

Figure 32 – Ethics training certificate



02/08/2024, 17:05 Email - Sawant, Manthan (Student) - Outlook

Decision - Ethics ETH2324-8580 : Mr Manthan Shashikant Sawant (Medium risk: Yellow) - Data Analytics: Impact of social media in humans' well-being

donotreplyethics=aru.ac.uk@mg.aru.ac.uk <donotreplyethics=aru.ac.uk@mg.aru.ac.uk>
on behalf of
Research Ethics <donotreplyethics@aru.ac.uk>
Fri 02/08/2024 13:02
To:Sawant, Manthan (Student) <MSS217@student.aru.ac.uk>

You don't often get email from donotreplyethics@aru.ac.uk. Learn why this is important

a.r.u. research ethics

02 Aug 2024

Dear Manthan Shashikant

Principal Investigator: Manthan Shashikant Sawant

Research ethics application number: ETH2324-8580

Project Title: Data Analytics: Impact of social media in humans' well-being

Risk Level: Yellow (Medium)

Confirmation that conditions from the ethics committee have been met

Thank you for submitting the changes requested by the ethics committee.

I have reviewed your changes and **confirm all the conditions have been met,** meaning that you can now start your research.

Please also ensure that you comply with the general conditions on your conditional approval letter from the ethics committee.

If you have any queries, please do not hesitate to contact me.

Yours sincerely,

Faraz Janan

ARU Chelmsford, Bishop Hall Lane, Chelmsford, CM1 1SQ 01245 493 131

ARU Cambridge, East Road, Cambridge, CB1 1PT 01245 493 131

Ethics ETH2324-8580 : Mr Manthan Shashikant Sawant (Medium risk: Yellow) - Data Analytics: Impact of social media in humans' well-being

https://outlook.office.com/mail/inbox/id/AAQkADE1ZGNhZDI1LTliZWQtNDZiMS04NmU1LTNmODg2NzkyYzY0ZAAQAI4u1qMeL2ZCpBiPddaQOv... 1/1

Figure 33 – Conditional approval confirmation certificate