**Mental Health Prediction using**

**Machine Learning**

**A Project Work Synopsis**

*Submitted in the partial fulfillment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE WITH SPECIALIZATION IN**

**INTERNET OF THINGS**

**Submitted by:**

21BCS4626 NIMIT GARG

21BCS5381 REYAAN SACHDEVA

**Under the Supervision of:**

MR ANKUR SHARMA



# Abstract

This project endeavors to create a predictive model for mental health estimation by analyzing diverse lifestyle factors, with a primary emphasis on stress prediction. The data set incorporates a wide array of participant information, spanning age, gender, and occupation (Student, Corporate, Others), along with sleep patterns, physical activity levels, dietary habits, substance use, and more. Employing a machine learning regression model, the project aims to establish correlations between these factors, self-reported stress scale responses, and physiological indicators. Ethical considerations are fundamental to the project's scope, ensuring participant privacy and responsible data handling. By focusing on stress prediction as a proxy for mental well-being, the project addresses the challenges associated with directly predicting mental health outcomes, offering a nuanced approach to ethical concerns. Real-world applications extend to personalized interventions in counseling, corporate wellness programs, and health care strategies. Acknowledging the growing importance of mental health awareness, the project recognizes the critical intersection of technology and mental health, contributing to a more comprehensive understanding of stress factors and mental well-being. This initiative seeks to bridge the gap between technological advancements and holistic mental health support, promoting a nuanced and personalized approach to individual well-being

**Keywords:** Machine Learning **,** Real Time data , Survey , Mental well-being.

# Table of Contents

|  |  |
| --- | --- |
| Title Page | i |
| Abstract | ii |
| 1.    Introduction |  |
| 1.1  Problem Definition |  |
| 1.2 Project Overview |  |
| 1.3 Hardware Specification |  |
| 1.4 Software Specification |  |
| 2.    Literature Survey |  |
| 2.1 Existing System |  |
| 2.2 Proposed System |  |
| 2.3 Literature Review Summary |  |
| 3.    Problem Formulation |  |
| 4.    Research Objective |  |
| 5.    Methodologies |  |
| 6.    Experimental Setup |  |
| 7.    Conclusion |  |
| 8.    Tentative Chapter Plan for the proposed work |  |
| 9.    Reference |  |

# 1. INTRODUCTION

## 1.1 Problem Definition

## In contemporary society, the omnipresence of smart phones has led to concerns about its potential impact on mental health, particularly among students and professionals. Excessive usage has been associated with heightened stress levels and anxiety. The project's problem definition centers on addressing this intricate relationship between technology use, lifestyle factors, and mental well-being. By understanding and predicting stress levels, the project aims to contribute valuable insights to the broader field of mental health analysis.

## 1.2 Problem Overview

## In the contemporary digital landscape, the ubiquitous presence of smart phones has become an integral part of daily life, profoundly influencing various aspects of human behavior and well-being. While these technological advancements offer unparalleled connectivity and convenience, concerns about their impact on mental health, particularly among students and professionals, have gained prominence. The core problem lies in the potential negative effects of excessive smart phone usage on mental well-being, specifically manifesting as increased stress levels and heightened anxiety.

## As individuals increasingly rely on smart phones for communication, information access, and entertainment, understanding the nuanced relationship between their usage patterns and mental health outcomes becomes imperative. Excessive smart phone use has been associated with a range of mental health challenges, including disrupted sleep patterns, decreased physical activity, and potential addiction-like behaviors.

## The problem overview involves a thorough exploration of the intricate connections between lifestyle factors, stress levels, and overall mental well-being. It recognizes the need for a nuanced understanding that goes beyond mere correlation, aiming to provide actionable insights for individuals, mental health professionals, and relevant stakeholders. By predicting stress levels based on a diverse set of lifestyle factors, the project aims to contribute to the early identification of potential mental health challenges and facilitate proactive interventions.

## 1.3 Hardware Specification

The hardware specifications for this project are minimal, requiring standard computing equipment such as laptops or desktops. The emphasis is on accessibility and ease of implementation, ensuring that the project can be executed without the need for specialized or high-performance hardware.

## 1.4 Software Specification

The software specification for this project includes:

**Google Collab:**

Google Collaboratory (Collab) is a cloud-based platform that facilitates collaborative and interactive development of machine learning models. With its seamless integration with Google Drive, it provides a free and convenient environment for running Jupyter notebooks, enabling easy access to shared resources and collaborative work on machine learning projects.

**Visual Studio Code (VSCode):**

Visual Studio Code is a versatile and lightweight code editor that supports various programming languages, including Python. Its rich set of features, extensions, and integrated Git support make it an ideal choice for coding, debugging, and version control in machine learning projects. VSCode provides a user-friendly interface and extensible for enhancing the development experience.

**Jupyter Notebook:**

Jupyter Notebook is an open-source web application that allows the creation and sharing of live code, equations, visualizations, and narrative text. Widely used in data science and machine learning, Jupyter Notebooks provide an interactive environment for prototyping and presenting code, making it an indispensable tool for exploratory data analysis, model development, and collaborative research.

# 2. LITERATURE SURVEY

## 2.1 Existing System

## 2.2 Proposed System

The proposed system aims to develop a comprehensive framework for assessing and addressing mental health issues among university students, focusing on the interplay between smart phone usage behaviors and depression and anxiety symptoms. The system comprises several key components: data collection, machine learning models, feature selection, prediction and analysis, intervention strategies, and evaluation and feedback. Data will be collected through surveys or questionnaires, incorporating demographic information, smartphone usage patterns, and mental health indicators. Machine learning algorithms will analyze the collected data to predict scores on standardized scales such as the PHQ-9, GAD-7, and SAS. Relevant features associated with smart phone usage and mental health outcomes will be identified through feature selection techniques. Predicted depression and anxiety scores will be analyzed to identify patterns or correlations with smart phone usage behaviors. Based on these findings, targeted intervention strategies will be developed to support students’ mental well-being. Continuous evaluation and feedback will ensure ongoing improvement and adaptation of the system to meet the evolving needs of university students.

## 2.3 Literature Review Summary (Minimum 7 articles should refer)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# 3. PROBLEM FORMULATION

Mental health issues, particularly depression and anxiety, remain significant challenges for university students. Due to the widespread use of smart phones in this population, there is a growing interest in understanding the relationship between smart phone usage patterns and mental health effects. The goal of this study is to use machine learning techniques to assess symptoms of depression and anxiety in college students using data collected from a variety of characteristics, including smart phone use and mental health indicators. We specifically focus on predicting scores on three widely used mental health rating scales: the Patient Health Questionnaire-9 (PHQ-9), the Generalized Anxiety Disorder-7 (GAD-7), and the Self-Rating Anxiety Scale (SAS). To achieve our goals, we use machine learning algorithms including Gaussian Naive Bayes, Random Forest, and Support Vector Machine (SVM) to predict PHQ-9, GAD-7, and SAS scores from the provided data set. These models are trained using measures of anxiety/stress levels, social engagement, and well-being, among others.The data set used in the study includes demographic data, mobile phone usage patterns, social data.This study contributes to an understanding of the complex interaction between smart phone behaviors and mental health outcomes among college students. Using machine learning techniques, we aim to uncover valuable insights that can guide future research and inform evidence-based interventions to support student mental health.\

# 4. OBJECTIVES

The objectives of this project are as follows:

**Develop a Predictive Model:** Design and implement a machine learning regression model capable of predicting stress levels based on diverse lifestyle factors. The model will serve as a tool to estimate an individual's mental well-being indirectly through stress predictions.

**Data Collection and Diversity:** Gather a comprehensive dataset by designing detailed questionnaires and ensuring participation across diverse demographic groups, including various age ranges, genders, and occupations. This objective aims to capture a representative sample for robust model training.

**Real-world Applicability:** Integrate the predictive model into practical applications by collaborating with counseling departments, psychologists, and health care professionals. Develop personalized interventions, wellness programs, and health care strategies based on the stress predictions to address real-world mental health challenges.

**Focus on Stress as a Proxy:** Concentrate on stress prediction as a valuable proxy for mental well-being assessment. By avoiding direct mental health predictions, the project aims to navigate ethical concerns while still providing meaningful insights for mental health professionals and individuals.

**Technology and Mental Health Intersection**: Recognize and explore the critical intersection of technology and mental health. Leverage technological advancements to contribute to a deeper understanding of stress factors, offering insights that can be applied in counseling, psychology, and medical science.

# METHODOLOGY

The study collected information from 1024 participants, including instructors and students from a range of university academic programme. 24 records were eliminated following an initial screening procedure because of errors or inconsistencies. The age range of the participants was 18–46 years old. With 471 participants, the gender distribution was slightly skewed towards men, with the remainder participants being women. A smaller cohort of teachers and 618 pupils were also included in the data set. The attendees came from a range of academic backgrounds, with the most common streams being business administration, marketing, and computer science.

Here is the methodology of the Proposed Model:

* **Data Collection:** Gathered data from 1024 university students and faculty across multiple disciplines. Excluded 24 records due to data quality issues. Final data-set consists of 1000 entries with age ranging from 18 to 46 years. Gender distribution: 471 males, 529 females.
* **Data Pre-Processing**: Handled missing values, encoded categorical variables, and scaled features.
* **Model Selection**: Chose Random Forest and Support Vector Machine (SVM) algorithms for classification tasks.
* **Data Visualization**: Utilized histograms, heat maps, and correlation matrices to visualize relationships between variables.
* **Model Evaluation**: Assessed model performance using classification reports, confusion matrices, and evaluation metrics (accuracy, precision, recall, F1-score).
* **Model Validation**: Employed holdout validation and k-fold cross-validation for model validation. Optimized models based on validation results to ensure reliability.

# 6.EXPERIMENTAL SETUP

# 7.CONCLUSION

In conclusion, this project addresses the pressing issue of understanding and mitigating the impact of smart phone usage on mental well-being. The developed predictive model, centered on stress prediction as a proxy for mental health, offers a valuable tool for estimating individuals' psychological states. Ethical considerations have been paramount throughout the project, ensuring participant privacy and responsible data handling.

The real-world applications of the predictive model extend to personalized interventions in counseling, wellness programs, and health care strategies. The iterative refinement process, guided by stakeholder feedback, underscores the commitment to continuous improvement and adaptability. By navigating the ethical challenges associated with mental health predictions and prioritizing collaboration with mental health professionals, this project strives to contribute meaningfully to mental health awareness.

The intersection of technology and mental health is a key focus, leveraging advancements for a more comprehensive understanding of the intricate relationship between lifestyle factors and psychological well-being. Ultimately, this project stands as a step forward in fostering a healthier and more informed approach to mental health in the digital age.

In essence, this project aspires to bridge the gap between technology and mental health, providing a nuanced and actionable approach to mental well-being in an increasingly digital world. Through responsible research practices, ethical considerations, and collaborative efforts with mental health professionals, it endeavors to contribute to a holistic understanding of the factors influencing our mental health and pave the way for improved interventions and support systems.

## 8. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

**CHAPTER 1: INTRODUCTION**

**CHAPTER 2: LITERATURE REVIEW**

**CHAPTER 3: OBJECTIVE**

**CHAPTER 4: METHODOLOGIES**

**CHAPTER 5: EXPERIMENTAL SETUP**

**CHAPTER 6: CONCLUSION AND FUTURE SCOPE**

## REFERENCES

1. Seo, J.G. and Park, S.P., 2015. Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. The journal of headache and pain, 16, pp.1-7.
2. Zhong, Q.Y., Gelaye, B., Zaslavsky, A.M., Fann, J.R., Rondon, M.B., Sanchez, S.E. and Williams, M.A., 2015. Diagnostic validity of the ´ generalized anxiety disorder-7 (GAD-7) among pregnant women. PloS one, 10(4), p.e0125096.
3. Sun, Y., Fu, Z., Bo, Q., Mao, Z., Ma, X. and Wang, C., 2020. The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. BMC psychiatry, 20, pp.1-7.
4. Spitzer, R. L., Kroenke, K., Williams, J. B., Lowe, B. (2006). A ¨ brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine, 166(10), 1092–1097.
5. Eisenberg, D., Hunt, J., Speer, N. (2012). Mental health in American colleges and universities: Variation across student subgroups and across campuses. Journal of Nervous and Mental Disease, 200(5), 1–7.
6. Thomee, S., 2018. Mobile phone use and mental health. A review of the ´ research that takes a psychological perspective on exposure. International journal of environmental research and public health, 15(12), p.2692.
7. Escalera-Chavez, M.E. and Rojas-Kramer, C.A., 2020. SAS-SV Smart- ´ phone Addiction Scale in Mexican University Students. Education Research International, 2020, pp.1-10.

[8] Phelan, E., Williams, B., Meeker, K., Bonn, K., Frederick, J., LoGerfo, J. and Snowden, M., 2010. A study of the diagnostic accuracy of the PHQ-9 in primary care elderly. BMC family practice, 11, pp.1-9