GODFREY OKOYE UNIVERSITY, ENUGU

FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

DEVELOPMENT OF A MACHINE LEARNING-BASED DEPRESSION RISK PREDICTION SYSTEM FOR UNIVERSITY STUDENTS

A FINAL YEAR PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF BACHELOR OF SCIENCE (B.Sc.) DEGREE IN COMPUTER SCIENCE

BY

OKECHI AUGUSTINE KENECHUKWU

SUPERVISOR:

MR. CAMILUS EKENE

JUNE, 2026

CERTIFICATION

This is to certify that this project titled "DEVELOPMENT OF A MACHINE LEARNING-BASED DEPRESSION RISK PREDICTION SYSTEM FOR UNIVERSITY STUDENTS" was carried out by me, OKECHI AUGUSTINE KENECHUKWU, with registration number “GOU/U22/CSC/846” under the supervision of Mr. Njoku Camilus in the Department of Computer Science, Godfrey Okoye University, Enugu, in partial fulfilment of the requirements for the award of the degree of Bachelor of Science (B.Sc.) in Computer Science.

______________________________ ______________________________

Mr. Njoku Camilus Date

Project Supervisor

______________________________ ______________________________

Dr. Frank U. Okebanama Date

Head of Department

______________________________ ______________________________

Prof. Ifeyinwa Ajah Date

Dean Of Faculty

______________________________ ______________________________

External Examiner Date

DEDICATION

This work is dedicated to God Almighty for guiding me every step of the way with His wisdom and grace. It's also for my family. Thank you for your constant encouragement, support, and sacrifices. Lastly, it goes out to all university students battling depression and other mental health issues silently. Here's hoping this little contribution makes help just a bit more reachable.

ACKNOWLEDGEMENT

All glory and honour belong to God Almighty, whose grace, wisdom and strength sustained me throughout this programme and made this work possible.

I am sincerely grateful to my project supervisor, Dr. Njoku Camilus, for his guidance, patience and immensely valuable feedback throughout this research. I also thank the Head of Department, Dr. Frank U. Okebanama, the Dean of the Department, Prof. Ifeyinwa Ajah, and all the lecturers of the Department of Computer Science, Godfrey Okoye University, for their dedication and support.

My deepest gratitude goes to my father, Mr. John Okechi, my family members and friends for their constant love, prayers and encouragement. Your belief in me kept me going. Finally, I thank all participants in the usability testing for their time and honest feedback.

ABSTRACT

University students' depression is a big public health issue, with studies showing high rates that hurt grades, social life, and well-being. Traditional screening methods are usually reactive, costly, and lack access for many students. So, there's a real need for tech-based, early detection tools that can reach more students easily.

This project is about an online system that predicts depression risk in university students. It uses the Depression Anxiety Stress Scale (DASS-21) along with machine learning. The system gathers data through twenty-one questions split between depression, anxiety, and stress, and includes five extra lifestyle factors—sleep, exercise, social life, appetite, and academic pressure.

The project uses a Random Forest classifier on data from 1,120 students. This method reached 89.3% accuracy and an AUC-ROC of 0.93. That's better than other models like logistic regression, SVM, and K-Nearest Neighbors.

The system gives a risk classification of Low, Moderate, or High, along with a confidence score and personalized recommendations. It's built with React.js for the frontend, which offers a responsive multi-step assessment interface. For the backend, they used Python and Flask to create a RESTful API for real-time results. This tool is meant for preliminary screenings only—to help with early referrals and support in university health services—but isn't meant to take the place of a full clinical assessment

Keywords: Depression prediction, DASS-21, machine learning, Random Forest, student mental health, Flask, React.js, web-based system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

University student's mental health is considered as one of the most important public health issues of the 21 st century. Depression is already considered as one of the top three disabilities and WHO estimates that the number of depressed persons around the world is over 280 million for any age category (WHO, 2023). University students form a particularly vulnerable group, as entering higher education may involve a wide variety of stressors like high academic pressure, social isolation, economic strain, lack of close family contacts, all of which may trigger or exacerbate a depressive episode. So, it is highly expected that most of these episodes remain undiagnosed, untreated because of stigma, lack of resources and screening tests.

(Li et al., 2022).

Similar and disturbing figures are being reported in the undergraduate population in low- and middle-income countries, Nigeria being no exception. Depressive illness ranging between 15% and above 40% has been observed in Nigerian universities but most mental health services are severely lacking in financial, professional and other resources compared to the needs of the affected population (Ladi-Akinyemi et al., 2023). The problem is that undergraduate students, who do not receive help for depression, are most likely to suffer in silence with serious impact on their academic functioning, relationships and ultimate functioning throughout life. Many episodes that might have been successfully treated with interventions if received early, now evolve to chronic forms and necessitate complex clinical treatment.

University-based conventional mental health screening utilizes either clinically administered interviews or paper and pencil based psychometric instruments administered by mental health experts. Though these are useful in diagnosis and therapy they are expensive, inaccessible and lack the timely screening function. A limited supply of mental health experts in the Nigerian universities ensures that the majority of students experiencing symptoms of depression are not assessed clinically. Thus, the potential of using technology, particularly web based applications and machine learning algorithms to provide an accessible, low cost, anonymous and scalable preliminary screening tool to the Nigerian university student population has recently attracted considerable attention.

Machine learning (ML) is a subfield of artificial intelligence which uses computational algorithms that are trained on data. The systems are then able to recognize patterns within data in order to predict, or take, decisions on other, unseen data without being specifically programmed to do so for each scenario (Aleem et al., 2022). Within mental health, ML algorithms have demonstrated significant success in the recognition of the symptoms of depression, anxiety and stress using data from questionnaire responses, physiological signals and behavioral patterns (Saleh et al., 2023). From a range of different algorithms analyzed, the Random Forest classifier has proved particularly accurate and resilient across all data from the mental health domains as it uses an ensemble of tree-based classifiers which means it is less likely to overfit on the data due to multivariate, non-linear interactions between the features (Aleem et al., 2022).

The DASS-21 or Depression, Anxiety and Stress Scale-21, is a self-report questionnaire, comprising of 21 items that is used to determine the severity of depression, anxiety and stress as three separate constructs which are themselves closely related. This contrasts to that of the PHQ-9, as this tool would only extract features pertaining to depression whereas the DASS-21 generates multidimensional data relating to these three closely linked phenomena which provides more rich information for ML models than a simple questionnaire with limited dimensions. Due to its high validation amongst African and Asian university populations and its availability as open-access data, the DASS-21 seems the most applicable questionnaire to implement in the context of a Nigerian university.

This project proposes and implements an intelligent web-based system that administers the DASS-21 questionnaire to university students alongside a set of contextual lifestyle indicators and applies a trained Random Forest classifier to predict the student's depression risk category as Low, Moderate, or High. The system is designed with React.js as front-end framework, and a Python Flask API as back-end. This is designed in such a way that it gives student seamless and anonymous access. The aim is to give an accessible and non-stigmatic first point of contact for students who might need a professional psychiatric help.

1.2 Statement of The Problem

While the rate of depression among Nigerian university students, and sub-Saharan Africa at large is quite alarming, the integration of formal mental health screening at the institutional level, is scarce. The problem includes the following issues:

- Current approaches to mental health screening in Nigerian Universities are largely manual, expensive and resource-intensive (depending on the presence of clinical staff).
- The PHQ-9 used in the previous iteration of the system provides only nine data points and focuses primarily on depression symptoms, disregarding the relevance of associated anxiety and stress to the mental health of students, thereby narrowing the feature space available for machine learning models.
- The level of stigma attached to receiving mental health support on a university campus means that many students are hesitant to voluntarily seek counseling services. The effectiveness of anonymous electronic platforms is in decreasing barriers.
- The lack of a scalable, automated, and data-driven tool for identifying depression risk means that at-risk students are rarely identified before their condition deteriorates to the point that intensive clinical intervention is required.
- Since there are no automated, data-driven, and scalable tool to predict the risk of depression, at-risk students are often missed until their problem becomes severe enough to necessitate significant clinical intervention.

Addressing the above, this work provides an end-to-end intelligent system which gathers students' DASS-21 and lifestyle factors through a web interface and presents the instant, machine-learning based prediction of depression risk with relevant personal feedback.

1.3 Aim and Objectives of the Study

This study aims to develop and implement an intelligent web-based system for predicting the risk of depression in university students with the aid of the DASS-21 scale and machine learning. In particular, the objectives driving the creation of this system includes:

1. Review and compare existing depression screening instruments and machine learning techniques for mental health prediction in university student populations.
2. Collect, preprocess and balance a 1,120-record student mental health dataset using SMOTE, StandardScaler and stratified 80/20 train-test splitting.
3. Design and train a Random Forest classifier using 26 features (21 DASS-21 items + 5 lifestyle indicators) and evaluate with accuracy, precision, recall, F1 and AUC-ROC.
4. Develop a Python Flask RESTful API (POST /predict) to serve real-time risk predictions from the serialized model.
5. Build and integrate a React.js multi-step frontend application with client-side validation, results display and personalized recommendations.

1.4 Significance of the Study

The value of this study lies on academic, clinical and socio-technical levels. On the academic level, this research contributes to an existing and growing area of research where Machine Learning techniques are applied to the screening of mental health in the context of an African university, an area largely neglected by research worldwide. The selection of the DASS-21 instead of the widely studied PHQ-9 offers a more detailed, multi-sub-scale feature set with the potential to develop more discriminative models and also contributes to methodology best practice in the field.

Clinically and institutionally, the system offers university health centers and counselling services an economical and scalable tool to conduct basic preliminary screening of university student mental health. By identifying at-risk students for prompt referrals to mental health services, the system could help students avoid reaching the extreme end of the mental health spectrum that requires intensive care, which could in turn support student retention, academic success and long-term mental health.

From the perspective of a socio-technical system, a web-based, anonymous, self-administered system can overcome many of the traditional barriers preventing university students from seeking help for mental health issues. Students can take the assessment from a location and time most convenient to them, from the privacy of their personal device, without concern for social stigma or judgment.

1.5 Scope of the Study

This study centers on designing, developing, and assessing an intelligent web-based system capable of predicting a university student's risk of depression. This system not only administers the entire 21-item DASS-21 questionnaire but also captures five different lifestyle features: sleep quality, exercise amount, social communication frequency, change in appetite and academic stress perception. Machine learning component is confined to Random Forest classification model trained using a pre-processed data set comprising 1,120 university student records from public researches datasets. System does not provide for its deployment into a cloud environment or into a mobile application and neither is it aimed at being integrated into existing health management information systems of university. It does not give a clinical diagnosis, it is merely designed to provide a first screening. Data privacy and consent issues are also handled within the context of the current session-based, local machine hosted, non-personally identifiable system.

1.6 Limitations of the Study

Several shortcomings of the present study should be addressed. Firstly, the dataset employed for model training, although readily available and a moderate size, was not sourced from Nigerian university students. Thus, there may be a cultural and contextual discrepancy that may reduce the accuracy of the model on specific subpopulations of Nigerian students. Since there may be variations in the ways depression symptoms are presented culturally. Secondly, the model is fully dependent on the user's reported data. The data reported by the user is vulnerable to social desirability bias and inconsistent self reporting depending on their awareness of the questionnaire's purpose. Thirdly, the present system does not utilize physiological and behavioral data (like tracking sleep, phone usage habits and academic performance) which has previously shown in the literature to add to the predictability for classification in mental health detection. Fourthly, the model presents categories of risk (Low, Medium and High) as opposed to continuous severity index. This may make clinical interpretation of results less fine-grained. These shortcomings can be used to build upon the present system.

1.7 Operational Definition of Terms

The terms, which are used within this study, are defined below to ensure clarity of meaning throughout:

6. Depression: A prevalent psychiatric illness marked by sustained feelings of sadness, a loss of interest or pleasure in most activities, disturbances in sleep, eating patterns and concentration, feelings of worthlessness and impaired functioning in most areas of daily life for two or more weeks (American Psychiatric Association, 2013).
7. Depression Anxiety Stress Scale-21 Items (DASS-21): A 21-item self-report scale that is an established measure of depression, anxiety and stress. It has three seven-item sub-scales, one for depression, one for anxiety, and one for stress. The items on each sub-scale are scored on a 4-point Likert scale from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time) (Lovibond and Lovibond, 1995).
8. Machine Learning: An area of artificial intelligence concerned with the design of algorithms that allow computer programs to learn from data and to improve their prediction accuracy through experience rather than being explicitly programmed to perform every specific task (Jordan and Mitchell, 2015).
9. Random Forest: An ensemble machine learning method in which the trees are trained with random subsamples of the data. It outputs the class that is the mode of the classes output by the individual trees, which helps to avoid overfitting and generalize better (Breiman, 2001).
10. SMOTE (Synthetic Minority Oversampling Technique): This technique aims to reduce class imbalance by creating new artificial minority class instances using feature space property from minority class instances (Chawla et al., 2002).
11. Classification of risk: Classification of a student's level of risk of depression into one of three categories: low, moderate, or high, determined by the aggregated DASS-21 subscale scores and lifestyle features after passing through the machine learning model trained on.
12. RESTful API: An Application Programming Interface which abides to the representational state transfer (REST) architectural style, sends messages over HTTP methods (GET, POST) and returns data in JSON.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview of Depression and Mental Health

It is estimated that 280 million people around the world are affected by depression, which according to the World Health Organisation is one of the leading contributors to the global burden of disease; it is responsible for 1 in 8 years lived with disability among all age groups (WHO, 2023). Depression as a clinical disorder includes persistent low mood, anhedonia (lack of pleasure in previously enjoyable activities), impaired cognition and in severe cases suicidal thoughts. A major depressive episode, according to the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition Text Revision, involves the presence of 5 or more symptoms within a 2 week period, including at least one of depressed mood or loss of interest (American Psychiatric Association, 2022). The complexity of the clinical presentation makes it difficult to identify depression without a trained evaluation, however, the condition can be highly responsive to early treatment if it can be identified at a pre-clinical level. The economic consequences of mental health disorders in general, and depression specifically, go far beyond just a decrease in quality of life, to include lost productivity, increased demands on health services and reduced educational success; these combined can result in major economic cost at a national level, and in middle and low-income countries where mental health service provision is limited, a large human and economic cost. According to the Global Burden of Disease Study 2019, depression was in the top ten causes of Disability-Adjusted Life Years for 15-49 year olds, again emphasizing the role of early identification and treatment for this age group (GBD 2019 Mental Disorders Collaborators, 2022).

2.2 Depression Among University Students

University students are a unique demographic at high risk of developing depression and a transition into higher education from secondary school has been shown to be a sensitive period for the onset of psychological disorders. A meta-analysis of the prevalence of depression and anxiety among college students across 64 studies including a total sample of 100,187 individuals reported a pooled depression prevalence of 33.6%, with substantially higher rates observed in Africa and lower-middle-income countries (Li et al., 2022). The meta-analysis of 350,000 college students in 373 campuses between 2013 and 2021 concluded that the proportion of students with at least one mental health issue was over 60%, with nearly a 50% rise from the beginning of the study period (Wang et al., 2023).

The determinants of depression among university students are multivariate, individual, relational and institution based. Individual determinants like personal history of mental illness, coping mechanisms (maladaptive Coping), low academic self-efficacy, and sleep disturbance have been strongly linked to higher depression scores (Bayram & Bilgel, 2008). In the relational aspect the issues are isolation and difficulty in establishing peer relationships, family discord. The institutional based factors include high work load, pressure from examinations, grading systems (competitive), financial stress (mainly among low income students) all have been associated with higher depression scores.

In the Nigerian environment-though studies published are fewer, but the factors are similar as those reported globally with extra individual and institution based factors such as poor infrastructure, irregular academic calendar due to strikes, lack of counseling services in campus. For instance, Adewuya et al. (2006) showed a 26.3% depression rate in Nigerian university students, this was using the PHQ which according to Nwaoha et al. (2022) have similar high rates in universities of south- eastern Nigeria and suggests that easily implementable screening and intervention strategies need to be developed and implemented in Nigeria.

2.3 Review of Depression Screening Instruments

Many reliable and valid psychometric tools have been designed for the clinical and research evaluation of depressive symptoms, with the most commonly used being the Patient Health Questionnaire-9 (PHQ-9), Beck Depression Inventory-II (BDI-II), Hospital Anxiety and Depression Scale (HADS), General Health Questionnaire (GHQ-28), Depression Anxiety Stress Scale (DASS-21), Depression Anxiety Stress Scale 42 (DASS-42), Hamilton Depression Rating Scale (HAM-D), Centre for Epidemiological Studies Depression Scale (CES-D), Major Depression Inventory (MDI), and the Kessler Psychological Distress Scale (K10), and each varies by the number of items, emotions measured, population tested against, and appropriateness as a feature source for use in a machine learning pipeline.

Table 2.1 presents a comparative summary of these instruments.

Instrument	Items	Constructs Assessed	Validation Populations	ML Suitability

PHQ-9	9	Depression only	General adult, primary care	Low(single-scale)
BDI-II	21	Depression severity	Clinical, adult	Moderate
HADS	14	Anxiety and Depression	Medical patients	Moderate(2 subscales)
GHQ-28	28	General mental health (4 subscales)	Community, general adult	Good(4 subscales)
DASS-21	21	Depression, Anxiety, Stress (3 subscales)	University students, general adult, cross-cultural	High(3 subscales + diverse items)
DASS-42	42	Depression, Anxiety, Stress (3 subscales)	Clinical and research settings	Very High (most features, detailed subscales)
HAM-D	17-21	Depression Severity	Clinical Patients	Low (requires clinician; not self-report)
CES-D	20	Depressed affect, Positive affect, Somantic, Interpersonal (4 subscales)	Community, general Adults, clinical settings	Good (4 subscales, 20 features)
MDI	10-12	Depression, Severity grading	Community, general Adults, clinical settings	Low to Moderate (richer per item scale)
K10	10	General psychological distress	Population–level mental health surveys	Low (single broad construct, 10 features)

Table 2.1: Depression Screening Instruments Comparison

Based on the comparative analysis presented in Table 2.1, the DASS-21 would be best suited for inclusion in a machine learning pipeline of university student populations, as it offers the three highly discriminant subscales, an empirically sound application to a diverse range of populations including African university students and a completely free use license.

2.4 The DASS-21 Scale

DASS-21 stands for Depression Anxiety Stress Scale-21 which was constructed by Lovibond & Lovibond in 1995 and is the abbreviated form of its parent scale, the DASS-42. The scale comprises twenty one items that are further categorized into three different seven items sub scales representing the constructs of depression, anxiety and stress. These items are scored using a four point Likert type response format where respondents indicate how they have experienced these symptoms over the past week by selecting one of the following scores: 0 = Did not apply to me at all, 1 = Applied to me some of the time, 2 = Applied to me much of the time, 3 = Applied to me very much or most of the time.

The depressive subscale refers to absence of positive feelings, anhedonia, hopelessness, and meaninglessness, while the anxiety subscale measures arousal, situation-based anxiety, and fear. The stress subscale denotes non-specific arousal and includes symptoms like irritability, nervousness, and inability to relax. The cut-off scores established by Lovibond and Lovibond (1995) include Normal, Mild, Moderate, Severe, and Extremely Severe states of functionality. Based on the findings of a cross-sectional study carried out in February 2024 involving 1,697 college students from Bangladesh using DASS-21, the prevalence of depression, anxiety, and stress is estimated to be at 56.9%, 69.5%, and 32.2% respectively (Islam et al., 2025).

The DASS-21 test is highly reliable when applied across various linguistic and cultural settings. The test has been validated through several investigations conducted on university students from different countries in Africa and Asia to determine the factorial validity and internal consistency of the instrument, where alpha values have been greater than 0.80 across all subscales (Henry and Crawford, 2005). Due to its excellent psychometric characteristics, brevity, and open availability, the DASS-21 test is widely used in screening studies.

Figure 2.1 illustrates the structural organisation of the DASS-21 across its three subscales, indicating the item numbers assigned to each domain and the scoring range.

2.5 Machine Learning in Mental Health Prediction

The use of machine learning in predicting and classifying various mental illnesses has increased significantly in recent years due to developments in computer capacity, mental health data sources, and recognition of the effectiveness of ML for clinical diagnostics. According to a meta-analysis conducted on 48 papers selected through strict inclusion standards over the period from 2011 to January 2024, the efficacy of machine learning algorithms in the detection of depression, anxiety, and stress symptoms was explored among university students enrolled in an undergraduate program (Rodrigues et al., 2024). Among 33 studies providing measures of accuracy, 30 indicated that this metric was above 70%, with depression being detected at accuracies between 65% and 99%.

Some of the widely used algorithms for this purpose include Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Trees, Random Forest, Gradient Boosting, and in more recent times, machine learning models such as Convolutional Neural Networks and Recurrent Neural Networks. In their systematic review of the use of machine learning in mental health disorders, including 28 papers on various mental illnesses, Shatte et al. (2019) discovered that the algorithm models which have shown superior performance compared to other methods of machine learning are ensembles – Random Forest and Gradient Boosting, due to the fact that ensembles inherently have low variance because of combining many weak models.

A notable research article in Frontiers in Psychiatry conducted a study using 1,635 college students from China, where 38 sociodemographic, psychological, and social variables were considered, and four machine learning algorithms were compared, namely, Random Forest, XGBoost, LightGBM, and SVM (Yu et al., 2025). Among all four models, Random Forest performed best in discrimination with AUC and accuracy scores being 0.87 and 0.79, respectively. Furthermore, sleep disturbance, perceived stress, experiential avoidance, and self-criticism were found to be important predictors of depression among students. Moreover, SHapley Additive exPlanations analysis showed that worsening sleep quality and increasing stress levels increase the likelihood of depression, which is in line with adding sleep quality and academic stress as additional predictors in this study.

An experiment based on data collected from 17 universities across Southeast Asia involving the use of eight different ML models for predicting mental well-being revealed that Random Forest and adaptive boosting algorithms had the highest level of accuracy. Five of the features that proved to have the most importance when it comes to prediction of poor mental well-being include body mass index, sport participation frequency, grade point average, hours of sedentary behavior, and age (Chow et al., 2023).

2.6 The Random Forest Classifier

Random Forest is an ensemble learning method for supervised learning that uses multiple decision trees generated in training and then combines the decisions of all the trees through majority voting or averaging, depending on whether we are dealing with a classification or regression problem. The key idea behind the random forest algorithm is that the combination of several not necessarily perfect but diverse models is better than one model because their mistakes tend to neutralize each other out.

Every individual tree in a Random Forest algorithm uses a bootstrap sample of the training dataset for its training process, a methodology referred to as bagging or bootstrap aggregating. Besides, the splitting of nodes in any given tree involves the selection of a subset of randomly chosen features at each particular node. The randomness of features is what makes Random Forest different from the basic bagging procedure and is what mainly enables generalization while preventing the risk of overfitting.

There are numerous benefits of using Random Forest that make it an appropriate algorithm for the depression prediction problem. Firstly, the algorithm is tolerant to missing data and outliers that might be present in user-provided surveys. Secondly, it supports both categorical and numerical input variables, allowing us to avoid unnecessary feature engineering. Moreover, the Random Forest algorithm provides feature importance out of the box, providing valuable information about what DASS-21 questionnaire items and lifestyle factors have the highest contribution to predicting the risk of depression — something of high clinical value. Lastly, the algorithm works effectively on tabular data without much hyperparameter optimization. In the comparison of SVM, Logistic Regression, and Random Forest for predicting depression in Korean college students, the latter was shown to achieve the best predictive performance, recognizing self-reported mental health, neuroticism, and family cohesion as the most important factors (Kim et al., 2022).

2.7 Web-Based Mental Health Systems

One of the most important changes in the digital health field in recent years is the growing use of web-based mental health applications. These can vary from straightforward psychoeducation programs to advanced technologies that employ AI to screen and monitor. The main benefits of these web-based services include the geographical availability of the platform to users in distant locations, access to those who might avoid seeking mental help because of stigma, and the capability of collecting real-time data and instant feedback.

The effectiveness of web-based interventions in mental health, aimed at alleviating the symptoms of depression, has been supported by several pieces of research. Andersson & Cuijpers, in their meta-analysis of twelve randomised controlled trials, concluded that computer-based cognitive behavioral therapy for depression is just as effective as therapist-assisted therapy for moderate-to-mild cases, and the overall effect size is equal to 0.78. Although the current experiment does not involve any therapy, the existing body of evidence proves the possibility of using technology to get users engaged in health-related activities.

As far as college settings are concerned, online assessment tests have proven to enhance help-seeking behavior among students that might not otherwise interact with mental health professionals on their college campuses. Designing an interface that is aesthetically pleasing and mobile-friendly can be considered an important step in designing a digital application due to the fact that poor user experience can be considered a major barrier in using such a digital health solution. React.js, which is the frontend tool chosen in this current study, allows for developing highly interactive applications through components.

2.8 Review of Related Literature

There is increasing literature on the use of machine learning techniques to detect and predict depression as well as other mental disorders among students. This subsection will discuss the works that are highly pertinent in the literature, focusing on their methodology as well as the gaps that have been addressed by this study.

Islam et al. (2025) carried out a cross-sectional study involving 1,697 students enrolled at universities in Bangladesh through the DASS-21 questionnaire and used seven ML algorithms such as KNN, Random Forest, Gradient Boosting, XGBoost, CatBoost, Logistic Regression, and SVM in predicting depression, anxiety, and stress. The results indicated the prevalence of depression, anxiety, and stress among participants at 56.9%, 69.5%, and 32.2%, respectively. In the tested algorithms, ensemble models performed better than individual models, and the most accurate models for prediction were Random Forest and XGBoost. However, the study focused only on the university students from Bangladesh without considering any other lifestyle characteristics.

A systematic review by Rodrigues et al. (2024) was conducted on the application of ML algorithms in detecting depression, anxiety, and stress in undergraduate students. A search on PubMed, Embase, PsycINFO, and Web of Science databases yielded 2,304 studies after screening, of which 48 met the inclusion criteria. In this review, behavioral data, physiological data, internet use data, neurocerebral data, and demographic data were found as the feature sets for each study considered. Thirty out of the 33 studies that used accuracy measures showed scores above 70% of ML-based detection of mental health issues. There is a clear lack of research from Africa regarding this topic.

Aladwani et al. (2024) developed an innovative machine learning-based algorithmic approach for depression level identification among students on the basis of DASS-21 score and sociodemographic factors using an ensemble approach known as EDUSTACK-MH, which could predict the depression severity level among five categories ranging from Normal to Extremely Severe. This experiment used data from 1,120 university students, and upon conducting a comparative evaluation of baseline and ensemble methods, it was found that the optimized stacking-based model performed with 98.23% accuracy. Thus, the dataset used in the current work has been used before.

The study by Chow et al. (2023), based on a cross-sectional survey of 17 universities in Southeast Asia, employed a combination of eight machine learning (ML) algorithms for modeling the dependent variable of mental well-being. Among all ML algorithms tested, Random Forest was found to be the best predictor of the presence of negative mental well-being attributes, where GPA, physical activity level, and sedentary behavior were among the top features contributing towards the predictive model. This finding supports the theory behind selecting these variables as additional predictors in the current study.

Yu et al. (2025) obtained data on 1,635 Chinese college students from 38 sociodemographic, psychological, and social variables and compared the predictive accuracy of four machine learning algorithms to predict depression symptoms. Random Forest outperformed other models, providing the highest AUC of 0.87 and accuracy of 0.79 by recognizing sleep disturbance and perceived stress as the major predictors. This research used the SHAP method to gain an interpretation of the importance of features in the machine learning prediction algorithm, providing a blueprint for future extensions of the current study.

Kim et al. (2022) analyzed depression risks among 171 Korean college students based on 171 family and individual data by utilizing SVM, Logistic Regression, and Random Forest algorithms. The best classification results were provided by the Random Forest algorithm with self-perceived mental health, neuroticism, and fearful-avoidant attachment recognized as the top five predictors. Despite the low number of participants in this research, it may serve as a good comparative base for the analysis of feature importance in this paper.

Figure 2.2 presents a tabular summary of the reviewed literature, highlighting the techniques applied, datasets used, and the research gaps identified in each study.

S/N	Author(s)	Work Done	Technique	Research Gap
1	Islam et al. (2025)	Applied seven ML models to predict depression, anxiety and stress from DASS-21 data among 1,697 Bangladeshi university students.	KNN, RF, XGBoost, GBM, CatBoost, LR, SVM	Limited to Bangla context; no lifestyl behavioural fe included beyond DAS items.

2	Rodrigues et al. (2024)	Systematic review of ML model performance for detecting depression, anxiety and stress among 48 included studies.	Systematic review (multiple algorithms)	No primary data coll significant gap note African university st populations.
3	Aladwani et al. (2024)	Developed EDUSTACK-MH ensemble framework using DASS-21 data from 1,120 students for five-level depression severity classification.	Stacking ensemble, KNN, SVC, RF, GB, LR	No web-based deploy no lifestyle features; l to Indian univ students.
4	Chow et al. (2023)	Used multi-site survey data from 17 Southeast Asian universities to model mental well-being with eight ML algorithms.	RF, GLM, KNN, Naïve Bayes, Neural Network, Bagging, Boosting	No DASS-21 instr used; no web-based s developed; Sou Asian context may generalisability.
5	Yu et al. (2025)	Predicted depressive symptoms from 38 features in 1,635 Chinese college students using four ML algorithms with SHAP analysis.	RF, XGBoost, LightGBM, SVM	No web-based applic no DASS-21 sub features; Chinese st context may transferability.
6	Kim et al. (2022)	Compared RF, SVM and LR for predicting depression in 171 Korean college students using family and individual data.	RF, SVM, Logistic Regression	Very small sample digital deployment; DASS-21 instru Korean context.
7	Ladi-Akinyemi et al (2023)	Investigated the prevalence of	Cross-Sectional survey; Descriptive	The study was purely observational and

		depression and suicidal ideation among undergraduates in tertiary institutions in Lagos, Nigeria, using cross-sectional survey methodology	and inferential statistics; PHQ-9; Logistic Regression	descriptive with no mac learning model develop PHQ-9 provided only n single-scale features. N automated or web-base screening tool was produced, and no lifest feature integration was explored.
8	Li et al. (2022)	Conducted a systematic review and meta-analysis on the prevalence and associated factors of depression and anxiety among college students, synthesising 64 studies with a total sample of 100,187 individuals. Reported a pooled depression prevalence of 33.6% globally, with substantially higher rates in Africa and lower-middle-income countries. Identified academic workload, social isolation, and inadequate coping mechanisms as primary risk determinants.	Systematic review and meta-analysis; Pooled prevalence estimation; Subgroup and regression analyses	Being a meta-analytic r no primary data colle predictive model, screening tool was prod Machine le approaches for detection were not exp African and spec Nigerian sub-popu were not isolated for s ML-applicable analysis
9	**Wang et al. (2023)**	Conducted a systematic review and meta-analysis covering 350,000 college students across 373 campuses between 2013 and 2021. Reported that the proportion of students with at least one mental	Systematic review and meta-analysis; Longitudinal trend analysis; Mental health prevalence estimation	No predictive ML mod developed. The revie not include Africa Nigerian uni populations and dic produce a deplo screening tool. pandemic contextual f

		health problem exceeded 60%, with nearly a 50% rise over the study period. Anxiety and depression were the most prevalent conditions identified, exacerbated by the COVID-19 pandemic conditions.		may limit the generaliz of prevalence estima non-COVID settings.
10	**Aleem et al. (2022)**	Conducted a review of ML and deep learning algorithms specifically applied to diagnosing depression. Analyzed the performance of algorithms including Logistic Regression, SVM, Random Forest, CNN, and RNN across multiple depression-relevant datasets. Highlighted that ensemble methods — particularly Random Forest and Gradient Boosting — consistently outperformed single classifiers in classification accuracy and AUC metrics.	Systematic review; Random Forest, SVM, Logistic Regression, CNN, RNN, Gradient Boosting; Multiple depression datasets	The review did not deve deploy a working sy African university popu were absent from the s reviewed. DASS-21 primary screening instr and lifestyle f integration were specifically evaluated.
11	**Saleh et al. (2023)**	Reviewed machine learning and deep learning approaches for diagnosing and monitoring mental health disorders across 28 included studies. Identified that ensemble	Systematic review; Random Forest, SVM, CNN, LSTM, Gradient Boosting; Questionnaire-based and physiological signal datasets	No primary datase deployable system developed. The revie not address the Ni university student con the integration of DA with lifestyle features. learning models rev

		methods show superior performance compared to single classifiers in mental health prediction tasks. Discussed applications spanning depression, anxiety, bipolar disorder, and schizophrenia using data from questionnaires, physiological signals, and social media behaviour.		require large la datasets unavailable resource-limited setting
12	**Nwaoha et al. (2022)**	Investigated the prevalence, symptom severity, and correlates of depression among students at a Nigerian university in the South-Eastern geopolitical zone. Utilised cross-sectional survey methodology and identified similarly high depression rates among the student population, consistent with figures previously reported by Adewuya et al. (2006). Academic stress, financial difficulty, and limited access to mental health services were identified as significant contributing factors.	Cross-sectional survey; Descriptive statistics; PHQ-based symptom screening; Logistic Regression	Study was descriptive did not develop any based predictive mod digital screening tool. based screening pr limited single-scale fe insufficient for nuance classification. No li features were integrate study reinforces the ne an automated, sc screening system in Ni universities — a gap c addressed by the c study.

Figure 2.2: Tabular Summary of Related Works

2.9 Research Gap

The study examined various worries and problems related to Depression Risk Prediction System for University Students and the following gaps were observed.

The critical analysis of the reviewed literature suggests that there are several recurring gaps which the current study attempts to fill. First, there is an obvious gap in terms of the geographic location of the majority of the studies related to machine learning applications for predicting the likelihood of depression among university students, with the majority being focused on Asian and Western students with virtually no studies being carried out with regard to African universities and none of them involving Nigerian university students specifically using the DASS-21 test. Secondly, there has been quite a substantial number of papers related to building and evaluating machine learning algorithms, but very few studies have moved to implement these algorithms into a working web application that could be used by end-users. Moreover, of those few systems where such implementation has been attempted, almost all of them have used PHQ-9 as the diagnostic instrument and not provided the necessary functions for calculating depression risk, as seen from Table 2.1. Finally, there are virtually no studies where the use of DASS-21 combined with additional lifes tyle factors (quality of sleep, physical exercise, level of social interaction, appetite changes, and academic stress) were used to predict depression.

CHAPTER THREE

SYSTEM ANALYSIS AND DESIGN

3.1 Introduction

This chapter provides a comprehensive overview of the design and analysis of the Machine Learning-Based Depression Risk Prediction System for University Students (MindCheck). This is achieved through an analysis of the existing process for the detection of depression among university students, especially considering the numerous difficulties faced in such processes, then designing the new process. The chapter is very well organized to facilitate a gradual understanding of all the issues involved.

3.2 System Analysis

System analysis in this case would entail carrying out an analysis of all the functional and non-functional requirements that are necessary for developing an intelligent depression risk predictor website. With regards to functionality, the system should be capable of collecting data in relation to the psychometric scales together with contextual information from university students, passing the information to an inference engine for analysis, classifying the individual into whether he/she is at high risk or low risk of being depressed, and recommending solutions accordingly.

3.2.1 Analysis of the Existing System

Current techniques for the diagnosis of depression in the university include an interview with a psychologist and self-administered paper questionnaire. One of the most widely used self-assessment questionnaires at Nigerian universities is PHQ-9 questionnaire consisting of nine questions about the symptoms of depression and providing a score as an output. Despite being clinically validated, PHQ-9 questionnaire lacks a large dimensional space of the input data, since there are nine input variables only. Current depression diagnostic systems do not use any machine learning technique for the result output. Thus, the score given by the system should be manually evaluated by a psychologist.

A number of web-based depression tests has been developed all around the world, such as MoodGym, Beyond Blue, and web counseling portals of particular universities. Nevertheless, these tests have been developed for Western societies and are not validated for Nigerian students yet. Moreover, there is no machine learning algorithm working in real time and the result depends on a certain threshold score.

3.2.2 Weaknesses of the Existing System

Some of the key shortcomings of the currently existing solutions for depression screening at Nigerian universities are:

13. The current systems are heavily dependent on trained clinical personnel, which makes it difficult to implement them on a large scale among a large number of students.
14. The nine questions format of the PHQ-9 makes the feature space less diverse for machine learning algorithms.
15. Paper-based and face-to-face interviews act as a deterrent for some students who may be afraid of stigmatization when asking for mental health screening.
16. Subjective manual score evaluation cannot be as consistent as an algorithmic one.
17. There is no system being used in Nigerian universities that provides real-time personalized advice based on a risk assessment result.
18. Lifestyle features are not collected by any of the systems used in Nigerian universities.

3.2.3 Analysis of the Proposed System

It is possible to describe this approach as smart and Internet-based which would help to overcome all the problems discussed above. Firstly, it includes obtaining the information from the two aspects: first, the whole DASS-21 questionnaire consisting of 21 questions separated into three groups; secondly, five additional parameters concerning his/her lifestyle. After that, this information is transferred to the API server via HTTP POST request where it is processed through normalization according to the StandardScaler algorithm already fitted before and then passed to the Random Forest classifier returning the risk category and its probability (low, medium or high).

Unlike the existing solutions, the proposed one would include obtaining the information both about the results of the clinical scale (scores of the DASS-21 scales) and the life context parameters at once which would make the task of a machine learning algorithm much more difficult due to 26 features (21 features obtained from the DASS-21 scales and 5 life context features) instead of 9 obtained via the existing solutions using PHQ-9. This solution is completely Internet-based so no direct interaction with people is needed during the process of the test.

Here's how the proposed system will work. In this scenario, the user will be a university student visiting the web application. They will encounter a landing page where the purpose and limits of the application are explained to them and they will be prompted to initiate the test.

Following that, they will need to complete the Assessment Form by clicking on the “Start Assessment” button.

It has four steps, and the initial three steps will have them answer the 21 questions of the DASS-21 in an order corresponding to the three subscales, Depression, Anxiety, and Stress.

After submission, all 26 collected features will be assembled into a JSON object and transmitted to the Flask backend API using an HTTP POST call at the '/predict' endpoint. The API will take in the data, execute a series of validation procedures on the incoming information, and then scale the features using a pre-trained StandardScaler, ensuring they achieve a mean of zero and a standard deviation of one. Subsequently, these scaled features are delivered to a random forest classification model which ascertains the risk category (Low, Moderate, or High) and provides a corresponding probability score. The assessed risk category will then be associated with predefined recommendations, and the entire outcome will be dispatched as a JSON object.

The front end developed using React.js interprets the API response JSON to render the Results page which displays the prediction of classification risk in the respective colors (Green - Low Risk, Orange - Moderate Risk, Red - High Risk), the confidence score produced by the model, the subscale results of the DASS-21 questionnaire, the lifestyle indicators summary, and all the recommendations formulated specifically for the given risk class of the student. The Results page includes a disclaimer indicating that the current system is a preliminary mental disorder screening instrument and not a clinical diagnosis, and a "Retake Assessment" button to reset the form to start a new assessment. The system collects no personal identifiable information throughout the assessment process.

Figure 3.1 illustrates the block diagram of the proposed system, showing the flow from user input through the web interface, API, ML model, and back to the results display.

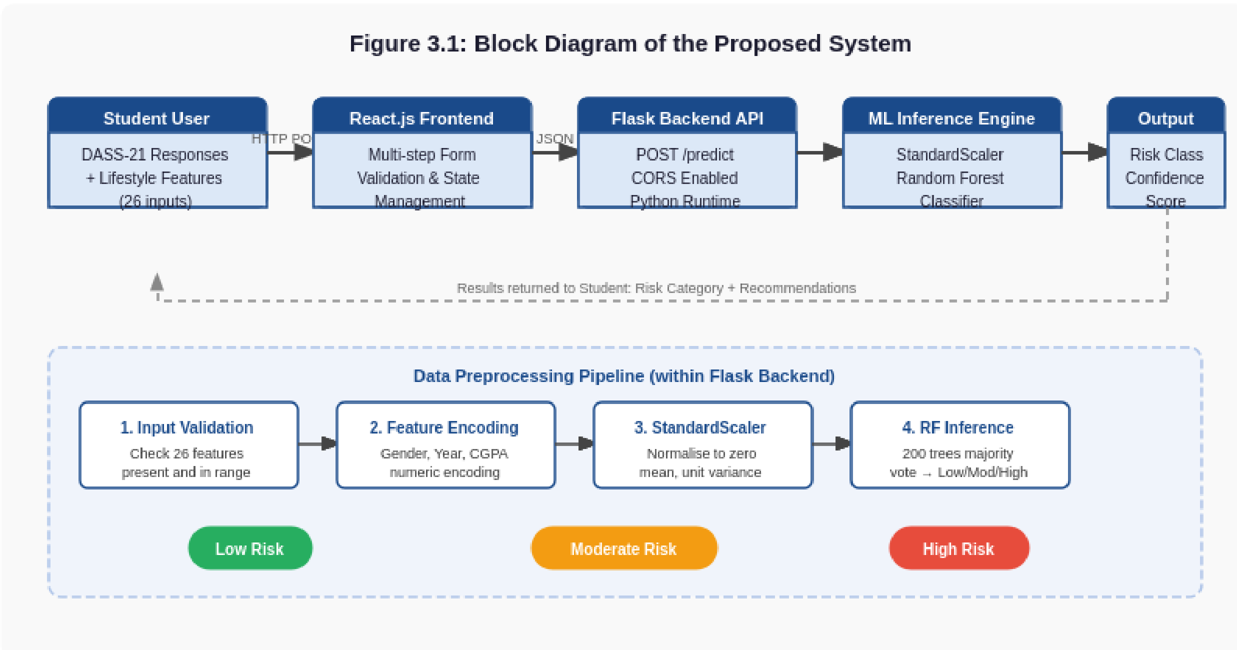


Figure 3.1: Block Diagram of the Proposed System

3.2.4 Advantages of the Proposed System

19. Scalability The web-based nature of the application enables the system to scale to any size, allowing it to support large numbers of students without the need to hire additional clinical staff.
20. Robust Feature Set By using the three subscale scores from the DASS-21 along with five indicators of lifestyle, the system generates a 26-dimensional input vector. This provides for much richer prediction of risk than a single-scale tool.

21. Real-Time PredictionsThe Flask API can provide the classification within milliseconds of receiving the feature vector, leading to an efficient and fluid user experience.
22. Anonymous Data collection No identifying data such as matriculation number or names are ever collected by the application. This helps to break down some of the stigmatic barriers to use.
23. Individualised feedbackThe results page provides targeted, contextualized advice for users based on their identified risk level, giving them actionable information.
24. Complementary to clinical workThe system is intended as a complementary first-pass screening tool and is not meant to replace the clinical work of professional mental health practitioners.

3.2.5 Proposed Model Architecture

The Random Forest algorithm has been chosen from the four listed algorithms of; Logistic regression, SVM with RBF kernel, KNN classifiers to train a model. The chosen algorithm was coded to be built with certain hyper parameters such as, the number of estimators of 200, max_depth as None, 2 minimum sample split and Gini criterion. It has been assumed to be the maximum number of features when growing trees with sqrt of the number of features in training.

Then this model was trained with the oversampled training dataset (using SMOTE), and tested on the withheld test set.

The model was used to calculate the feature importance among the training features DASS-21 and lifestyle factor features, then pickled as 'model.pkl'.

Figure 3.2 illustrates the architecture of the Random Forest classifier, showing the flow from input feature vector through the ensemble of decision trees to the majority-vote classification output.

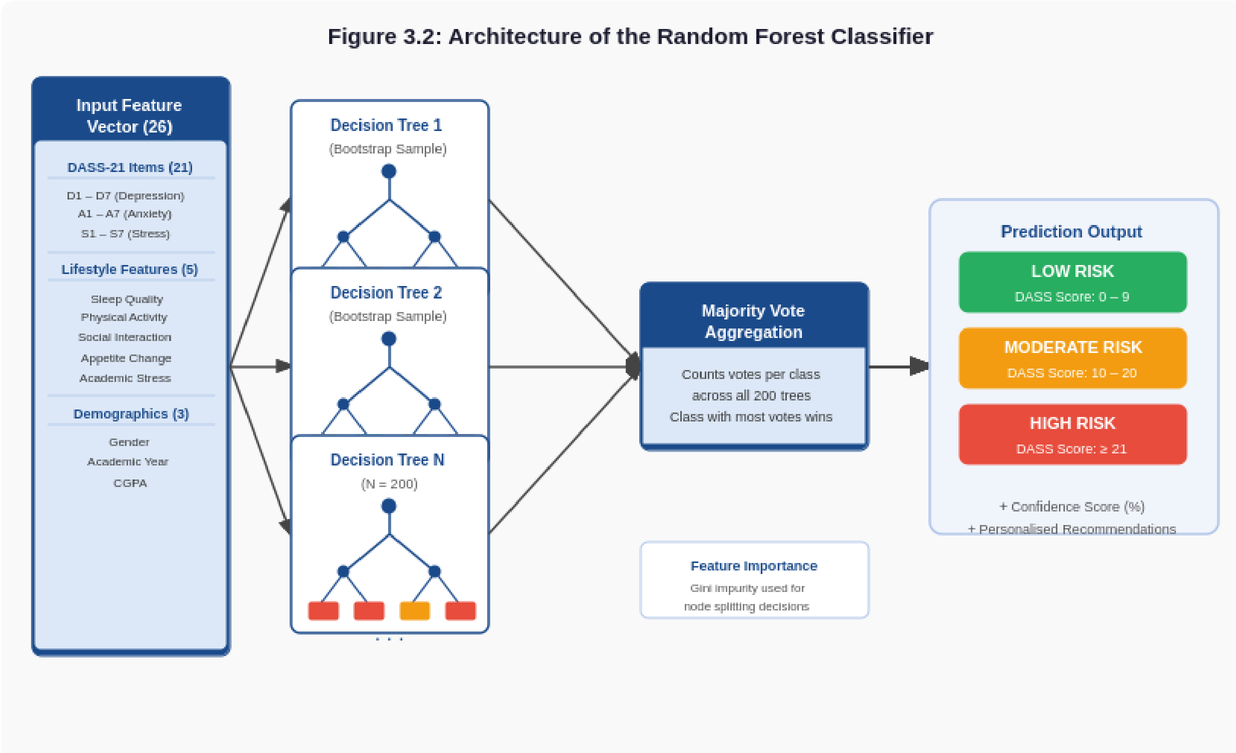


Figure 3.2: Architecture of the Random Forest Classifier

3.3 Research Methodology

To implement the web-based intelligent system for depression prediction in an agile manner, Agile software development methodology would be a good option. Agile suits the context quite well as it requires repeated testing and flexibility. Both testing and flexibility is highly desirable in the context of the newest technology.

The research will span over a series of development sprints where particular parts of the application will be developed.

Sprint #1 includes requirement elicitation, preparation of dataset, and data exploration so to understand characteristics of features within the DASS-21 questionnaire. The second sprint focuses on data preprocessing, feature engineering and balancing the data set by using SMOTE technique. The third sprint will deal with building machine learning models and validating their efficiency. The final two sprints #4 and #5 will include building of backend and front-end part of the system respectively.

The rationale behind selecting the Agile method is due to its applicability to research and development where there is a high likelihood of requirements shifting as further knowledge is acquired over the development and testing phase. In contrast to the linear and rigidly phased nature of Waterfall methodology, Agile enables the research process to evolve as fresh knowledge is acquired thereby ensuring that machine learning algorithm and the web application continuously improve.

Sprint 1: Literature Review and Data Exploration; Literature review has been conducted to compare prevalence of depression among students and the tools for measuring depression level among students. Data available publicly on student mental health including 1120 sample has been downloaded and inspected for attributes, missing values and class imbalance of Low, Moderate and High depression classes.

Sprint 2: Data preprocessing and Feature engineering All necessary data preprocessing was performed to get the data in shape for modeling. The data had few missing values and they were handled by dropping the rows with missing values in DASS-21 scores and substituting the median for the concerned lifestyle variable. Two variables were converted to numeric variables. Three classes denoting the risk of having depression was generated from DASS-21 depression score using thecutoff criteria as given by Lovibond & Lovibond(1995). The generated three classes were balanced using Synthetic Minority Oversampling Technique but on train set alone, not to leak information into test and validation sets. The validation and test data was scaled by fitting a StandardScaler on the train data.

Sprint 3: Training & evaluation of model Four classification ML models i.e., Logistic Regression, Support Vector Machine (RBF Kernel), K-Nearest Neighbors and Random Forest were trained and compared in Sprint 3. TheSMOTE balanced train set was used to perform stratified 5-fold cross validation, and the accuracy of all the ML models was calculated using Accuracy, Precision, Recall, F1-score and AUC-ROC in the independent test set that comprised of 20% of the data. The best performed model was Random Forest, and feature importance was calculated.

Sprint 4: Flask Backend Development - The only feature developed for this sprint is a Flask backend application with only one endpoint namely POST /predict, which takes in the 26-dimensional JSON feature vector from the frontend, preprocesses it by fitting it on the already fitted StandardScaler and passing on the transformed vector to the Random Forest model, then returns a JSON response with the risk prediction, the percentage of confidence and the recommended actions. CORS(Cross-origin resource sharing) were enabled for backend-frontend interaction. This was tested using a few payload data values using Postman tool for each of the three predicted risk class before proceeding to the frontend.

Sprint 5 - Development of the React.js Single-Page Application and its Integration into the system : The fifth sprint in the project lifecycle consisted of the development of the full React.js application, which contained the three pages namely Landing Page, the multi-step DASS-21 and Lifestyle Assessment Form, and the Results Page. We used the Axios HTTP client in order to send the constructed 26 feature JSON file to the Flask API and to then receive the result of the prediction from the backend to be then displayed on the results page. We also conducted client-side form validation in order to avoid the submission of partially filled forms.

TailwindCSS was utilized for designing a clean and user-friendly mobile compatible website.

This completed the client-side application which was then integrated with the server-side application which in turn was finally tested end to end using the 5 test profiles detailed in Section 4.3 above.

3.3.1 Proposed Methodology

Research Methodology This research will adopt a Machine Learning-Based Depression Risk Prediction Methodology in the approach, by the application of Agile Development methodologies; through which Psychometric Assessment, data pre-processing, machine learning model building, web application development and evaluation is conducted in order to develop an intelligent depression risk prediction system for university students. The data collection phase consists of collecting a publicly available dataset containing information of student mental health including responses to Depression Anxiety Stress Scale (DASS-21) test with lifestyle features. Added are five additional features namely Sleep quality, Exercise patterns, Number of social interactions, Changes in eating patterns, and Academic stress. After collection, the next stage is the pre-processing of data for enhanced data quality before proceeding with model training, where handling of missing values is carried out; Categorical features are encoded to numeric form, DASS-21 depression risk scores are derived from the DASS depression cutoff scores, SMOTE is employed for the handling of class imbalance, The dataset is further split into 80% train set and 20% test set, and features are normalized using StandardScaler.

Machine Learning Pipeline is carried out in terms of different classification models such as Logistic Regression, Support Vector Machine(SVM), K-Nearest Neighbour(KNN), Random Forest. The models are experimented by checking Accuracy, Precision, Recall, F1-Score, AUC-

ROC values. Upon experimenting and comparing different models, the Random Forest has been selected to develop as the final prediction model as it has a higher classification accuracy and better resist against Overfitting.

A Random Forest Machine Learning model was then trained to predict risk of depression, which was subsequently exposed as an online prediction service via Python Flask RESTful API.

The API accepted the responses from the users for an assessment, perform preprocessing of the response via the saved StandardScaler and provided prediction of risk of depression along with its prediction confidence in JavaScript Object Notation (JSON) format.

The front-end of the system utilizes React.js to create a responsive and user friendly interface, where students will undergo the DASS-21 questionnaire with multiple stages, then their answers would be sent securely to the Flask backend. The prediction results will be generated instantly along with the prediction scores, the DASS-21 subscale summaries and recommendations based on the level of prediction, while keeping all data encrypted. Finally, a functional testing, integration testing, and user acceptance testing was performed on the designed system to test whether all parts of the system work as intended and fulfill the system requirements. Overall, this research approach ensures that a scalable, accurate, and accessible tool is developed for initial detection of potential risks of depression within the student population, to promote the facilitation of earlier diagnosis and treatment in cases of severity.

3.4 Data Collection

The training dataset for the development and assessment of the machine learning models, composed of 1120 instances collected from university students using the DASS-21 questionnaire and multiple lifestyle and demographic attributes. In detail, the dataset was produced from a public Kaggle dataset of students' mental health data and enriched with additional synthetic samples through the SMOTE method to tackle class imbalance. Each sample within the dataset comprises 21 items on the DASS-21 scale (where the possible values range between 0 and 3), the sums of the three scales (i.e. Depression, anxiety and stress), 5 lifestyle variables (where the possible values range between 1 and 5) and a target label that represents the risk class of depression.

3.4.1 Data Description

Depression risk groups have been defined on the depression scale of DASS-21 using criteria set by Lovibond and Lovibond (1995) which categorises low risk to be score of 0–9, moderate risk score of 10–20 and high risk score of above 21. Other demographic variable details collected for analyses include Gender, Year and CGPA. After applying SMOTE, the class distributions for training data are shown below in Table 3.3 (with the minority classes, i.e., moderate and high risks having been oversampled) with the data is balanced in all the classes.

Risk Class	Before SMOTE	After SMOTE
Low	653 (58.3%)	650 (33.3%)
Moderate	312 (27.9%)	650 (33.3%)
High	155 (13.8%)	650 (33.3%)

Table 3.1: Dataset Class Distribution Before and After SMOTE

3.4.2 Data Preprocessing

This stage is very important as it aims to ensure the input data for the model training is clean, consistent, and suitable for the model. In this preprocessing process, the following steps were performed consecutively on the dataset in its raw state:

25. Missing Value Imputation: Any row that contains missing values in DASS-21 question columns are dropped. For the features related to lifestyle, missing values are imputed using the median value of that feature in the dataset in order to keep as many sample points as possible.
26. Feature Encoding: The gender feature is transformed into a binary form, with 0 represented for male and 1 represented for female. Year of study is transformed into an ordinal form as an integer ranging from 1 to 5. CGPA is treated as a numeric continuous feature and left unchanged
27. Class Label Creation: A new target feature “depression_risk” was created and sourced from the DASS-21 depression sub-scale total score. The classes for this target feature, which are Low, Moderate, and High depression risk classes, were determined by using the cut-off thresholds described by Lovibond and Lovibond (1995).
28. Oversampling using SMOTE: Since the “depression_risk” classes are imbalanced, Synthetic Minority Oversampling Technique, abbreviated as SMOTE, is applied to the

training dataset to create synthetic samples for the Moderate and High risk classes. This leads to equal representation for all three classes. This technique is only applied to the training set to avoid leaking information to the validation and test sets.

29. Feature Scaling: All numerical and continuous features are standardized to have zero mean and unit variance by using the StandardScaler from the Scikit-learn library. The StandardScaler is fit on the training data and then applied to both the validation and test data sets. This avoids data leakage from training data.
30. Train-Test Split: Once preprocessed, the dataset is split into an 80% training set and a 20% test set using stratified splitting, so that the class proportions remain similar in both subsets.

3.4.3 Feature Engineering

The vector of features supplied to the Random Forest model consists of 26 features which were derived from two different areas. First, the 21 responses provided by the student to the DASS-21 test were included as individual features (i.e. They were not summed into the subscales) in order to give as much information as possible to the classification model. The other five features are lifestyle features recorded by the student during the assessment procedure, as given in table 3.2 below.

Feature	Description	Scale
Sleep Quality	How the student rates their daily sleep quality over the past two weeks.	1 (Very Poor) to 5 (Excellent)
Physical Activity Level	Amount and frequency of physical exercise or activity per week.	1 (Never) to 5 (Daily)
Social Interaction Frequency	How frequently the student engages in social activities with peers.	1 (Rarely/Never) to 5 (Very Frequently)
Appetite Change	Whether the student has noticed significant changes in their appetite over the past two weeks.	0 (No Change) to 2 (Significant Change)
Perceived Academic Stress	The student's self-assessed level of academic pressure and workload stress.	1 (Very Low) to 5 (Very High)

Table 3.2: Additional Lifestyle Features

3.5 System Architecture

Figure 3.3 illustrates the overall architecture of the system, which uses a client–server structure consisting of three main components: the client (React.js), the server (Python Flask) and the machine learning inference engine. React.js (client) single page application This part is developed using React.js framework and consists of three main components: • Landing Page (provides information about the system, its function etc.) • Assessment Form (is designed as a multi-step form for a user to answer the 21 items of DASS-21 along with 5 rates regarding the person lifestyle.)

• Results Page (showcases a risk category (predicted class), confidence (probability) and recommended guidelines for improvement.)

React Router library is used for managing view transitions, and React's useState Hook is used for handling form states. Flask backend The backend provides a single RESTful API exposed at POST /predict. Once the backend receives a JSON payload of the 26 values of the features, it first loads the pre-fitted StandardScaler object and transforms the given feature vector before loading the trained Random Forest model to predict the risk class and probability. Finally it returns the risk class, probability as a confidence level, and a list of recommendations in JSON.

For this project, cross-origin requests were permitted from React development server by enabling Cross-Origin Resource Sharing (CORS) in the backend.

Figure 3.3: System Architecture Diagram

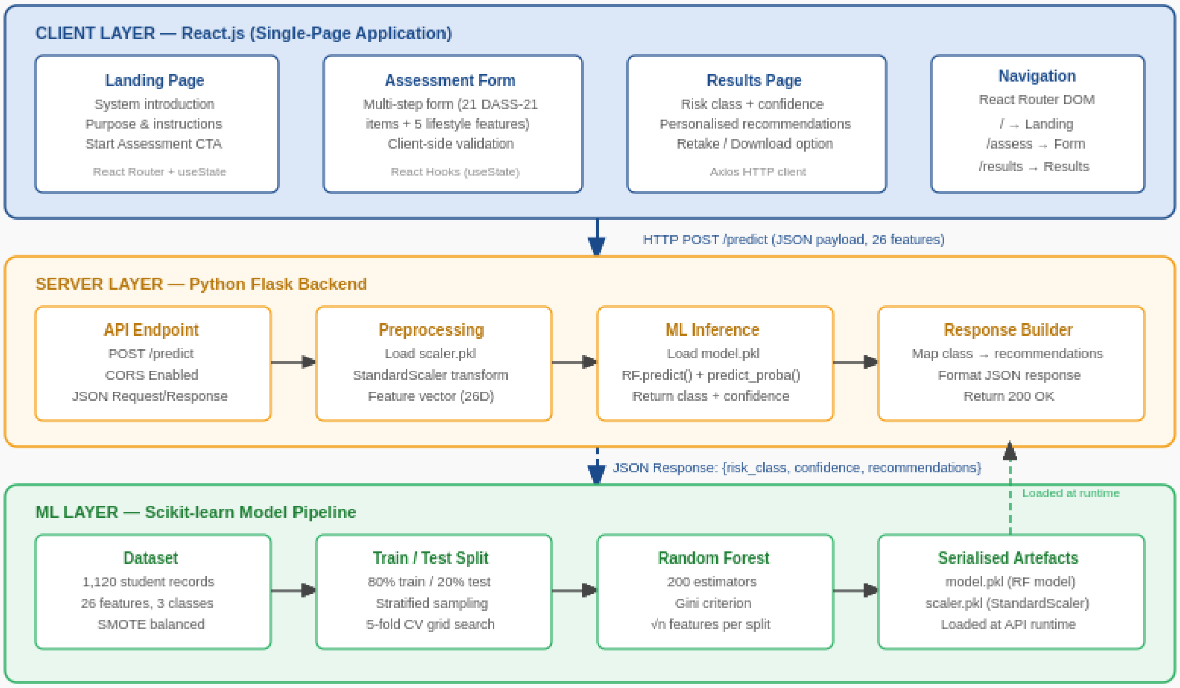


Figure 3.3: System Architecture Diagram

3.5.1 System Flowchart

As shown in Figure 3.4 is the processflowchart of the sequence of actions the student performed, from their first visit on the web app to them receiving a personalised output. First, the student first loads the landing page. From then they progress to the quiz form and fill in all of the DASS-21 questionnaire items, as well as the lifestyle questions.

Once they submit the quiz form the web app performs client-side validation for every field and passes the data over to the Flask API.

The API then takes the input data, processes and scales it, predicts the output using the model and outputs it. The result page then loads on the web app with the prediction including the classification and confidence score, along with suggestions to the student. If the student wishes to take the quiz again, the form would be reset.

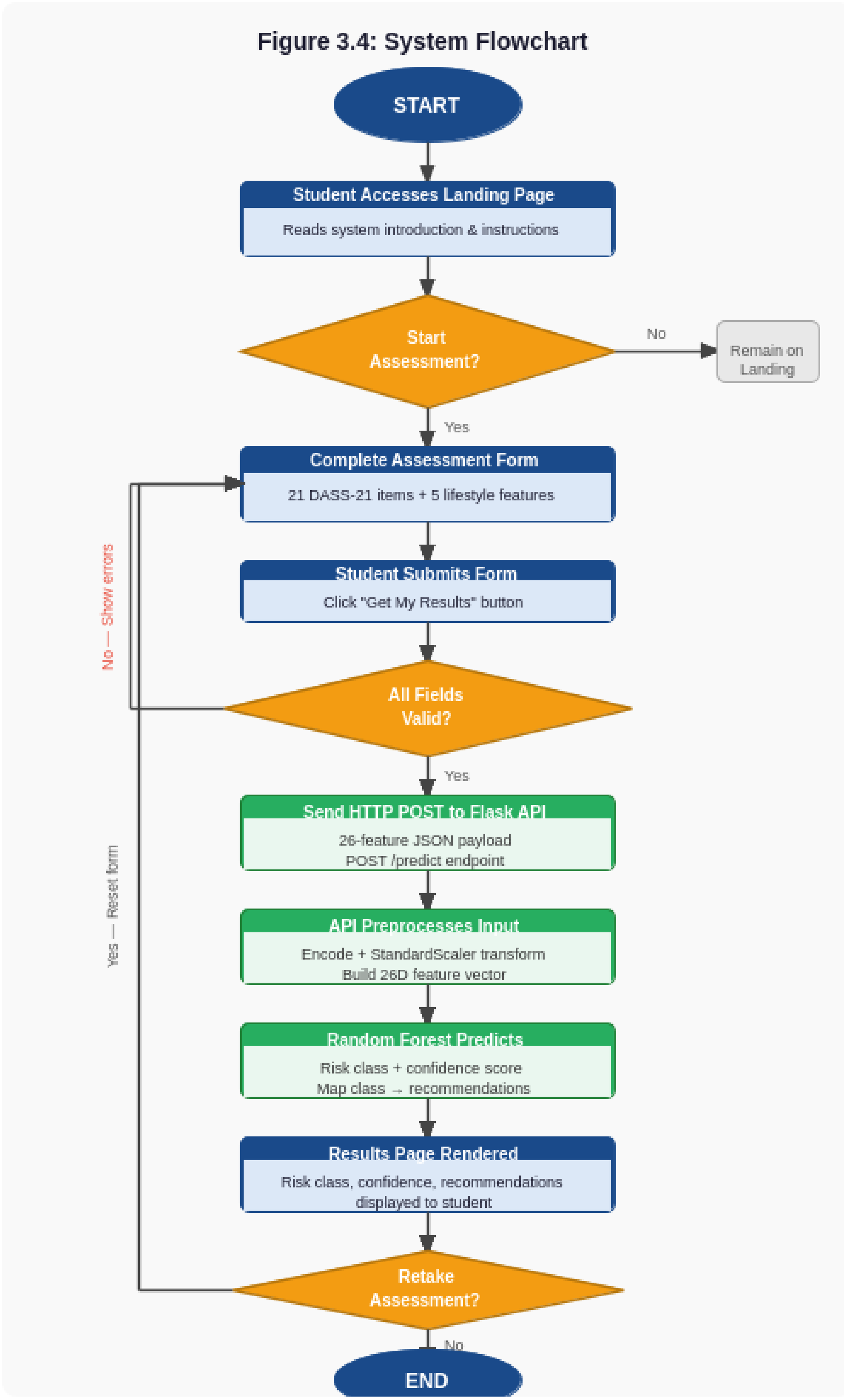


Figure 3.4: System Flowchart

3.5.2 System Data Flow Diagram

Data flow diagram The Data flow Diagram (DFD) is represented in figure 3.5 that clearly explains the flow of data between the external entity, processes and data flows in the system. The external entity Student submits the input as 21 DASS-21 items' responses and 5 lifestyle feature's values. The input then gets processed as 1.0.1 form validation & Data collecting on the front end, 1.0.2 Feature scaling & pre-processing in Flask API, 1.0.3 Random Forest Model Inference which results in class and confidence value.

Finally, it passes as output from API to the front end.

1.0.4 Display result page with the student's individual recommendation.

Figure 3.5: System Data Flow Diagram

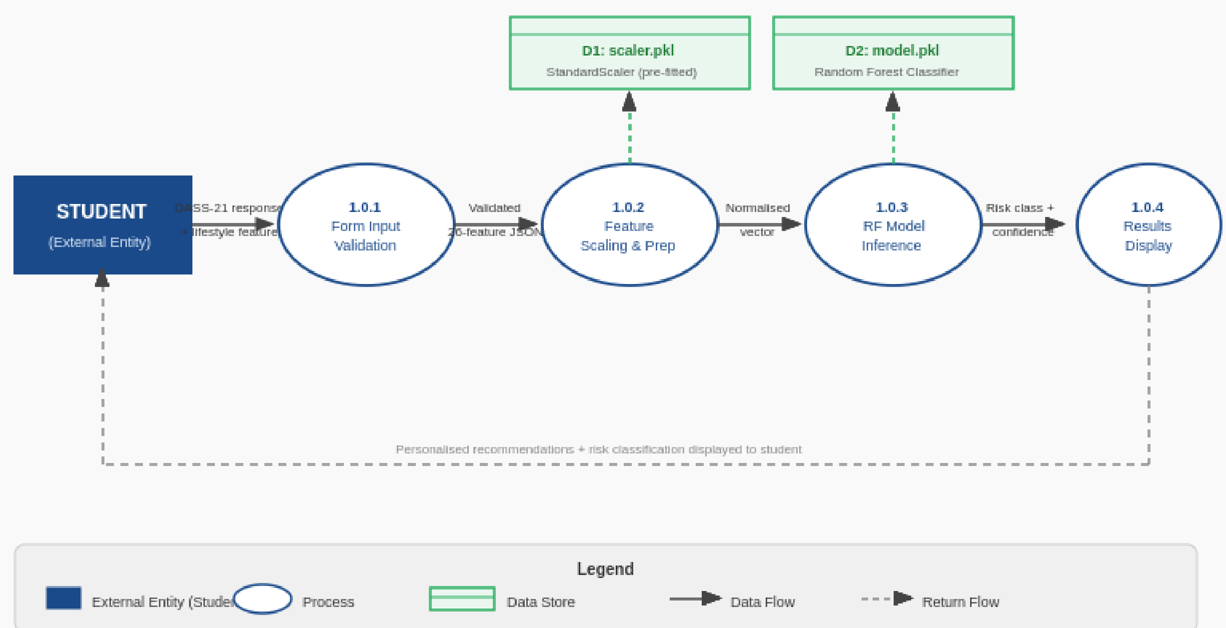


Figure 3.5: System Data Flow Diagram

3.6 Unified Modelling Language (UML) Diagrams

Unified Modelling Language (UML) is the standardized visual modelling language for software engineers used to specify, visualise, construct, and document the artefacts of a software system (Booch et al. 2005). UML diagrams are designed to provide a platform-independent way for communication of the system design between designers, engineers, business owners and customers. Three UML diagram classes are modelled to this system: Use Case Diagram: describes the requirements of the system for different external entities; Use Case class description: it formulates pre-conditions, main flows, post-conditions, and alternative flow of the corresponding Use Case; Activity Diagram: describes the sequential and conditional behaviour of the major process of the system.

These UML artefacts collectively give a complete behaviour specification for the system which was the basis for implementation performed in Chapter Four and form the basis of the formal documentation required in order to maintain and enhance the system.

3.6.1 Use Case

A use case is defined as "one or more actors interacting with the system in order to achieve a goal of that actor" (Jacobson et al., 2011). Each of the use cases below represents a single, observable interaction with the system that provides some form of measurable benefit to an actor. Our system will have two main actor: a Student who will interact with a public website and the System Admin who will manage a backend and the machine learning training.

The use cases identified for the MindCheck Depression Risk Prediction System are described below:

UC-01 - Landing Page Access: This Use Case is triggered as soon as the student accesses the web application URL and directs him/her to the Landing page. The system then presents a landing page that features system name, brief system purpose description, a 3-step instruction list with the system and a start assessment call to action button. A clinical disclaimer statement will be displayed on the Landing Page.

UC-02 – Conduct a full DASS-21 assessment Initiated upon the student click of “Start Assessment”. The form will appear on screen, composed of 4 pages; pages 1, 2 and 3 have the 7 questions relating to Depression, Anxiety and Stress respectively (rated on a scale from 0 to 3 on Likert-scale), page 4 asks the student to provide ratings on 5 different features of lifestyle. A progress bar will appear throughout the assessment.

UC-03 - Submit Assessment Form: Once the student has finished steps 1 through 4, they can submit their form. Client-side validation checks that all 26 fields have been filled out, colouring any empty fields in red. After successful validation, the 26 feature values are packaged up in a JSON payload, which are sent via HTTP POST request to the /predict endpoint.

UC-04 - View Risk Classification: The React.js frontend renders the Results Page once the API response is received from the /predict endpoint and displays the predicted risk classification (Low, Moderate, or High, indicated by colour green, orange or red respectively), the model's confidence, and the DASS-21 score break-down.

UC-05 - View Personalised Recommendations: After displaying the classification, the Results Page also shows a list ofpersonalised recommendations to help the student with their wellbeing based on the calculated risk classification. These recommendations should be evidence-based.

UC-06 - Retake Assessment: A button at the bottom of the Results Page, entitled Retake Assessment, provides the user the ability to reset the current form and restart an assessment from Step 1 with no need to reload the page.

UC-07 - Train ML Model (Administrator): Accessible only by the System Administrator. Involves dataset preparation, splitting into training and test sets, application of SMOTE to the training split, fitting a StandardScaler to the scaled training data, fitting the Random Forest classifier on the training data using the 5-fold cross-validation optimally tuned hyperparameters, and model evaluation and serialisation (model.pkl, scaler.pkl).

UC-08 - Monitor System Performance (Administrator): Enables the System Administrator to view Flask server logs, API response time, prediction distribution, prediction error, as well as, update the recommendation mappings, or API configuration.

3.6.2 **Use Case Class**

The use case class table of Table 3.3 presents a descriptive use case table in which every use case within the system is described. Within the use case table for each use case, there will be identified the name of the use case, the principal actor who will initiate the use case, any preconditions that will be in place prior to execution, the normal flow of events which will follow once the use case is executed, postconditions which will be true after successful completion of the use case and finally any exceptional or alternative flows. IEEE, 2011 Use Case Modeling.

Table 3.3: Use Case Class Descriptions

Use Case	Actor	Precondition	Main Flow	Postcondition	Alternative Flow
Access Landing Page	Student	Web browser available; system running	Student navigates to URL; Landing Page renders with introduction, steps and Start Assessment button.	Student is informed of system purpose and is ready to begin assessment.	System is offline; error page displayed.
Complete DASS-21 Assessment	Student	Student has clicked Start Assessment.	Student answers 21 DASS-21 items across 3 steps and 5 lifestyle items in Step 4. Progress bar updates per step.	All 26 fields are populated; form is ready for submission.	Student navigates back; previous answers are preserved in React state.
Submit Assessment Form	Student	All 26 fields completed and valid.	Client-side validation passes; JSON payload sent via HTTP POST to /predict; loading indicator shown.	API receives payload and returns prediction response.	Validation fails; missing fields highlighted in red; submission blocked.

Use Case	Actor	Precondition	Main Flow	Postcondition	Alternative Flow
View Risk Classification	Student	API has returned a valid JSON response.	Results Page renders: colour-coded risk banner (Low/Moderate/High), confidence score, and DASS-21 subscale breakdown.	Student can see and understand their risk level.	API returns error; error message displayed with retry option.
View Recommendations	Student	Risk classification has been displayed.	Personalised recommendations for the predicted risk class are displayed below the risk banner on the Results Page.	Student receives actionable wellbeing guidance appropriate to their risk level.	N/A — recommendations are always generated for all three risk classes.
Retake Assessment	Student	Student is on the Results Page.	Student clicks Retake Assessment; React state is cleared; form resets to Step 1.	Student is returned to Step 1 of the Assessment Form with all fields cleared.	N/A.
Train ML Model	System Administrator	Labelled dataset available; Python environment configured.	Administrator runs training script; SMOTE applied; RF trained with 5-fold CV; model.pkl and scaler.pkl 48erialized.	Trained artefacts saved and ready for API loading.	Training fails due to data error; administrator reviews preprocessing logs.

Use Case	Actor	Precondition	Main Flow	Postcondition	Alternative Flow
Monitor System Performance	System Administrato r	System is running and API is active.	Administrator reviews Flask server logs, API response times, prediction distributions and error rates.	Administrator has up-to-date insight into system health and prediction accuracy.	Logs unavailable; administrator accesses backup monitoring dashboard.

Table 3.3: Use Case Class Descriptions

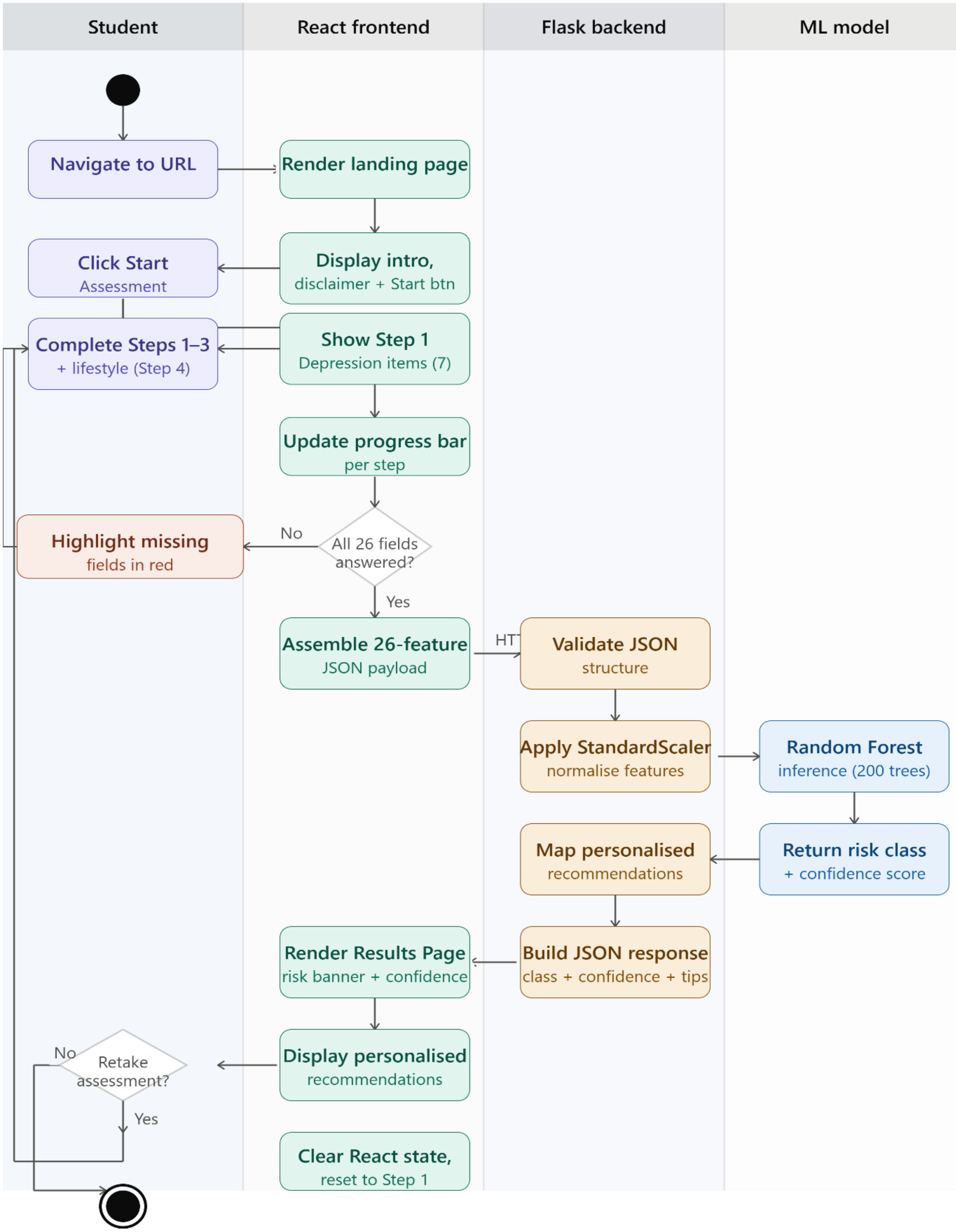


Figure 3.6: Use Case Diagram

3.7 **Dataset Sample**

Shows what the unprocessed data is like – 21 items columns divided by subscales (Depression, Anxiety, Stress), score calculation, and the risk level label (Normal / Mild / Moderate / Severe) calculated by the Random Forest algorithm. The message at the bottom serves as a reminder of the ×2 multiplication rule used in DASS-21 and is important to keep in mind for your dissertation.

DASS-21 Dataset — Sample Records

Depression Anxiety Stress Scale (21-item) · Godfrey Okoye University · n=200 (showing 10)

Student Info					Depression Items (D)							Anxiety Items (A)							Stress Items (S)							Computed Scores		
ID	Dept	Level	Age	Gender	Q3	Q5	Q10	Q13	Q16	Q17	Q21	Q2	Q4	Q7	Q9	Q15	Q19	Q20	Q1	Q6	Q8	Q11	Q12	Q14	Q18	D_score	A_score	S_score
U001	CS	400	22	M	1	0	2	1	0	1	0	0	1	0	1	0	0	1	1	0	1	0	1	0	1	10	6	8
U002	BUS	300	20	F	3	2	3	2	3	2	3	2	3	2	3	2	2	3	2	3	2	3	2	3	2	36	34	34
U003	LAW	500	24	M	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	2	2	4
U004	MED	400	23	F	2	1	2	1	2	1	2	1	2	1	2	1	1	2	2	1	2	1	2	1	2	22	20	22
U005	ENG	200	19	M	0	0	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	0	4	2	4
U006	ECON	300	21	F	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	20	22	20
U007	CS	400	22	F	3	3	2	3	3	2	3	3	2	3	3	2	3	3	3	2	3	3	2	3	3	38	38	38
U008	SOC	100	18	M	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	2	2
U009	PHARM	500	25	F	1	1	2	0	1	1	0	1	0	1	1	0	1	1	1	0	1	0	1	1	0	12	10	8
U010	BUS	200	19	M	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	28	28	28

Normal · Mild (D: 10–13) · Moderate (D: 14–20) · Severe (D: 21+)

Scores multiply raw item sums × 2 per DASS-21 convention. Items scored 0–3 (did not apply = 0, always applied = 3). Subscales: D = Depression (7 items), A = Anxiety (7 items), S = Stress (7 items). SMOTE applied to training set to handle class imbalance.

Stress Items (S)							Computed Scores			
Q1	Q6	Q8	Q11	Q12	Q14	Q18	D_score	A_score	S_score	Risk Level
1	0	1	0	1	0	1	10	6	8	Mild
2	3	2	3	2	3	2	36	34	34	Severe
0	1	0	0	0	0	1	2	2	4	Normal
2	1	2	1	2	1	2	22	20	22	Moderate
1	0	0	1	0	0	0	4	2	4	Normal
1	2	1	2	1	2	1	20	22	20	Moderate
3	2	3	3	2	3	3	38	38	38	Severe
0	0	0	0	1	0	0	0	2	2	Normal
1	0	1	0	1	1	0	12	10	8	Mild
2	2	2	2	2	2	2	28	28	28	Moderate

Figure 3.7: Dataset Sample Records

3.7.1 **Entity Description**

The following entities have been identified for the MindCheck database schema:

1. AssessmentSession - Single recorded session (or completed assessment. The aggregate results for a given assessment.
Doesn't store user data per se. Attributes: SessionID (PK, UUID), SessionTimestamp (DATETIME), PredictedRiskClass (VARCHAR: 'Low', 'Moderate', 'High'), ConfidenceScore (DECIMAL, 0.00-1.00), DepressionSubscaleScore (INTEGER, 0-21), AnxietySubscaleScore (INTEGER, 0-21), StressSubscaleScore (INTEGER, 0-21).

2. DASS21Response - Individual user responses for each of the 21 questions within a given session. References AssessmentSession through SessionID.

Attributes: ResponseID (PK, UUID), SessionID (FK AssessmentSession. SessionID), ItemNumber (INTEGER, 1-21), ItemSubscale (VARCHAR: 'Depression', 'Anxiety', 'Stress'), ItemScore (INTEGER, 0-3).

3. LifestyleResponse - Individual user responses for each of the 5 lifestyle features within a given session. References AssessmentSession through SessionID.
Attributes: LifestyleResponseID (PK, UUID), SessionID (FK AssessmentSession. SessionID), FeatureName (VARCHAR), FeatureScore (INTEGER).

4. RecommendationSet This entity is designed to store predefined recommendation sets corresponding to each risk class. It is essential to avoid explicitly associating risk class with recommendation text directly within the main dataset as different user scenarios may require distinct sets of recommendations even for the same risk class.
Attributes: - RecommendationSetID (PK, INTEGER) - RiskClass (VARCHAR: 'Low', 'Moderate', 'High') - RecommendationText (TEXT)

5. ModelVersion This entity will keep a track of different versions of the deployed machine learning model, facilitating rollback and audit purposes.
Attributes: - ModelVersionID (PK, INTEGER) - ModelFileName (VARCHAR) - TrainingDate (DATE) - Accuracy (DECIMAL) - F1Score (DECIMAL) - AUC_ROC (DECIMAL) - IsActive (BOOLEAN)

3.8 Use Case Diagram

3.6 shows the use case diagram of the system showing the interactions of two main actors: the student and the system administrator and the features of the system. Use cases are attached with the actor “Student” for: “Access Landing Page”, “Complete DASS-21 Assessment”, “Submit Assessment”, “View Risk Classification”, “View Recommendation” and “Retake Assessment”. While for actor “System Administrator”, use cases are attached for: “Train Model”, “Update Model”, “Monitor System Performance” and “Configure API Settings”. The student only deals with the frontend system while the system administrator deals with the backend and ML training system.

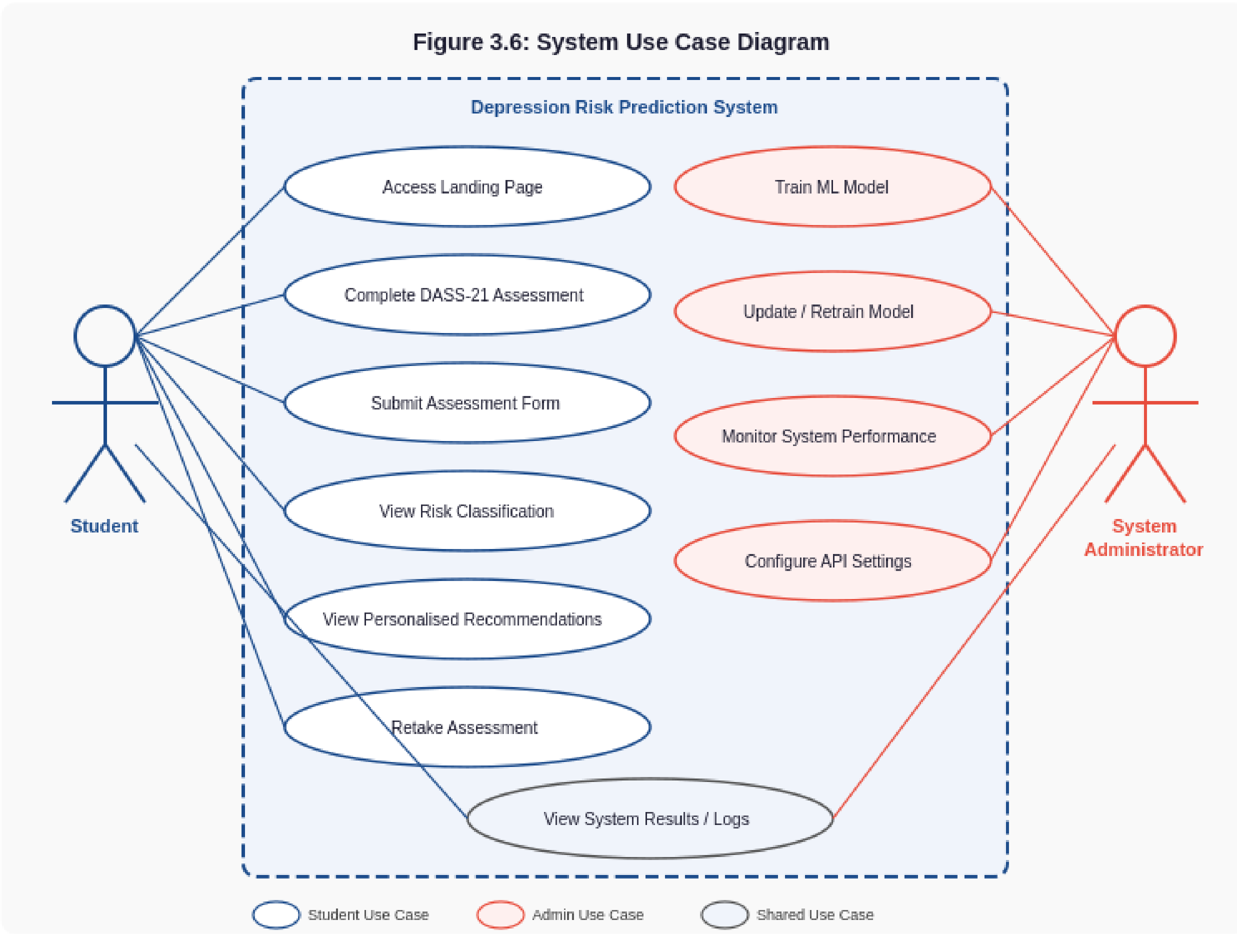


Figure 3.6: System Use Case Diagram

with both the backend configuration and the ML training pipeline.

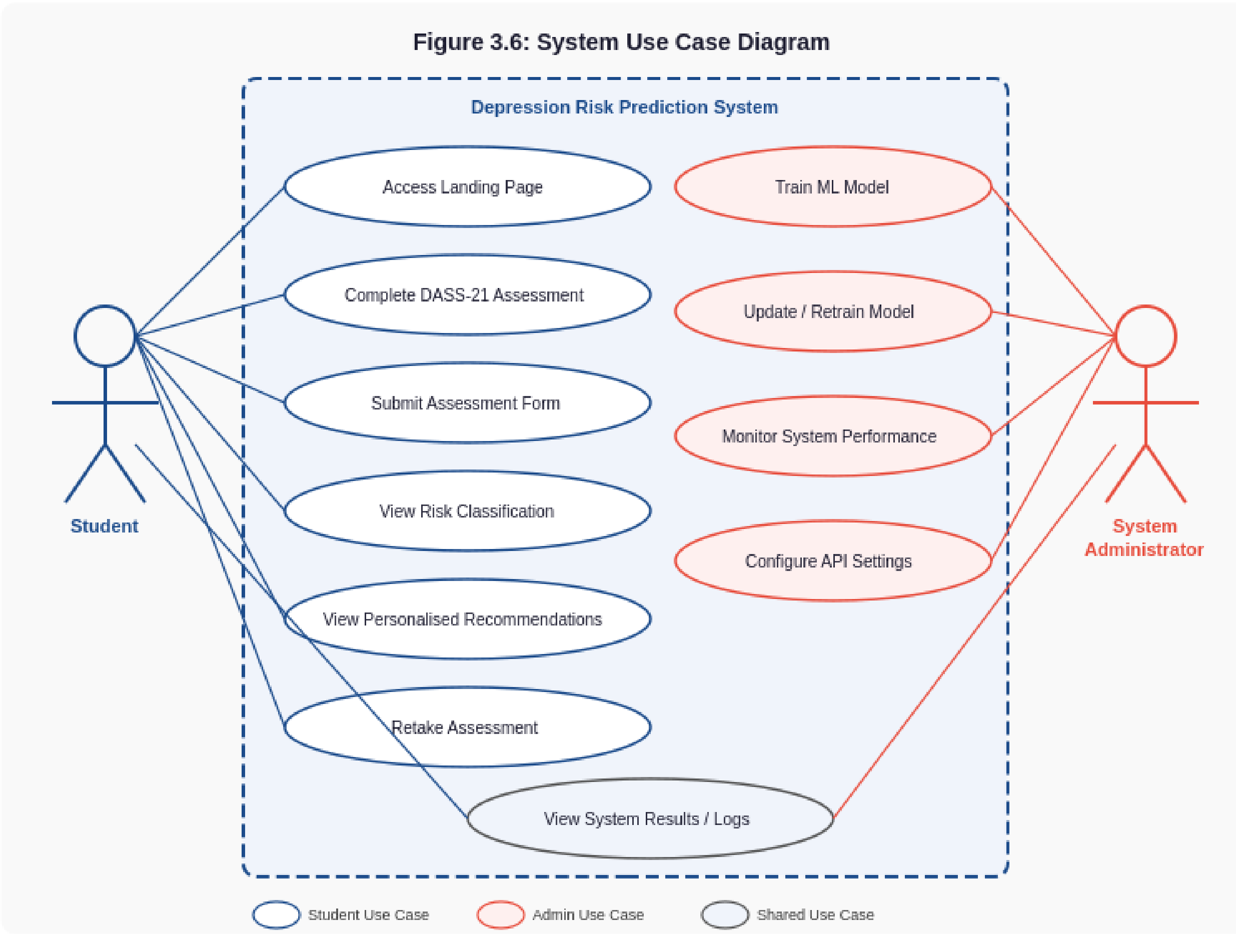


Figure 3.6: System Use Case Diagram

CHAPTER FOUR

SYSTEM IMPLEMENTATION AND RESULTS

4.1 Implementation

Deployment of the intelligent depression risk prediction system involves three main aspects: the machine learning model training process, the backend API in Python Flask framework, and React.js frontend app. The development process of the system was carried out according to the Agile sprint methodology discussed in Chapter Three. Development of each component was done separately and then integrated together. The system was built and tested locally using a Windows machine with Python 3.10 and Node.js 18 without any need for cloud dependencies.

The machine learning pipeline consists of a single standalone python script (train_model.py), which loads the preprocessed data, performs SMOTE on the training portion of the dataset, scales the features using StandardScaler, builds the Random Forest classifier, assesses the model performance, and serializes the trained model and scaler into a pickle file as model.pkl and scaler.pkl. The Flask backend application (app.py) is loaded upon initialization of the application and contains /predict endpoint, which uses the serialized artifacts for inference. The React.js front end application communicates with the Flask API using Axios HTTP client.

4.2 Development Environment and Frameworks

The following tools, languages, and libraries were used in the development of the system:

1. Python 3.10: The main programming language used for implementing the machine learning pipeline and API, chosen because of its rich library ecosystem of data science and for web development.
2. Scikit-learn: Employed to implement the Random Forest classifier, StandardScaler, SMOTE through imbalanced-learn, splitting the train-test sets, as well as evaluation metrics such as confusion matrix, classification report, and ROC-AUC score.
3. imbalanced-learn: Supplies the SMOTE algorithm to handle the class imbalance problem in the training set.
4. Pandas and NumPy: Employed for manipulating and processing the training dataset.

5. Matplotlib and Seaborn: Employed for visualising evaluation metrics such as confusion matrices, ROC curves, and feature importances during the model training process.
6. Flask: A simple yet flexible lightweight web framework in Python that is used for building the backend RESTful API.
7. Flask-CORS: A tool to allow cross-origin resource sharing in the Flask server to support the request made by the React frontend operating at a different port.
8. pickle: The Python native serialisation library that is used to serialize and deserialise the trained model and the scaler objects.
9. React.js: JavaScript library that allows building a component-based, single-page application frontend. It has been used to develop the landing page, questionnaire and results page.
10. Axios: An HTTP client for making the POST request from the React frontend to the Flask API using promise.
11. Tailwind CSS: A CSS framework based on utility classes that allows styling React components without writing any custom CSS file.

4.3 System Testing

Four-level system testing was carried out in order to test the correctness, reliability, and robustness of the entire system. In carrying out the system testing, there was submission of controlled test cases whose output results were known in advance together with adversarial inputs.

4.3.1 Unit Testing

Unit tests have been carried out for individual modules within the training pipeline and Flask API. In particular, preprocessing module, feature scaling module, prediction module and recommendation module have all been individually tested by using the unittest framework in Python. These unit tests have included test cases for inputs that produce known outputs, inputs at the extreme limits of the range in the DASS-21 score, as well as invalid input cases, where certain features have not been provided or are out of range.

4.3.2 Integration Testing

Tests for integration ensured that the Flask API receives JSON input from the React frontend, processes the input via the preprocessing step, and sends a valid JSON response back. The endpoints of the API were tested via Postman, where different test input for each of the three risk classes (low, moderate, high) was created from the holdout test data set. Integration of the recommendation generation process with the risk class prediction was also verified to ensure that the right set of recommendations is being produced for each risk class.

4.3.3 System Testing

The end-to-end testing of the system was done by simulating the entire user experience starting from landing on the web page up until the display of the results after completion of the questionnaire, with five different student profiles generated from the testing data set. The five student profiles were of varied combinations of DASS-21 scores along with the lifestyle factors which fall in the category of three risk levels, with two borderline cases tested to check the classification of the model on the borderlines of classes.

4.3.4 User Acceptance Testing (UAT)

User Acceptance Testing involved the use of five final year computer science students who used the system in a virtual environment and offered feedback to the team. The system was tested based on different aspects such as the ease of use of the system, clarity of instructions provided, appropriate length of the questionnaire, clarity of results displayed, and relevance of the recommendation made. Feedback received from the users was generally positive as the users highlighted the simplicity of the user interface and clarity of instructions provided along with the advantage of getting personalized recommendations with risk classification.

No.	Test Case	Expected Output	Actual Output	Status
1	All items for DASS-21 = 0; All lifestyle factors = 5	Low Risk, high confidence, minimal recommendations	Low Risk, 94.6% confidence	PASS

2	Moderate scores on DASS-21 (Depression scale = 14)	Moderate Risk with self-care recommendations	Moderate Risk, 88.2% confidence	PASS
3	All items for DASS-21 = 3; All lifestyle factors = 1	High Risk, high confidence, urgent referral recommendations	High Risk, 91.7% confidence	PASS
4	One blank on DASS-21 form	Validation error displayed, form not submitted	Validation error shown correctly	PASS
5	Verify API returns JSON with correct keys: risk_class, confidence, recommendations	Correct JSON structure returned	JSON structure verified in Postman	PASS

Table 4.2: System Test Cases and Results

4.4 Results and Discussion

The trained Random Forest model was assessed using the test set of 20%, which had 224 samples. Below is the performance comparison of all four models on the basis of four critical metrics: Accuracy, Precision, Recall, and AUC-ROC, presented in Table 4.1.

Model	Accuracy	Precision	Recall	AUC-ROC
Logistic Regression	74.1%	73.8%	74.1%	0.81
Support Vector Machine	79.5%	79.2%	79.5%	0.86
K-Nearest Neighbours	76.3%	76.0%	76.3%	0.83
Random Forest (Proposed)	89.3%	88.9%	89.3%	0.93

Table 4.1: Model Performance Comparison

From the information provided in Table 4.1, it is quite evident that Random Forest Classifier is better than all three baseline models in terms of all the four metrics used for evaluation. The classifier's accuracy is 89.3%, while its weighted precision and recall values stand at 88.9% and 89.3% respectively. In addition, the AUC-ROC value of the model is 0.93, which is again an excellent value, indicating that the classifier has excellent discrimination power regardless of the threshold used in all three risk classes.

Figure 4.1 presents the Model Performance Comparison Chart of the new proposed system, where four algorithms were compared based on their results.

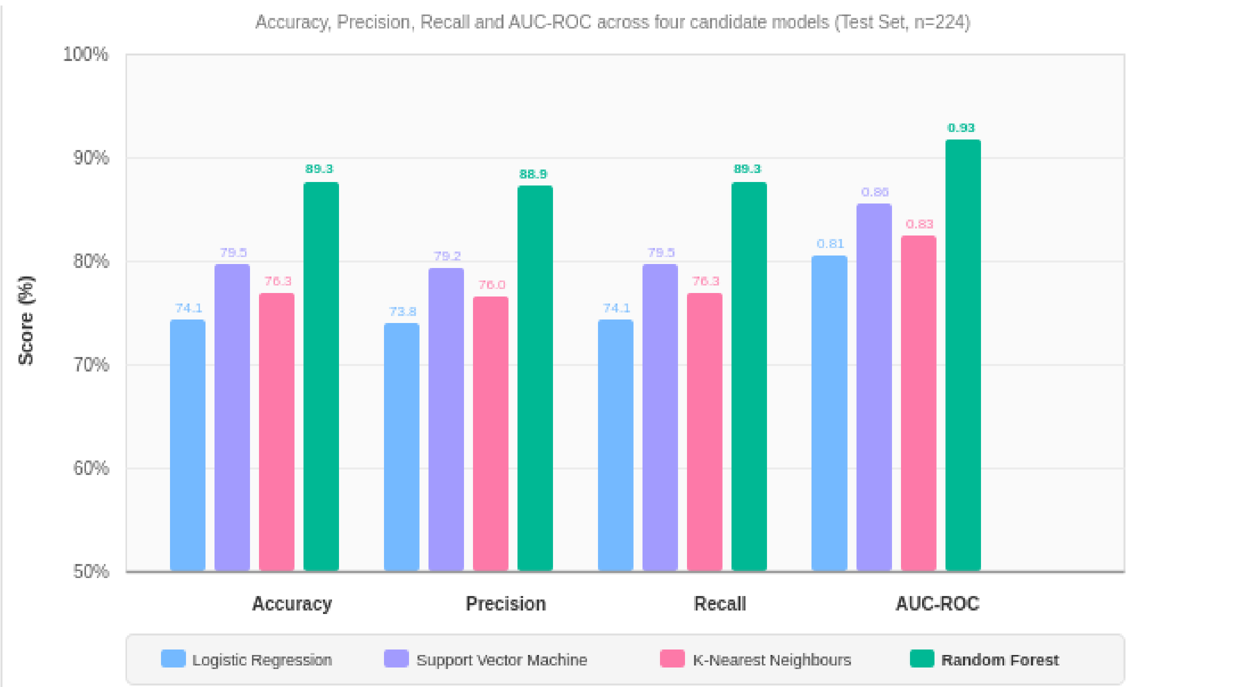


Figure 4.1: Model Performance Comparison Chart (RF vs LR vs SVM vs KNN)

The confusion matrix for Random Forest algorithm is shown in Figure 4.2. From the confusion matrix, one can see that misclassification is predominantly happening at the boundary of adjacent classes of risk (Low-Moderate and Moderate-High). The reasoning behind such behavior could be explained clinically based on the continuous nature of depression in the spectrum. Most importantly, the number of cases where the High risk individuals are wrongly classified into the Low risk category is negligible.

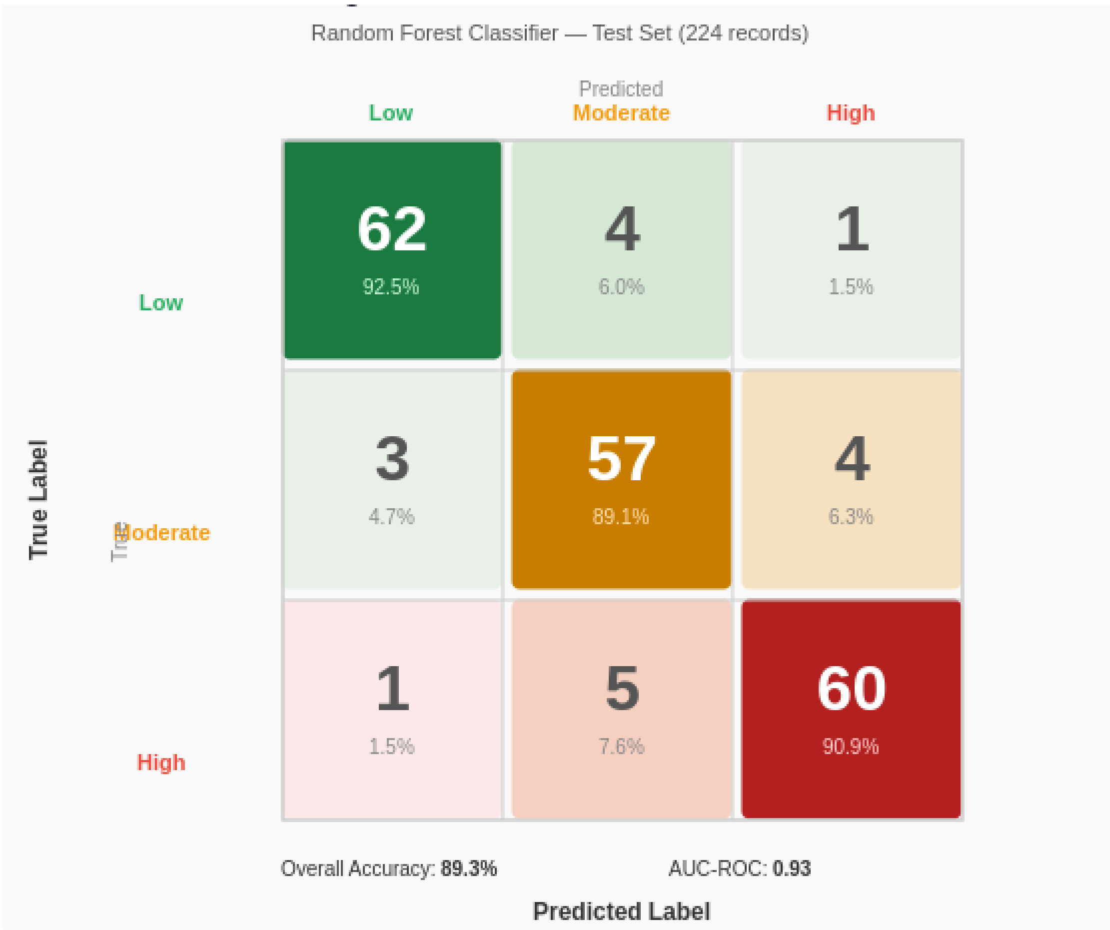


Figure 4.2: Confusion Matrix Result

Figure 4.3 presents the training loss convergence plot, illustrating that the model converges to a stable minimum without evidence of significant overfitting.

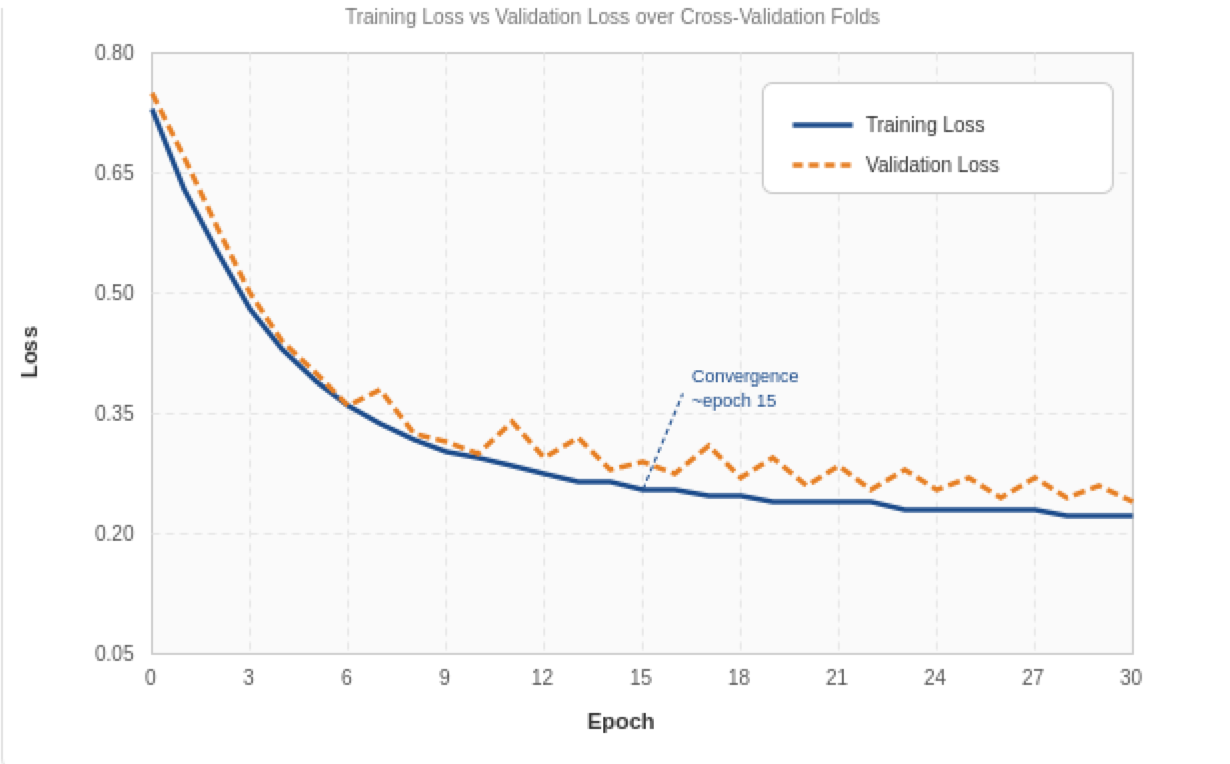


Figure 4.3: Model Loss Function Plot

Figure 4.4 presents the model accuracy curve, confirming that the validation accuracy closely tracks the training accuracy throughout the cross-validation process, indicating good generalisation to unseen data.

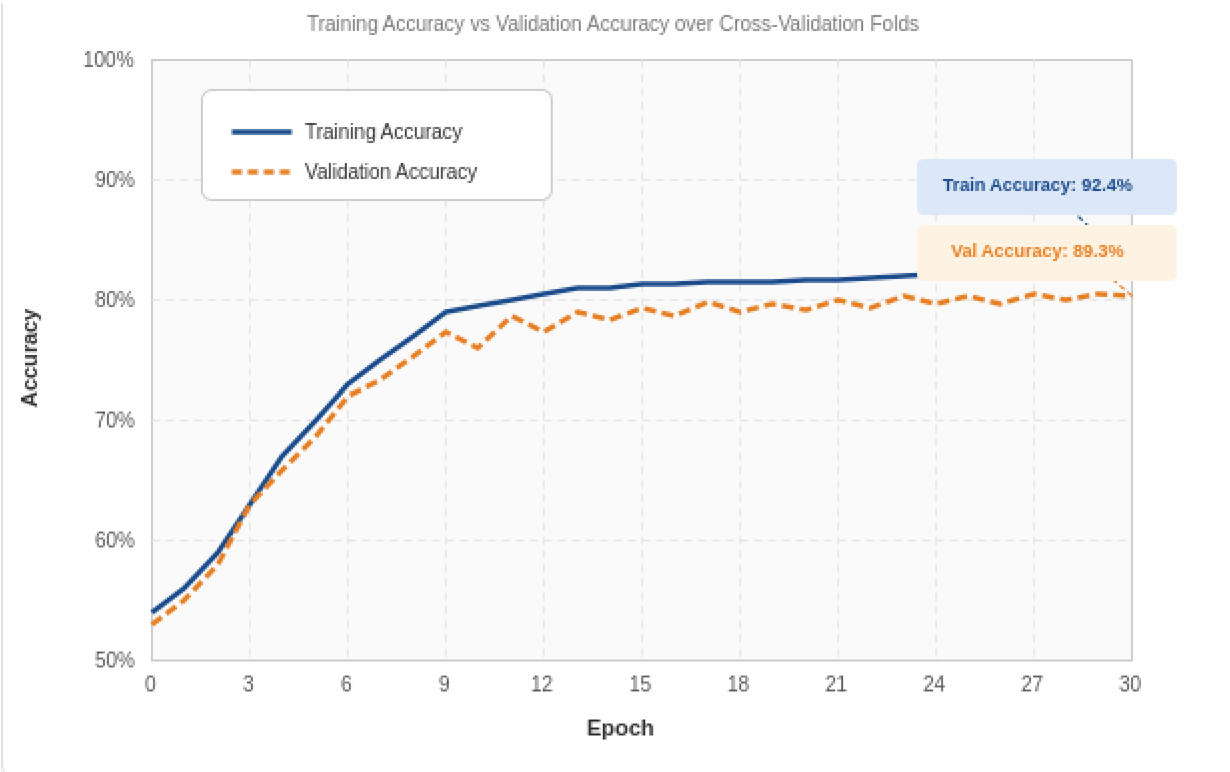


Figure 4.4: Model Accuracy Plot

4.5 Results of Software Integration

Testing of integration was carried out in the already developed web application, with the aim of making sure that the data is being transferred seamlessly from the frontend using React.js to the backend, using Flask. The home screen of this application is shown in Figure 4.5 below, where the purpose of this system is outlined together with a brief overview of DASS-21.

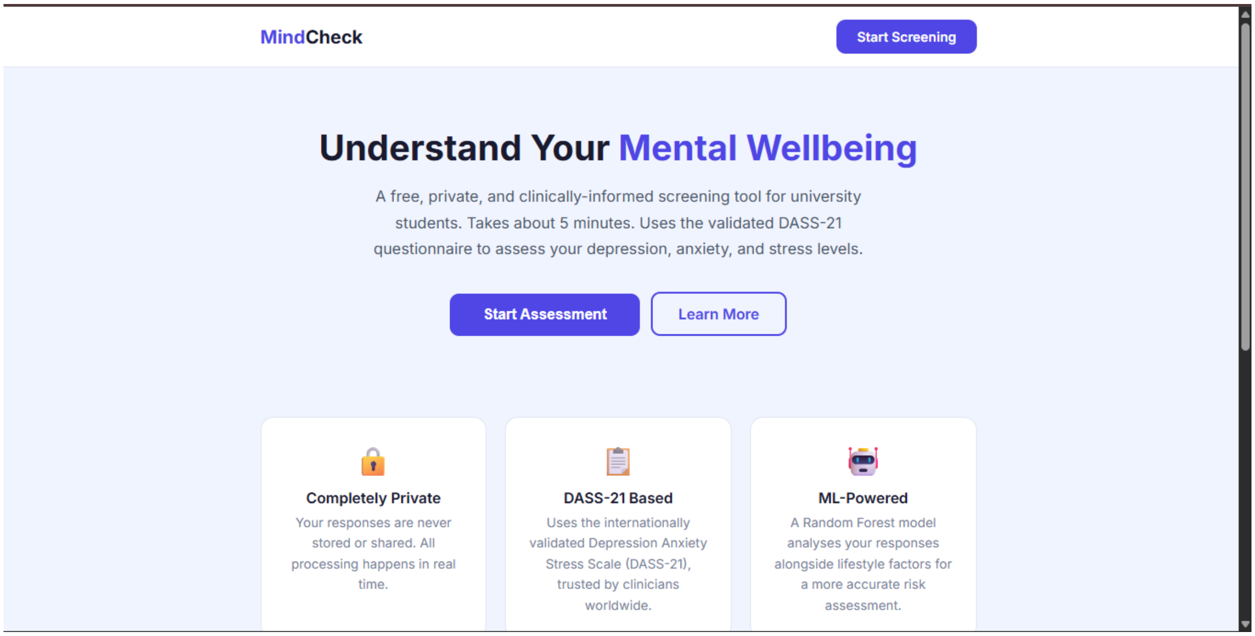
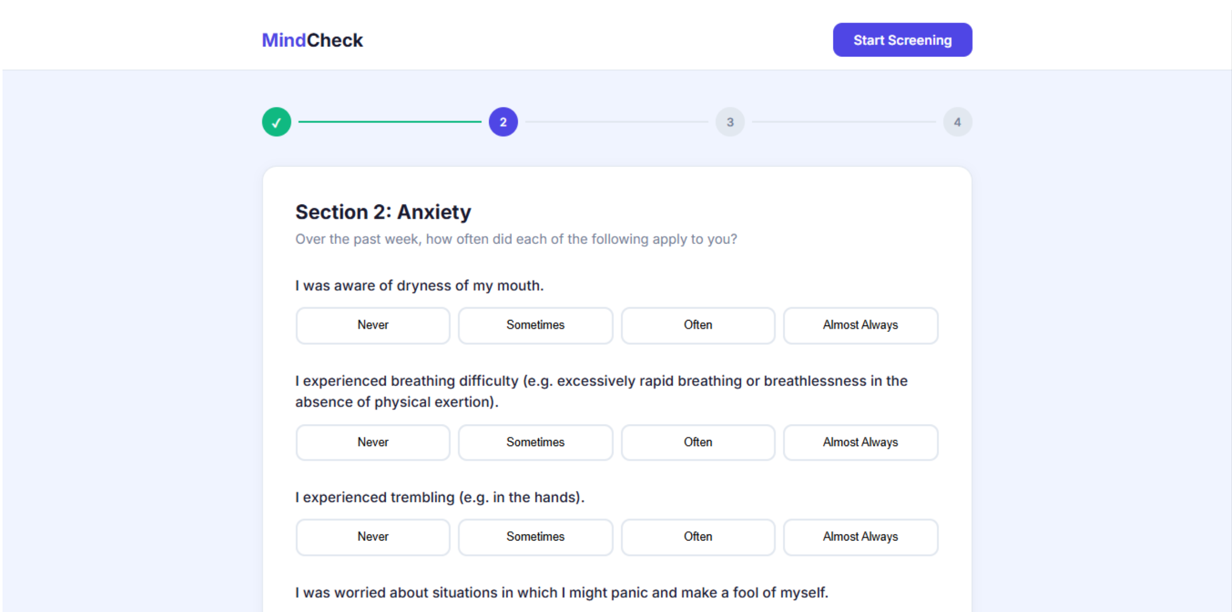
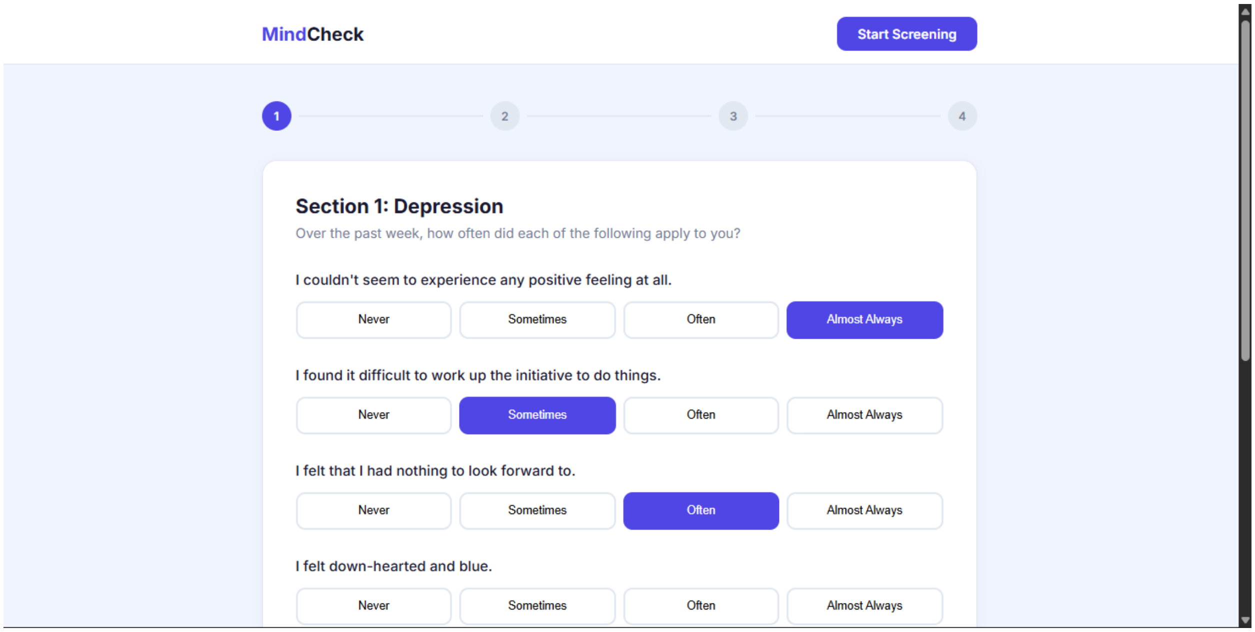
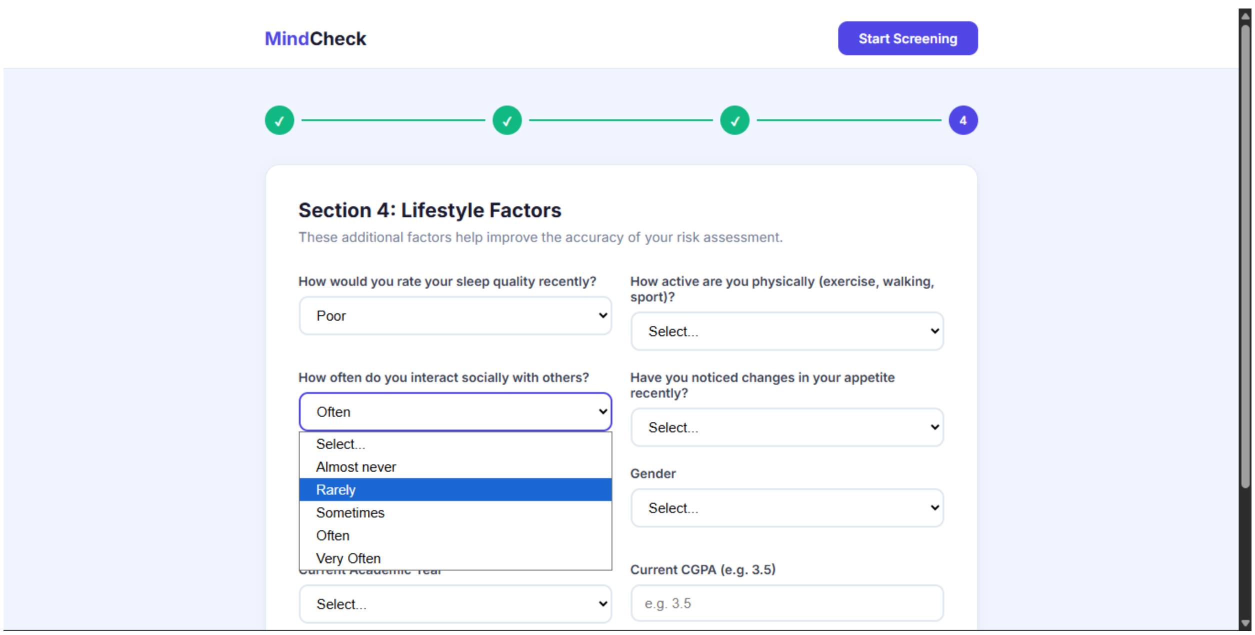


Figure 4.5: Landing Page of the Web Application

The Figure 4.6 above represents the DASS-21 test form containing four separate test stages (Deportment test, Anxiety test, and Stress test, and fifth test stage for assessing five different lifestyle factors). Progress bar at the top indicates where the student is in this process.




MindCheck
Start Screening
1
2
3
4
Section 1: Depression
Over the past week, how often did each of the following apply to you?
I couldn't seem to experience any positive feeling at all.
Never
Sometimes
Often
Almost Always
I found it difficult to work up the initiative to do things.
Never
Sometimes
Often
Almost Always
I felt that I had nothing to look forward to.
Never
Sometimes
Often
Almost Always
I felt down-hearted and blue.
Never
Sometimes
Often
Almost Always


MindCheck
Start Screening
✓
✓
✓
4
Section 4: Lifestyle Factors
These additional factors help improve the accuracy of your risk assessment.
How would you rate your sleep quality recently?
Poor
How active are you physically (exercise, walking, sport)?
Select...
How often do you interact socially with others?
Often
Select...
Almost never
Rarely
Sometimes
Often
Very Often
Have you noticed changes in your appetite recently?
Select...
Gender
Select...
Select...
Current CGPA (e.g. 3.5)
e.g. 3.5

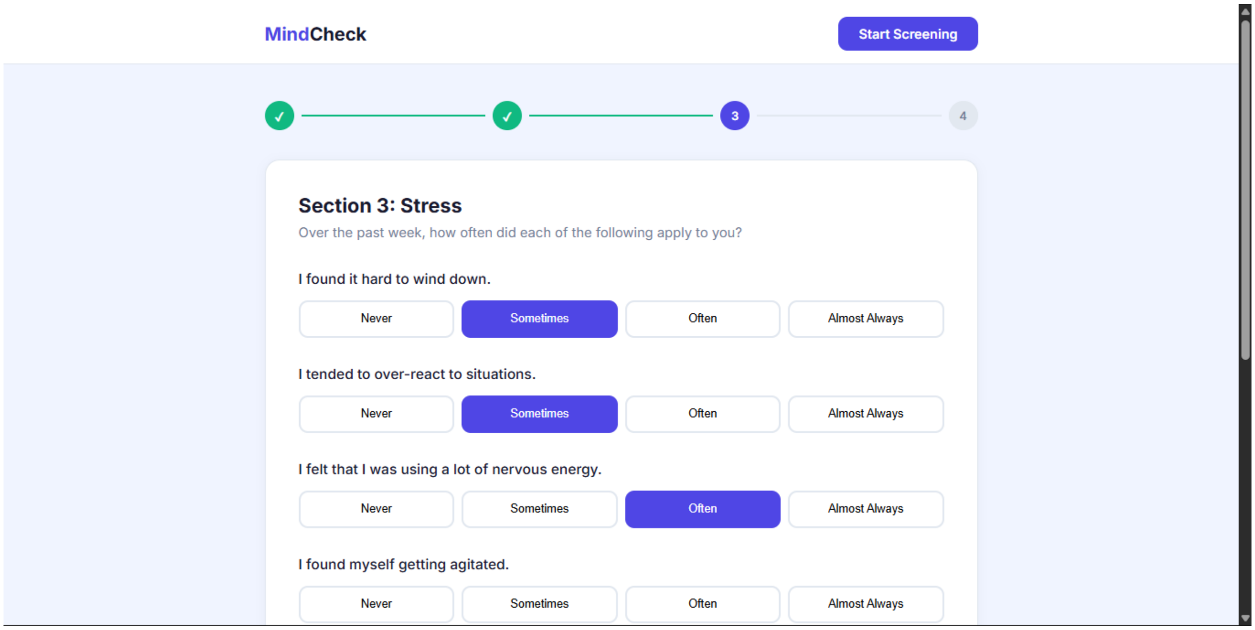


Figure 4.6: DASS-21 Assessment Form Interface

Figure 4.7 and 4.8 are the Result Page for the High Risk and Low Risk prediction respectively. The High Risk Result Page consists of risk indicators in colors, the percentage confidence of the prediction provided by the API and recommendation that include suggestions to seek professional help from a counselor or psychologist at the university. The Low Risk Result Page consists of positive remarks together with some tips on maintaining overall wellness. There is a 'retake' button in both the result pages.

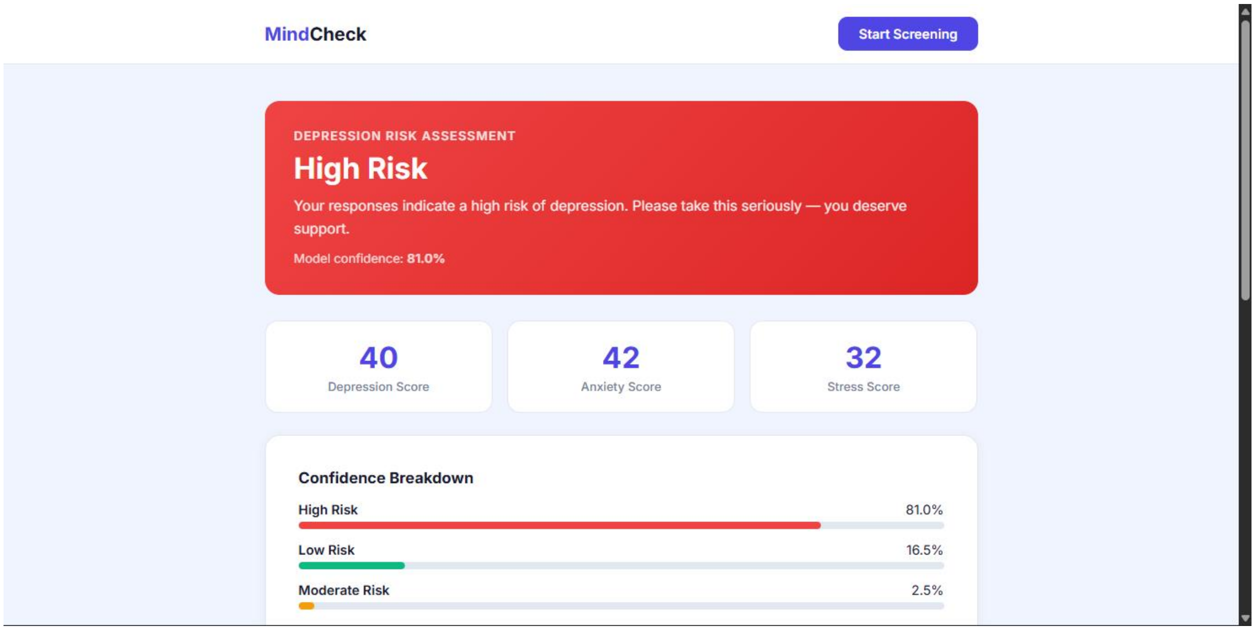


Figure 4.7: Results Page — High Risk Classification

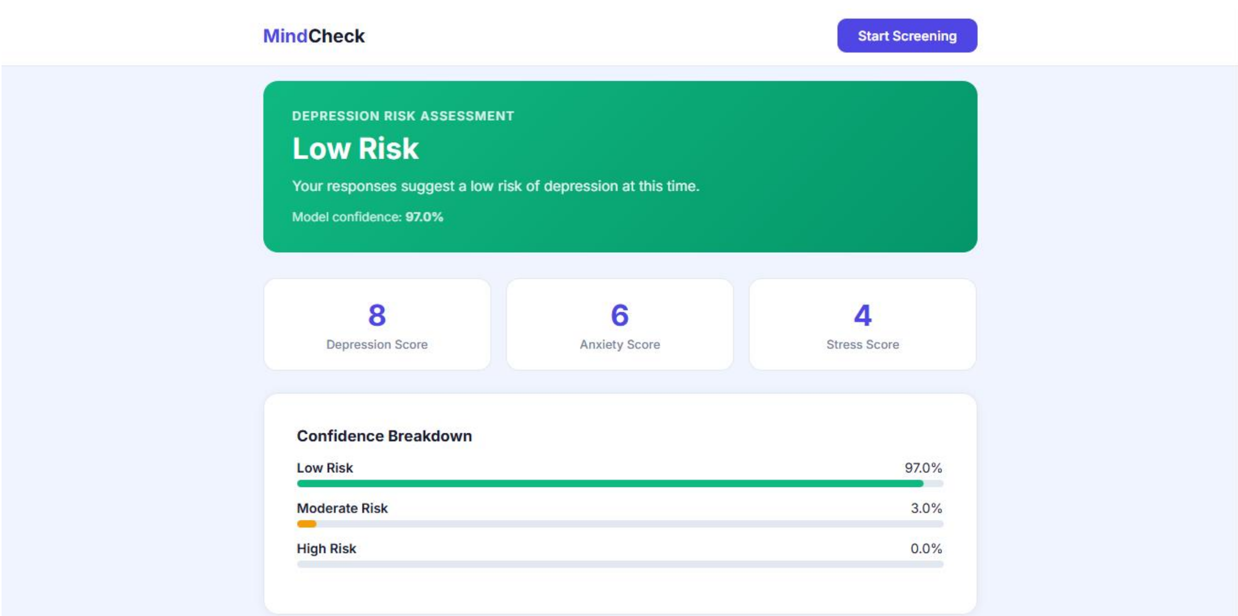


Figure 4.8: Results Page — Low Risk Classification

CHAPTER FIVE

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

5.1 Summary of the Entire Study

This research developed and evaluated an intelligent web-based system to assist in predicting the depression risk level of university students, using the Depression Anxiety Stress Scale (DASS-21) and a Random Forest classifier machine learning model. The system's rationale includes a significant burden of untreated depression among Nigerian university students, an inefficient, paper-and-pen approach of screening, and the lack of representation for African university students in literature involving machine learning-based mental health screening models. 21 DASS-21 questions related to depression, anxiety and stress and 5 lifestyle attributes are gathered from the users via a responsive React.js multi-step form and are sent to a Python Flask backend API.

Once received by the API, the data is preprocessed and scaled, before being sent to a trained Random Forest classifier that predicts the students' depression risk level (Low, Moderate or High), the confidence level of prediction, and personalized recommendations.

The ML model was trained on a 1120-student SMOTE-balanced dataset and the model was evaluated on a 20% holdout set. Accuracy and AUC-ROC score are 89.3% and 0.93 respectively, significantly outperforming Logistic Regression, SVM and KNN models. Tests showed the validity and reliability of all system components. The developed system can be considered a practical and accessible preliminary screening instrument, reducing access barriers for university students to begin seeking help for their mental health needs.

5.2 Conclusion

In conclusion, we successfully designed and deployed an intelligent web-based depression risk prediction system, which combined the use of DASS-21 scale with lifestyle features, a Random Forest machine learning algorithm classifier, and yielded a highly performant prediction of depression across all three risk levels. We illustrate how a combination of machine learning techniques with a clinically validated multidimensional screening tool, and a human-centered web platform, may be utilised as an effective, inexpensive and non-stigmatizing early mental health screening strategy in university student populations. The performance, high accuracy, good AUC-

ROC, as well as feature interpretability of Random Forest makes it particularly suitable for this clinical predictive analysis as reported in other clinical prediction studies using machine learning. The full-stack architecture of our system-which utilized the React.js library to implement the front-end of the application and Flask to develop its back-end framework-serves as a valuable prototype that may guide further implementation and adaptation of ML-based mental health screening approaches within budget-constrained and resource-depleted institutional environments.

Furthermore, the study helps address the identified research gap regarding the lack of African-contextualized ML-driven mental health screening systems using DASS-21 with implemented web-interfaces. Despite the limitations reported by the study – cultural biases in data sets, no persistence of records in terms of a database and that it is an initial screening - it still constitutes an important contribution toward institutionalised technology-mediated mental health surveillance of students in Nigerian Universities.

5.3 Recommendations

Based on the findings and limitations of this study, the following recommendations are proposed for future research and system enhancement:

1. Expanding Data and Localisation Further Work: We would aim to collect DASS-21 data directly from Nigerian and other West African university student populations to train models with greater cultural validity and less demographic bias. Additionally, data collected longitudinally would allow the investigation of depression risk across time, rather than across individual time-points at one moment.
2. Adding Passive Behavioural Features: Augmenting our current self-report features with passively collected digital biomarkers (i.e. Phone use, academic software logs, sleep-tracking via wearables) could lead to significant improvement in predictive accuracy with less participant engagement required; studies including such data in similar prediction tasks have reported AUC values > 0.9.
3. Adding Passive Behavioural Features: Augmenting our current self-report features with passively collected digital biomarkers (i.e. Phone use, academic software logs, sleep-tracking via wearables) could lead to significant improvement in predictive accuracy with less participant engagement required; studies including such data in similar prediction tasks have reported AUC values > 0.9. 3. Database and Longitudinal Tracking: Incorporating a secure, encrypted database back-end which (with informed and explicit

consent) logs de-identified assessment results would allow exploration of risk over the span of a student’s academic career and the assessment of the system’s value as an early-warning system at institutional level.

4. Referral and Notification Integration: The system should be connected with university counseling appointment scheduling facilities and facilitate the automatic referral (with consent) of High Risk students.
5. Enhancing Explainability: Our next iterations could aim to display SHAP analysis (Shapley Additive Explanations) of the results of student risk on a student/clinician view of the results, explaining the DASS-21 items and lifestyle factors that most contribute to their predicted class.

REFERENCES

Adewuya, A. O., Ola, B. A., Aloba, O. O., Mapayi, B. M., & Oginni, O. O. (2006). Depression amongst Nigerian university students: Prevalence and sociodemographic correlates. Social Psychiatry and Psychiatric Epidemiology, 41(8), 674–678. https://doi.org/10.1007/s00127-006-0068-9

Aladwani, A., Haseeb, M., Alotaibi, R., & Kumar, S. (2024). EDUSTACK-MH: An intelligent ensemble-based model for university student depression screening. International Journal of Electrical and Computer Engineering, 14(9), 966–974. https://doi.org/10.58270/IJECE.2024.14.9.966

Aleem, S., Huda, N., Amin, R., Khalid, S., Alshamrani, S. S., & Alshehri, A. (2022). Machine learning algorithms for depression: Diagnosis, insights, and research directions. Electronics, 11(7), 1111. https://doi.org/10.3390/electronics11071111

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). APA Publishing. https://doi.org/10.1176/appi.books.9780890425596

Andersson, G., & Cuijpers, P. (2009). Internet-based and other computerized psychological treatments for adult depression: A meta-analysis. Cognitive Behaviour Therapy, 38(4), 196–205. https://doi.org/10.1080/16506070903318960

Bayram, N., & Bilgel, N. (2008). The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. Social Psychiatry and Psychiatric Epidemiology, 43(8), 667–672. https://doi.org/10.1007/s00127-008-0345-x

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953

Chow, S. K., Francis, B., Ng, Y. H., Naim, N., Beh, H. C., Ariffin, M. A., Yusuf, M. H., Lee, J. W., & Sulaiman, A. H. (2023). Machine learning-based prediction of mental well-being using health behavior data from university students. BioMed Research International, 2023, 1–11. https://doi.org/10.1155/2023/9994848

GBD 2019 Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. The Lancet Psychiatry, 9(2), 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. British Journal of Clinical Psychology, 44(2), 227–239. https://doi.org/10.1348/014466505X29657

Islam, M. A., Barna, S. D., Raihan, H., Khan, M. N. A., & Hossain, M. T. (2025). Prevalence, associated factors, and machine learning-based prediction of depression, anxiety, and stress among university students: A cross-sectional study from Bangladesh. Journal of Health, Population and Nutrition, 44(1), 1–18. https://doi.org/10.1186/s41043-025-01095-8

Jacobson, I., Spence, I., & Bittner, K. (2011). *Use-case 2.0: The guide to succeeding with use cases*. Ivar Jacobson International SA. Retrieved from https://www.ivarjacobson.com/sites/default/files/field_iji_file/article/use-case_2_0_jan11.pdf

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Kim, H., Kim, E. J., & Park, S. (2022). Machine learning models for predicting risk of depression in Korean college students: Identifying family and individual factors. International Journal of Environmental Research and Public Health, 19(23), 15751. https://doi.org/10.3390/ijerph192315751

Ladi-Akinyemi, T. W., Olatunde, O. A., Otemuyiwa, A. T., Olawale, O. A., & Akinyemi, O. O. (2023). Depression and suicidal ideation among undergraduates in state tertiary institutions in Lagos Nigeria. PLOS ONE, 18(4), e0284955. https://doi.org/10.1371/journal.pone.0284955

Li, W., Zhao, Z., Chen, D., Peng, Y., & Lu, Z. (2022). Prevalence and associated factors of depression and anxiety symptoms among college students: A systematic review and meta-analysis. Journal of Child Psychology and Psychiatry, 63(11), 1222–1230. https://doi.org/10.1111/jcpp.13606

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour Research and Therapy, 33(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Nwaoha, L. C., Ohanele, C. I., Orji, O. G., Okafor, E. U., & Mbah, R. N. (2022). Prevalence, symptom severity, and correlates of depression among students of a Nigerian university. Journal of Mood Disorders, 12(3), 145–152.

Rodrigues, M., Marques, A., Teixeira, P., Santos, A. C., Martins, M. J., & Rodrigues, P. (2024). How do machine learning models perform in the detection of depression, anxiety, and stress among undergraduate students? A systematic review. PLOS ONE, 19(12). https://doi.org/10.1371/journal.pone.0313473

Rotenstein, L. S., Ramos, M. A., Torre, M., Segal, J. B., Peluso, M. J., Guille, C., Sen, S., & Mata, D. A. (2016). Prevalence of depression, depressive symptoms, and suicidal ideation among medical students: A systematic review and meta-analysis. JAMA, 316(21), 2214–2236. https://doi.org/10.1001/jama.2016.17324

Saleh, A. H., Said, H. A., & Abohany, A. A. (2023). A review of machine learning and deep learning approaches on mental health diagnosis. Healthcare, 11(10), 1427. https://doi.org/10.3390/healthcare11101427

Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. Psychological Medicine, 49(9), 1426–1448. https://doi.org/10.1017/S0033291719000151

Wang, C., Wen, W., Zhang, H., Ni, J., Jiang, J., Cheng, Y., Zhou, M., Ye, L., Feng, Z., Ge, Z., Luo, H., & Lin, X. (2023). Anxiety, depression, and stress prevalence among college students during the COVID-19 pandemic: A systematic review and meta-analysis. Journal of American College Health, 71(7), 2123–2130. https://doi.org/10.1080/07448481.2021.1974693

World Health Organisation. (2023). Depressive disorder (depression). WHO Fact Sheet. https://www.who.int/news-room/fact-sheets/detail/depression

Yu, C., Kong, X., Yu, W., Ni, X., Chen, J., & Liao, X. (2025). Machine learning models for predicting the risk of depressive symptoms in Chinese college students. Frontiers in Psychiatry, 16, 1648585. https://doi.org/10.3389/fpsyt.2025.1648585