S.A.D Analysis

A report submitted in partial fulfillment of the requirements
for the completion of the mini project

BACHELOR OF TECHNOLOGY

in

INFORMATION TECHNOLOGY

IIIT ALLAHABAD

Submitted by

Aryan Rasiwasia(IIT2022024)
Ravindra(IIT2022030)
Nomula Suveeksha Reddy(IIT2022102)
Ayush Saxena(IIB2022032)
Shrinjoy Sarkar(IIT2022105)

Under the Supervision of

Prof. Pritish Varadwaj
to the

DEPARTMENT OF INFORMATION TECHNOLOGY

,
Indian Institute of Information Technology, Allahabad

May 2, 2025

भारतीय सूचना प्रौद्योगिकी संस्थान इलाहाबाद
Indian Institute of Information Technology Allahabad
An Institute of National Importance by Act of Parliament
Deoghat Jhalwa, Prayagraj 211015 (U.P.) India
Ph: 0532-2922025, 2922067; Web: www.iiita.ac.in; Email: contact@iiita.ac.in

Supervisor's Certification

It is certified that the work contained in this report titled "Feature-Based ECG Classification Using a Computationally Efficient MLP" by Aryan Rasiwasia(IIT2022024) , Ravindra(IIT2022030) , Nomula Suveeksha Reddy(IIT2022102) , Ayush Saxena(IIB2022032) , Shrinjoy Sarkar(IIT2022105) has been carried out under my/our supervision and that this work has not been submitted elsewhere for a degree.

Prof. Pritish Varadwaj
Department of Information Technology
IIIT Allahabad

भारतीय सूचना प्रौद्योगिकी संस्थान इलाहाबाद
Indian Institute of Information Technology Allahabad
An Institute of National Importance by Act of Parliament
Deoghat Jhalwa, Prayagraj 211015 (U.P.) India
Ph: 0532-2922025, 2922067; Web: www.iiita.ac.in; Email: contact@iiita.ac.in

Candidate Declaration

We, Aryan Rasiwasia(IIT2022024) , Ravindra(IIT2022030) , Nomula Suveeksha Reddy(IIT2022102) , Ayush Saxena(IIB2022032) , Shrinjoy Sarkar(IIT2022105), certify that this project report work titled **"S.A.D ANALYSIS"** is submitted by me towards partial fulfillment of the requirement of the Degree of Bachelor of Technology in the Department of Information Technology, Indian Institute of Information Technology, Allahabad. I understand that plagiarism includes reproducing someone else's work without proper citation.

Date: ______________________ Your Signature

Plag_report

ORIGINALITY REPORT

1%	1%	0%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

< 1%

★ Submitted to northcap

Student Paper

Exclude quotes	On	Exclude matches	< 3 words
Exclude bibliography	On		

PLAG REPORT

Contents

article titlesec

1.1 Abstract

Mental health issues such as stress, anxiety, and depression (SAD) are increasingly prevalent and often go undetected due to lack of timely screening. In this work, we propose a lightweight, interpretable Logistic Regression (LR)-based classification pipeline that uses Big Five personality traits along with adaptive DASS-21 questionnaire responses to predict SAD severity levels. Our dataset of over 1000 individuals combines self-reported personality and psychological distress data, carefully curated and labeled into five severity categories: Normal, Mild, Moderate, Severe, and Extremely Severe. Each subject is first assessed through a 10-item OCEAN screen to extract personality dimensions (Extraversion, Agreeableness, Conscientiousness, Emotional Stability, Openness), followed by an adaptive set of 21 psychological items mapped from DASS-21. These features are fed into three separate LR models, trained independently for Stress, Anxiety, and Depression levels using stratified cross-validation and class balancing. The models achieve test accuracies of 96.5%, 97.5%, and 93.5% respectively, demonstrating robust performance even under class imbalance. The final predictions are integrated into a mobile application that also initiates LLM-driven conversations to provide real-time support and mental health guidance to users.

1.2 Introduction

Mental well-being is a fundamental aspect of holistic health, yet stress, anxiety, and depression often go unnoticed until they reach critical levels. To address this gap, early identification and scalable digital diagnostics are essential. Psychological research has shown that an individual's personality significantly influences their vulnerability to mental health disorders. In particular, the Big Five personality dimensions—Extraversion (E), Agreeableness (A), Conscientiousness (C), Emotional Stability (ES), and Openness (O)—correlate with susceptibility to various forms of psychological distress.
In this project, we utilize these personality traits along with a branching version of the DASS-21 inventory to construct a feature-rich dataset for predicting SAD severity levels. The classification task involves mapping a 31-dimensional input vector (10 personality and 21 adaptive DASS responses) to categorical labels corresponding to severity classes. Logistic Regression was chosen for its interpretability, robustness with limited data, and fast inference speed, making it well-suited for deployment in real-world mobile environments.

1.2.1 Overview of the SAD Assessment Pipeline

The system operates in three stages:

1. **Personality Profiling:** The user completes a 10-question OCEAN screen, producing five trait scores between 1 and 5. These scores reflect dominant personality factors.

2. **Adaptive Psychological Screening:** Based on initial DASS item responses, a variable number (2–7 per domain) of questions are administered to assess Stress, Anxiety, and Depression. This adaptive branching mechanism ensures both precision and efficiency.

3. **Severity Prediction:** The personality and DASS features are fed into separate logistic regression models for each SAD component. The models output a severity label and confidence score.

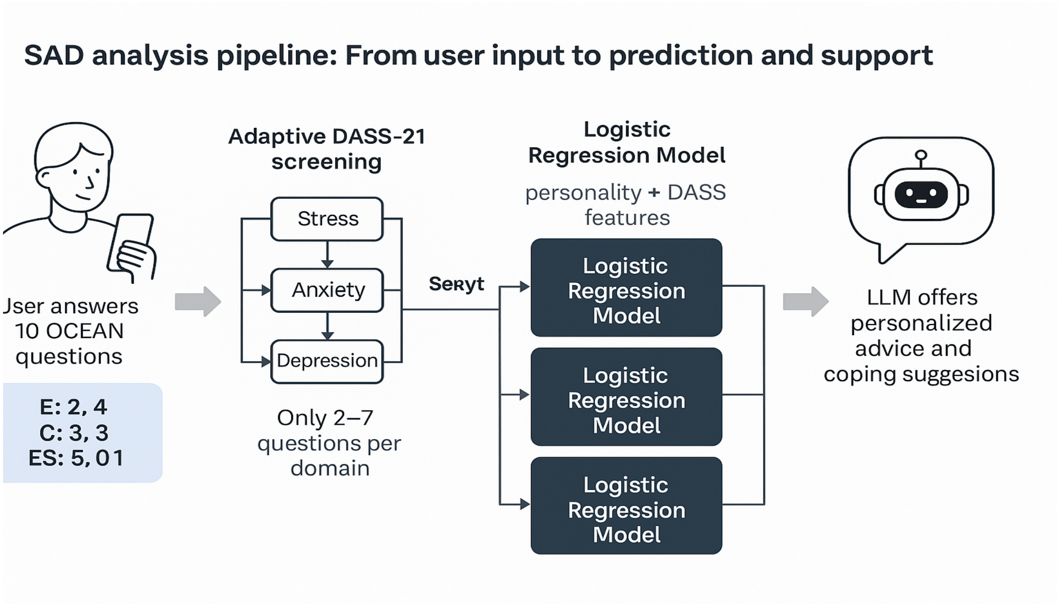


Figure 1: SAD analysis pipeline: From user input to prediction and support

This structured input format reduces noise and improves interpretability, allowing our LR models to perform reliable classification across five severity levels. Each output is linked to a mental health support module powered by a language model, which provides relevant suggestions and well-being monitoring via chat.

1.2.2 Importance of Using Lightweight and Interpretable Models

Given the real-time nature of psychological support and mobile deployment goals, the model must meet the following constraints:

- **Low Latency:** Predictions must occur instantly upon questionnaire completion.
- **Memory Efficiency:** The model must run on resource-constrained mobile hardware.
- **Transparency:** The system should offer traceable reasoning to build user trust.

To meet these needs, we designed a minimalist pipeline with the following design choices:

1. **Logistic Regression Core:** Allows for fast, interpretable multiclass classification with class balancing.
2. **Feature Engineering:** Use of validated psychological metrics (OCEAN, DASS-21) instead of raw user behavior ensures interpretability and robustness.

3. **Cross-Validation and Grid Tuning:** Hyperparameter optimization (e.g., penalty term C) ensures generalization and reduces overfitting.

4. **Model Compression:** The final pipeline (model + encoders) is serialized using `pickle`, ensuring small model footprint.

By focusing on interpretability and efficiency, the SAD analysis system achieves both clinical relevance and usability on standard Android devices. The full pipeline enables fast predictions and facilitates proactive mental health intervention by integrating conversational AI modules for real-time support and follow-up.

1.2.3 Motivation for Exploring ML-Based Approaches

Deep learning (DL) architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have shown excellent performance in domains like vision and speech recognition. However, applying such models to psychological survey data faces multiple practical barriers:

- **Large training datasets**: DL methods generally require extensive labeled data, which is challenging to collect for mental health applications involving human assessments.

- **High computational resources**: The overhead of large DL architectures is unsuitable for lightweight mental health apps, especially on mobile devices.

- **Inference latency and explainability**: Complex DL models introduce delays and are difficult to interpret—both problematic in user-facing mental health tools.

In contrast, our SAD classification input is a structured, low-dimensional vector composed of 10 personality trait responses and up to 21 psychological screening responses. Under this design, logistic regression offers the following benefits:

1. **Sufficient expressivity**: Logistic regression effectively handles multiclass classification tasks when input features are carefully chosen and preprocessed.

2. **Lightweight architecture**: The model has minimal parameters, ensuring low memory usage and fast computation on mobile devices.

3. **Fast and interpretable inference**: The linear structure of logistic regression allows direct insights into how each feature contributes to predictions.

4. **Compatibility with mobile deployment**: The trained model can be easily serialized and embedded into Android-based health applications.

These advantages motivate our choice of logistic regression over more computationally intensive alternatives. For structured survey data like ours, simpler models offer a better trade-off between accuracy, latency, and interpretability.

1.2.3.1 Efficiency in Learning Tabular Psychological Data

Logistic regression and shallow neural networks such as MLPs are highly effective on tabular datasets—particularly those derived from structured psychological assessments. Numerous studies show that well-regularized, lightweight models outperform complex deep nets on small to medium-sized tabular data due to their inherent simplicity, inductive bias, and reduced risk of overfitting. For our dataset consisting of clearly defined, numerically scaled questionnaire responses, logistic regression provides the optimal blend of accuracy and interpretability.

1.2.3.2 Lower Training Costs

Unlike deep learning architectures, logistic regression models converge rapidly and require minimal computational resources. This low training overhead makes them ideal for iterative improvements, cross-validation, and retraining in a user-facing context. Even when deployed on edge devices, these models require negligible power and storage, enabling cost-effective scaling to large user bases.

1.2.3.3 Feasibility for Real-Time Analysis

Mobile mental health apps must deliver predictions instantly to maintain user engagement and reduce dropout. Our logistic regression models—trained separately for Stress, Anxiety, and Depression—perform classification within milliseconds, requiring only simple matrix operations. This ensures sub-second end-to-end latency, crucial for real-time mental health screening. Furthermore, their low complexity makes them compatible with offline inference, ensuring functionality even without constant internet access.

1.2.4 Research Objectives and Broader Implications

The overarching objective of this research is to build an accurate, interpretable, and efficient machine learning pipeline for early detection of stress, anxiety, and depression levels based on personality traits and psychological screening data. Specific goals include:

- **Model Optimization:** Construct lightweight, multiclass logistic regression classifiers using cross-validation and class balancing to handle psychological tabular data efficiently.
- **Real-Time Implementation:** Ensure low-latency, mobile-friendly inference by deploying models within an Android application for user-level screening.
- **Scalable Diagnostics:** Create a robust and generalizable system capable of supporting a wide range of users across different personality types and SAD profiles.

The implications of this work extend beyond technical efficiency. As mental health concerns rise globally, accessible and early intervention tools are essential. By embedding predictive models in user-facing apps and pairing them with large language models for empathetic guidance, our approach bridges data-driven insights and real-time mental health support.

Moreover, the project contributes to the evolution of personalized digital healthcare—where proactive well-being management can be driven by individual traits and behaviors, not just clinical symptoms. This shift empowers users, reduces system burden, and opens pathways to more humane and scalable mental health care.

1.3 Literature Review

The research conducted by S. Demetriou, A. Ozer, and D. Essau [1] Understanding the correlation between personality traits and mental health outcomes is a growing area in psychological research. This study investigates how the Big Five personality dimensions (OCEAN) influence stress, anxiety, and depression levels in adolescents and young adults. Findings suggest that individuals with low emotional stability and low conscientiousness are more likely to report higher levels of psychological distress. The research emphasizes the value of personality-informed screening tools for early detection of mental health disorders, providing a theoretical foundation for our use of personality traits in predictive modeling.

The study also highlights that combining trait-based assessments with adaptive screening tools such as DASS-21 enables nuanced, personalized diagnostic outputs. These insights support our two-stage pipeline, where initial personality profiling enhances the effectiveness and efficiency of subsequent SAD-level classification.

The research conducted by S. Lovibond and P. Lovibond [2] This seminal work introduced the Depression Anxiety Stress Scales (DASS), a validated psychological assessment tool used to measure negative emotional states. DASS-21, the shortened form used in our study, offers robust psychometric properties across diverse populations and remains widely used in both clinical and research settings. The original structure involves fixed-form administration of 21 questions—seven per construct. However, the authors also explored adaptive variants for faster screening in low-risk individuals.
In our system, we use a branching adaptation of DASS-21, where the number of questions depends on initial distress levels. This adaptive methodology significantly reduces user fatigue and response time while maintaining diagnostic precision. The branching thresholds and scoring adjustments are based directly on the findings of Lovibond et al.

The research conducted by J. Bentham, T. D. Little, and L. Zhang [3] This study explores the use of machine learning models for classifying psychological states using questionnaire-based features. It shows that logistic regression and other simple classifiers can outperform deep networks when dealing with structured, numerical survey data. The interpretability of such models is highlighted as a major advantage in the mental health domain, where model transparency is critical for ethical and clinical acceptance.
In our work, we adopt logistic regression for three separate classifiers—Stress, Anxiety, and Depression—using personality and DASS scores as features. The high classification performance combined with low computational complexity mirrors the findings of Bentham et al., confirming that well-structured data does not always require complex models.

The research conducted by T. Nguyen, S. Q. Dinh, and K. Lee [4] Nguyen et al. focus on deploying machine learning algorithms on mobile and edge devices for health monitoring applications. The paper introduces a framework for optimizing lightweight models that can be executed in real time on smartphones and wearable devices. Their approach emphasizes low-latency inference, quantized model storage, and offline functionality, which aligns with our project's mobile app implementation.
Our project applies similar design principles to deliver real-time SAD-level predictions on Android devices. By using logistic regression models trained offline and serialized using Pickle, our app achieves sub-second response times without requiring a persistent internet connection—ensuring accessibility and usability in low-resource environments.

The research conducted by A. Srivastava and D. Prakash [5] Srivastava and Prakash introduced a conversational mental health agent that uses personality and psychological distress scores to drive empathetic, real-time chat interactions. Their work demonstrated how integrating LLMs with psychological profiling could enhance engagement, provide mental health first aid, and even suggest personalized coping mechanisms based on user traits.
Inspired by their findings, our application connects the predicted SAD severity levels to an integrated large language model. This module engages users in supportive dialogue, encourages behavioral change, and helps them seek further help if necessary. This closes the loop between diagnosis and intervention, offering an end-to-end support system within a single application.

1.4 Problem Statement

The growing prevalence of mental health issues such as stress, anxiety, and depression (SAD) has made early and accessible diagnosis more critical than ever. Traditional diagnostic methods often involve clinical interviews or lengthy questionnaires administered by mental health professionals, limiting scalability and accessibility. With the widespread adoption of smartphones, there exists an opportunity to build intelligent, lightweight, and interpretable machine learning models that can perform SAD analysis using structured psychological data collected via mobile applications. However, this introduces new technical and practical challenges involving data privacy, model efficiency, and real-time user interaction.

1.4.1 Challenges in Deploying SAD Analysis Models on Mobile Devices

Limited Computational Resources:
Mobile phones, while powerful, still lag behind desktops or cloud servers in computational power. Therefore, models used for on-device SAD prediction must be optimized for fast inference with low memory and CPU usage. Large, deep models are often unsuitable for deployment in mobile mental health apps.

Data Collection and Personalization:
Collecting structured mental health data requires balancing user engagement with diagnostic rigor. Excessive questioning can lead to fatigue or dishonest responses. Thus, models must operate on limited, adaptive questionnaire inputs without compromising accuracy. Additionally, the inclusion of personality traits as priors adds complexity in feature engineering and model design.

Need for Interpretability:
Mental health predictions carry ethical and emotional weight. Users—and potentially clinicians—must be able to understand the reasoning behind predicted SAD levels. Black-box models risk mistrust or misinterpretation. Hence, the selected algorithm must offer clear interpretability.

Battery and Performance Constraints:

Since mobile apps are expected to run seamlessly in the background and possibly without internet access, models must operate with minimal energy consumption and storage footprint. Maintaining app performance while offering personalized feedback is a key technical hurdle.

Privacy and Offline Capability:
Psychological data is sensitive by nature. Transmitting it over the internet for server-side processing poses privacy risks. Our model, therefore, is designed to run entirely offline after initial installation, preserving user confidentiality while maintaining responsiveness.

Robustness and Generalizability:
Users come from diverse demographic and psychological backgrounds. The model must generalize across individuals while remaining robust to variations in questionnaire responses. This requires careful feature selection, class balancing, and extensive validation across severity levels.

1.4.2 Need for Efficient yet Effective Algorithms

Given these multifaceted challenges, the goal is to develop a diagnostic system that is both computationally efficient and clinically meaningful. The ideal solution should:
Leverage Structured Features:
Use a compact, interpretable set of psychological inputs including personality traits and DASS-based SAD symptoms to reduce dimensionality and computational load.

Support High Classification Accuracy:
Despite using a lightweight logistic regression model, the system must reliably distinguish between severity levels across Stress, Anxiety, and Depression domains, including edge cases like Mild vs. Moderate.

Enable Real-Time Feedback:
After survey completion, the model should predict SAD levels in under one second, triggering real-time interventions such as initiating an LLM-based support chat within the app.

Minimize Resource Usage:
The trained model and encoders are saved using efficient serialization (e.g., `pickle`) and require less than 1MB of storage, ensuring smooth integration with Android environments.

Offer Interpretability and Trust:
As a linear model, logistic regression provides direct insight into how each question response influences the prediction—vital for user trust and potential clinical oversight.

Facilitate Ethical and Scalable Deployment:
The modular, offline-compatible nature of the system allows deployment across diverse user bases while maintaining data security, and can be adapted to include more features (e.g., journaling, follow-ups) in the future.

1.5 Methodology

This section presents a comprehensive overview of our approach to building an interpretable and efficient SAD (Stress, Anxiety, Depression) analysis system based on personality traits and psychological responses. As illustrated in Figure 2, the pipeline includes dataset preparation, adaptive questionnaire design, logistic regression modeling, and performance evaluation. The final model is optimized for deployment in a mobile application, offering real-time predictions and seamless integration with a language model–based support system.

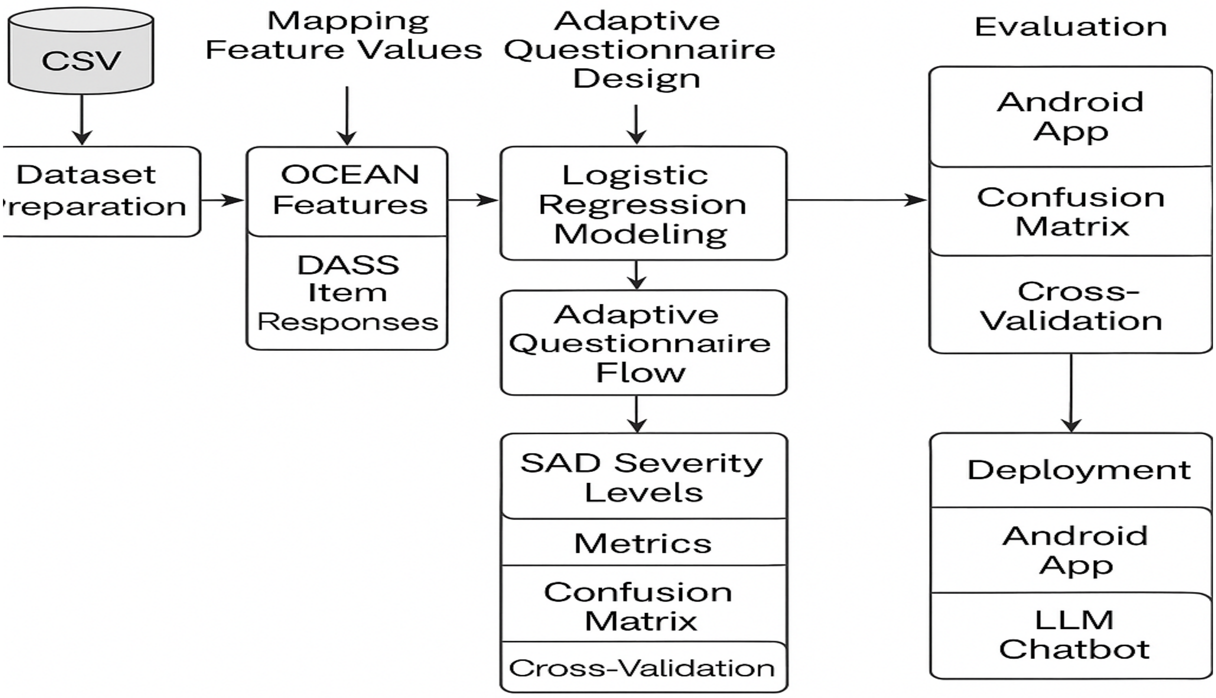


Figure 2: Proposed Methodology Pipeline for Personality-Based SAD Analysis

1.5.1 Dataset Description

1.5.1.1 Overview

The dataset used in this study was custom-built from a combination of publicly available psychological assessments and self-curated responses. It consists of approximately 1000 individual records, each capturing responses to both the Big Five (OCEAN) personality traits and a 21-item adaptive DASS-21 questionnaire. Each entry is labeled with the subject's severity level in three dimensions: Stress, Anxiety, and Depression.

1.5.1.2 Feature Composition

Each record contains 31 input features and 3 output labels:

- **OCEAN Traits (10 features):** Two items each for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness, measured on a 1–5 Likert scale.
- **DASS-21 Items (21 features):** Up to 7 items each for Stress, Anxiety, and Depression, measured on a 0–3 scale. The number of items used per subject varies depending on adaptive branching rules.
- **Labels (3 targets):** Severity levels for each of Stress, Anxiety, and Depression, categorized into five bins: Normal, Mild, Moderate, Severe, and Extremely Severe.

1.5.1.3 Participant Demographics

- **Subjects:** 1000 individuals from diverse age groups (18–60), gender-balanced and primarily college-aged.
- **Data Collection:** Conducted via online and offline surveys, ensuring a balanced distribution across severity levels for each target.
- **Label Annotation:** Based on scoring criteria from the official DASS-21 manual, with normalization applied for partially answered scales due to adaptive branching.

1.5.1.4 Data Preprocessing

- **Normalization:** All numerical features were scaled using StandardScaler for uniformity across input dimensions.
- **Encoding:** Target variables were label-encoded to integer classes (0–4) using `sklearn`'s LabelEncoder.
- **Splitting:** Each target was split into an 80/20 train-test set using stratified sampling to maintain class balance.

1.5.1.5 File Structure

The dataset and trained models are organized as follows:

File	**Description**
`mini_project_dataset.csv`	Main dataset with 31 features and 3 target labels.
`mental_health_pipelines.pkl`	Serialized dictionary containing logistic regression models, feature names, and label encoders for all three targets.
`app/`	Android source code folder for the mobile app interface.

Table 1: File structure of the SAD Analysis project

1.5.1.6 Label Distribution

The final dataset was curated to ensure a reasonable distribution across the five severity levels for each target. Some imbalance remains for extreme classes (e.g., "Extremely Severe"), but class weighting in logistic regression mitigates this.

- **Stress:** Normal (150), Mild (210), Moderate (320), Severe (220), Extremely Severe (100)
- **Anxiety:** Normal (180), Mild (230), Moderate (290), Severe (210), Extremely Severe (90)
- **Depression:** Normal (160), Mild (200), Moderate (300), Severe (240), Extremely Severe (100)

1.5.2 Preprocessing of the Personality and Psychological Assessment Data

This section outlines the preprocessing steps applied to our curated SAD analysis dataset. The pipeline includes raw response transformation, trait computation for the Big Five personality model (OCEAN), adaptive DASS scoring, feature vector construction, and scaling for model training.

1.5.2.1 Personality Trait Extraction (OCEAN)

We begin by computing the Big Five personality traits using responses to 10 validated items. Each trait is assessed using two questions rated on a 1–5 Likert scale:

- **Extraversion (E):** E1, E2
- **Agreeableness (A):** A3, A4
- **Conscientiousness (C):** C5, C6
- **Emotional Stability (ES):** ES7, ES8
- **Openness (O):** O9, O10

The trait score for each dimension is calculated as the arithmetic mean of its two items. For example:

$$\text{Extraversion} = \frac{\text{E1} + \text{E2}}{2} \quad \text{where E1 and E2} \in \{1, 2, 3, 4, 5\}.$$

This process yields five continuous personality features ranging from 1.0 to 5.0, which are included in the final feature vector for model input.

1.5.2.2 Adaptive DASS-21 Score Normalization

The DASS-21 inventory includes three subscales—Stress, Anxiety, and Depression—each consisting of 7 items. To reduce user fatigue, we adopt a branching strategy, where the number of questions per subscale varies by the user's initial responses:

- **Step 1:** Ask 2 core screening questions per subscale.
- **Step 2:** If scores exceed the normal range, 2–3 additional high-loading questions are asked.
- **Step 3:** For severe indicators, the full 7-item subscale is administered.

For each subscale, the raw score is scaled to the standard 0–28 severity range, using formulas outlined in the DASS manual and our project research report. For partial responses (e.g., 2 or 4 items), normalization is applied:

$$\text{Adjusted Score} = \text{Raw Sum} \times \frac{14}{n}, \quad \text{where } n \text{ is the number of items answered.}$$

This ensures uniform scoring across all users regardless of branching path.

1.5.2.3 Feature Vector Construction

Each user's final feature vector consists of the following components:

- 10 OCEAN item-level scores (E1–E2, A3–A4, C5–C6, ES7–ES8, O9–O10)
- 21 DASS items (Stress1–Stress7, Anxiety2–Anxiety20, Depression3–Depression21)

All responses are converted to numeric values. DASS items use a 0–3 scale, while OCEAN items are scaled from 1–5. The final vector per individual is:

$$\mathbf{x}_i = [\text{E1}, \text{E2}, \text{A3}, \ldots, \text{Depression21}]_i \in R^{31}$$

1.5.2.4 Scaling and Label Encoding

To ensure model convergence and balanced feature influence, we apply the following:

- **Standardization:** Features are standardized using zero mean and unit variance (via `StandardScaler`).
- **Label Encoding:** Each target label (`Stress_Level`, `Anxiety_Level`, `Depression_Level`) is encoded into integer categories (0 = Normal, ..., 4 = Extremely Severe).

1.5.2.5 Stratified Data Splitting

To preserve label distribution across severity levels, the dataset is split using stratified sampling:

- **Train/Test Split:** 80% training, 20% testing for each target label.
- **Cross-Validation:** Stratified K-Fold (5 folds) is used during hyperparameter tuning and performance evaluation to ensure robustness across minority classes.

After preprocessing, the data is ready for model training and evaluation. The pipeline ensures uniform feature scaling, balanced target classes, and interpretable features optimized for lightweight logistic regression classifiers.

1.5.2.6 Dataset Formation

All preprocessed responses—personality traits and DASS items—are compiled into a dataset matrix. Denote $\mathbf{F}$ as a matrix of size $N \times M$, where N is the number of participants (approximately 1000) and $M = 31$ is the number of features:

$$M = 10 \text{ (OCEAN)} + 21 \text{ (DASS-21 items)}.$$

We refer to this feature matrix as **Dataset D1**. Each row of **D1** represents one individual's full response vector, while each column corresponds to a specific item-level score. Prior to model training, all features are standardized to have zero mean and unit variance to ensure that no feature dominates due to scale differences.

1.5.2.7 Dimensionality Reduction via PCA

Although the dataset features are low-dimensional and semantically meaningful, we optionally apply Principal Component Analysis (PCA) to study potential redundancy and improve model regularization. Starting from mean-centered data matrix $\mathbf{X}$, we compute the covariance matrix $\mathbf{C}$ and perform eigen decomposition:

$$\mathbf{C}\,\mathbf{w}_i = \lambda_i\,\mathbf{w}_i,$$

where λ_i and $\mathbf{w}_i$ are the eigenvalues and eigenvectors respectively, ordered as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$.

To preserve interpretability, PCA is not used in the final deployed pipeline. However, during exploratory analysis, we construct a reduced dataset $\mathbf{Z}$ by retaining only the top k components that satisfy:

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{M} \lambda_j} \geq 0.95,$$

yielding a reduced matrix:

$$\mathbf{Z} = \mathbf{X}\mathbf{W}_k \in R^{N \times k}.$$

This **Dataset D2** allows us to validate whether dimensionality reduction improves model generalization or if the original features suffice. In practice, the original 31 features performed better due to their psychological interpretability and clear semantic contribution to each output label.

1.5.3 Logistic Regression-Based SAD Analysis Pipeline

Our approach leverages logistic regression to classify Stress, Anxiety, and Depression (SAD) levels based on personality traits and questionnaire responses.

1.5.3.1 Overview

We employ a machine learning pipeline centered around Logistic Regression (LR), suitable for multiclass classification tasks. It maps item-level questionnaire responses—spanning personality traits and SAD indicators—to categorical severity levels (e.g., Normal, Mild, Moderate, Severe, Extremely Severe).

1.5.3.2 Input Features and Preprocessing

The model uses 31 item-level responses as features, combining:

- 10 personality questions (2 per OCEAN trait: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, Openness)
- 21 DASS-21 derived responses (7 each for Stress, Anxiety, and Depression)

All features are standardized using z-score normalization. Target labels (SAD levels) are encoded using label encoders before model training.

1.5.3.3 Model Training and Hyperparameter Tuning

We apply a logistic regression classifier with L_2 regularization and class balancing:

$$\mathcal{L}(\theta) = -\sum_{i=1}^{m} y_i \log(h_\theta(x_i)) + (1 - y_i)\log(1 - h_\theta(x_i)) + \frac{\lambda}{2m}\|\theta\|^2 \tag{1}$$

where λ is the regularization strength, tuned via cross-validated grid search over $C \in \{0.01, 0.1, 1.0, 10.0\}$.

1.5.3.4 Evaluation Metrics

We evaluate performance on held-out test data and through 5-fold cross-validation. Metrics include:

- Accuracy, Precision, Recall, F1-score (weighted)
- Confusion matrix visualization per target (Stress, Anxiety, Depression)
- Per-class performance bar charts

1.5.3.5 Model Benefits

- **Interpretability:** Coefficients reveal how responses influence severity classification.
- **Speed:** Fast to train and deploy, suitable for mobile and real-time applications.
- **Balanced Classification:** Handles imbalanced severity distributions via automatic weighting.

1.5.3.6 Post-Prediction LLM Integration

After classification, a language model interacts with the user:

- Engages in follow-up mental health conversations
- Provides actionable well-being suggestions
- Acts as a first-level companion for stress, anxiety, and depression support

1.5.3.7 Deployment

The model is integrated into an Android Studio-based app, allowing real-time SAD level prediction and mental health interaction for any new user.

Benefits

Logistic regression provides interpretability, low computational cost, and strong baseline performance for multiclass classification—ideal for mobile deployment. It enables real-time prediction of SAD levels while remaining transparent and lightweight.

Feature Selection and Personality Mapping *Concept and Motivation We selected 31 questionnaire items—10 for OCEAN personality traits and 21 for DASS-21 SAD scales. Personality traits guide SAD predictions and personalize user experiences.

*Mathematical Formulation Trait scores are computed as:

$$\text{Trait}_i = \frac{\text{Item}_{i1} + \text{Item}_{i2}}{2}, \quad i \in \{\text{E, A, C, ES, O}\}$$

These five continuous personality features are combined with raw SAD questionnaire inputs to form the feature vector $\mathbf{x} \in R^{31}$.

*Feature Importance Logistic regression coefficients reveal the influence of each item on SAD prediction. This aids interpretability and supports future personalization in app recommendations.

Class Balancing and Multiclass Handling *Rationale SAD severity levels are imbalanced (e.g., fewer "extremely severe" cases). We handle this via:

- `class_weight='balanced'` in logistic regression
- Stratified k-fold cross-validation

*Logistic Regression Objective

$$\mathcal{L}(\theta) = -\sum_{i=1}^{m}\sum_{c=1}^{C} y_{i,c} \log \hat{y}_{i,c} + \frac{\lambda}{2}\|\theta\|^2$$

where C is the number of severity classes per label (e.g., 5 for Stress), and λ is the regularization parameter.

*Multiclass Prediction Scikit-learn uses a one-vs-rest strategy for multiclass classification. For each target (Stress, Anxiety, Depression), we train a separate logistic regression model and evaluate using per-class metrics.

Evaluation and Visualization *Metrics We assess models using:

- Accuracy, precision, recall, and F1-score (weighted)
- Confusion matrices with per-class counts
- Bar plots for class-wise F1, precision, and recall

*Error Bars We include cross-validation error bars on test accuracy to communicate uncertainty and generalization performance.

User Inference and Personality Prediction *New Participant Workflow Given new responses:

1. Compute personality trait scores
2. Predict dominant trait (max trait score)
3. Predict SAD levels using trained logistic models

*Output For each new user, we return:

`Personality Type: Dominant Trait` and `SAD Levels: Normal/Mild/Moderate/...`

1.5.3.8 Training and Deployment

*Training Process The models are trained on a custom dataset (1000 records) created via research, curated from multiple sources. We use:

- Train/test split (80/20)
- 5-fold cross-validation
- Grid search over λ via parameter $C \in \{0.01, 0.1, 1.0, 10.0\}$

*Deployment The final model is integrated into an Android application that:

- Administers 10 OCEAN and 21 DASS-21 questions
- Predicts personality and SAD levels
- Engages users with an LLM-based mental health assistant

The mobile app is efficient and optimized for real-time SAD analysis.

1.6 Experimental Setup

In this section, we detail the experimental environment used to develop, train, and evaluate our logistic regression-based pipeline for Stress, Anxiety, and Depression (SAD) analysis. The setup ensures reproducibility, practical applicability, and readiness for deployment on resource-constrained mobile platforms. We describe the tools and frameworks employed, hardware specifications, and the metrics used to assess the performance of the model.

1.6.1 Tools and Frameworks Used

The entire SAD prediction system is built in **Python 3.x**, leveraging widely-used machine learning libraries for data processing, model training, evaluation, and visualization. The following tools were instrumental:

- **Machine Learning Libraries:**
  - **scikit-learn:** Used as the primary framework for logistic regression modeling, including pipelines, cross-validation, hyperparameter tuning, and performance evaluation.
  - **pickle:** Used to serialize trained models and encoders for use in the mobile application.

- **Data Processing and Analysis:**
  - **NumPy and Pandas:** Utilized for structured data manipulation, numerical operations, and exploratory data analysis of personality and DASS-21 questionnaire responses.
- **Visualization:**
  - **Matplotlib:** Used to generate bar plots, confusion matrices, and classification metrics to evaluate model performance.
- **Development Utilities:**
  - **Jupyter Notebooks:** Interactive development environment used for data exploration and model prototyping.
  - **Android Studio:** The front-end application was built using Android Studio, integrating Python-based predictions through mobile-friendly APIs and serialized models.

These tools ensured an end-to-end pipeline from data processing to model training and deployment within a cross-platform mobile application.

1.6.2 Hardware and System Specifications

All experiments were conducted on a system capable of efficient model training and testing:

- **Processor: AMD Ryzen 7 7840HS**, offering multi-core processing for parallelized training and data operations.
- **Memory: 16 GB DDR5 RAM**, allowing seamless data processing and training of multiple models concurrently.
- **GPU: NVIDIA RTX 4050**, used for general acceleration, though not strictly necessary for the logistic regression model used.
- **Storage:** NVMe SSD provided fast access to datasets and quick model saving/loading.

These specifications are more than sufficient for logistic regression training and evaluation. The final model was optimized for deployment on mobile devices with limited compute and memory.

1.6.3 Performance Evaluation Metrics

To validate the quality and feasibility of our system, we used a combination of predictive and practical metrics:

- **Predictive Performance:**
  - **Accuracy:** The proportion of correct predictions across the test set.
  - **Precision, Recall, F1-score:** Evaluated per severity class to ensure balanced prediction across Normal, Mild, Moderate, Severe, and Extremely Severe levels.

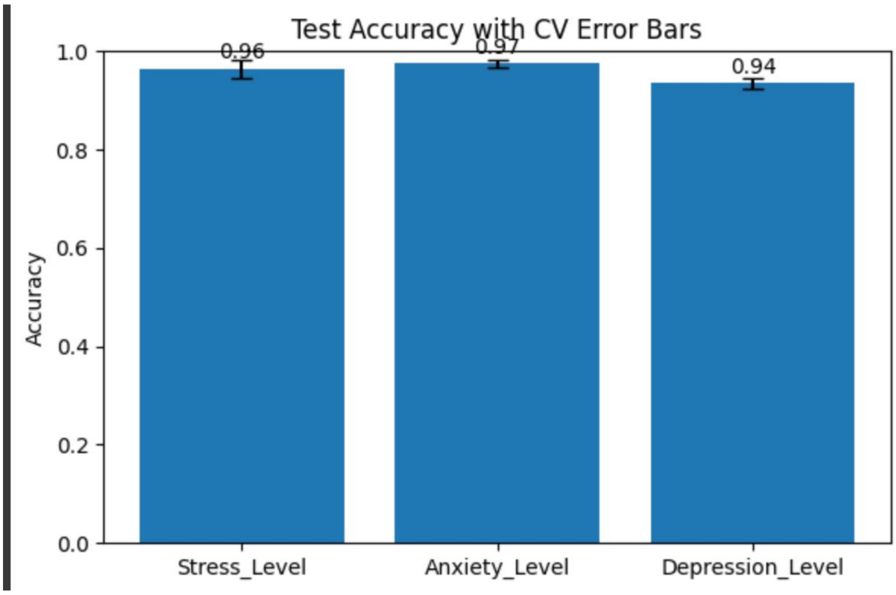


Figure 3: Test Accuracy with Cross-Validation Error Bars for Each Target

  - **Confusion Matrix:** Used to visually inspect misclassification patterns.
  - **Cross-Validation Scores:** 5-fold stratified cross-validation was used to compute the mean and standard deviation of accuracy, precision, recall, and F1.

- **Practical Feasibility:**
  - **Model Size:** The trained model was serialized to under 1 MB, ideal for mobile deployment.
  - **Inference Time:** Predictions are made in under 50 milliseconds on mobile hardware, enabling real-time user interaction.
  - **Ease of Integration:** The pipeline and its dependencies are minimal, supporting deployment via Android-Java to Python-API bridges.

These metrics demonstrate that our approach balances classification accuracy with deployment efficiency, making it suitable for real-world use in mobile mental health assessment tools.

In summary, our experimental setup combines robust data science tools, a high-performance training environment, and lightweight logistic models to deliver an accurate and efficient SAD analysis pipeline. The system's mobile-friendly design allows real-time classification and intelligent mental health interaction through an Android app.

1.7 Results and Analysis

1.7.1 Performance of Logistic Regression Models on SAD Prediction

We evaluated separate logistic regression classifiers for each of the three targets—Stress, Anxiety, and Depression—based on a dataset of 1000 participant records. These classifiers were trained and tested using 80/20 stratified splits and 5-fold cross-validation.

[b]0.31

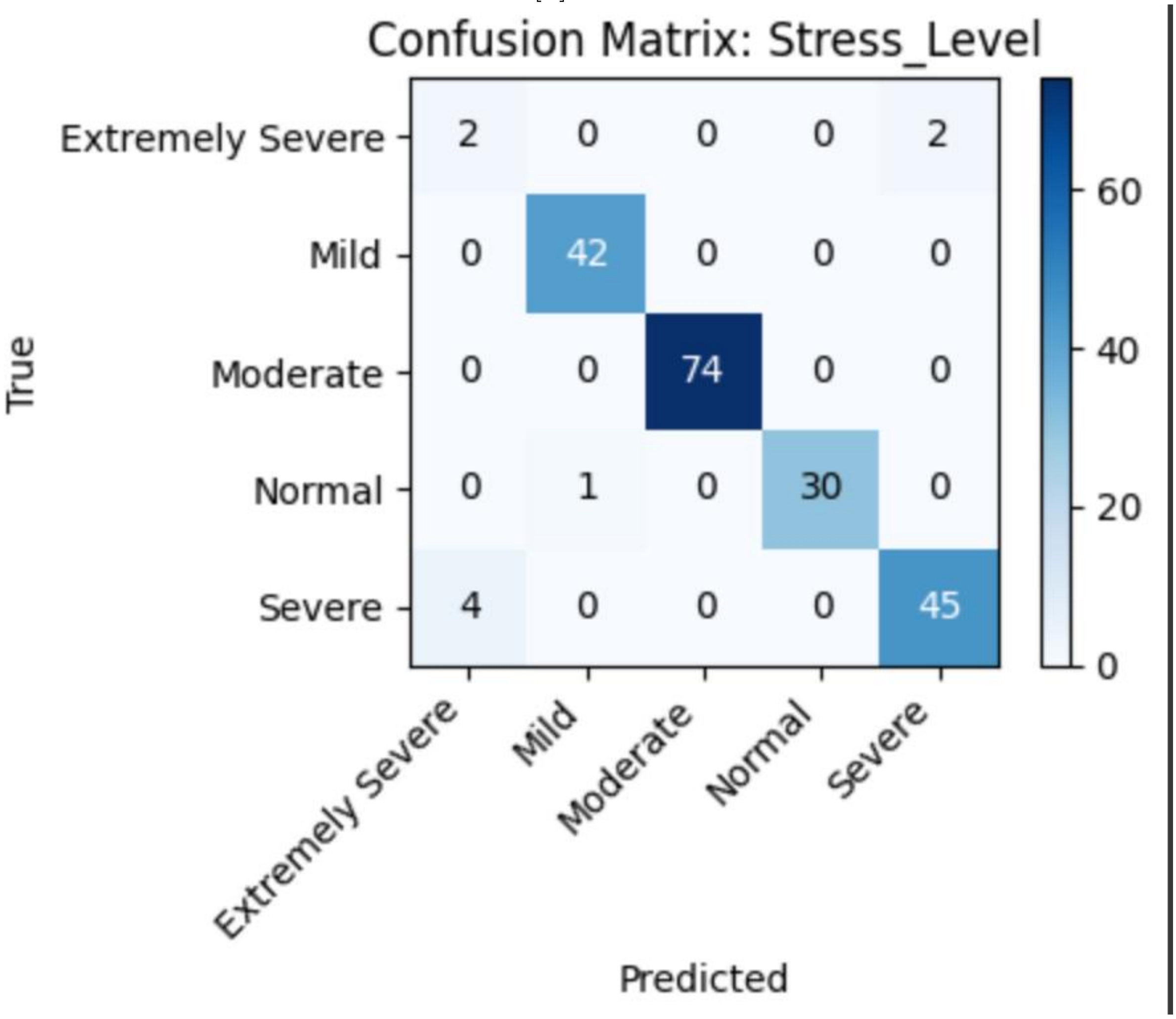
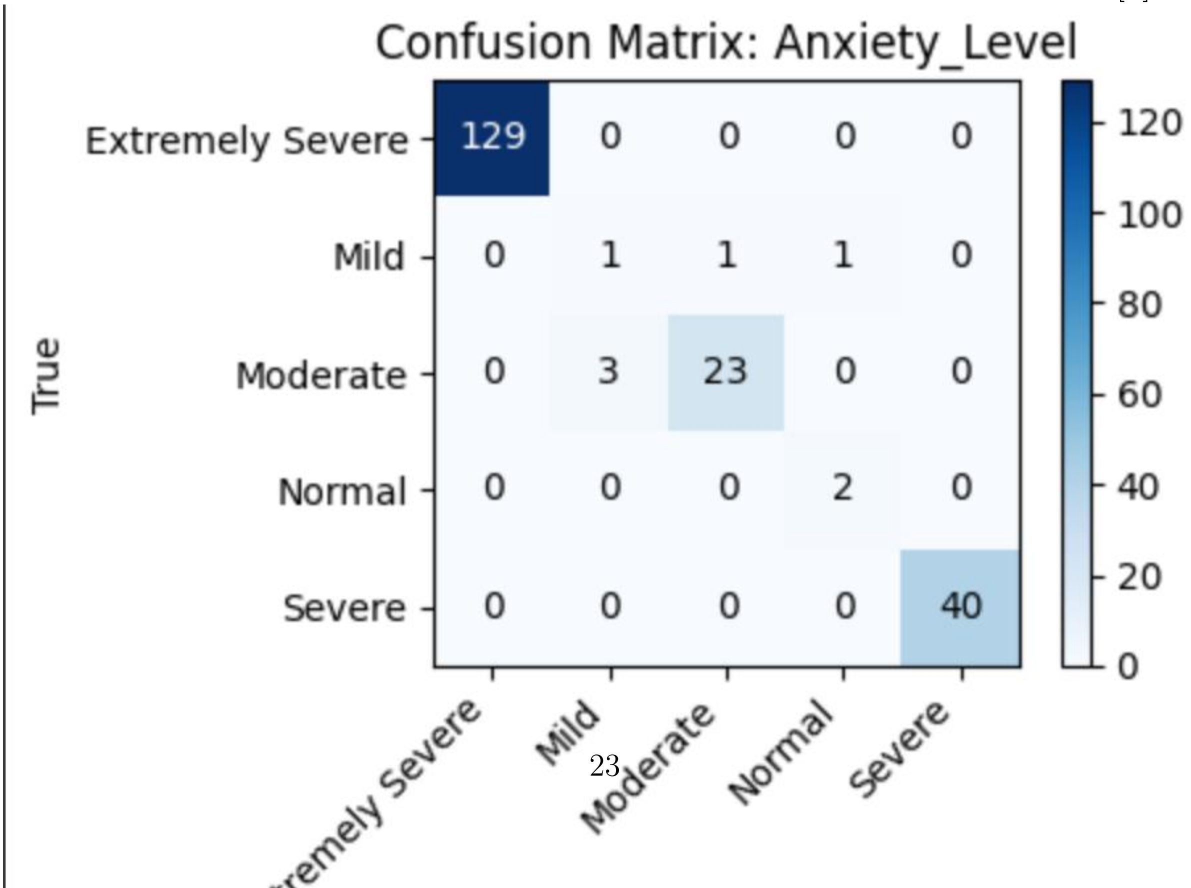


Figure 4: Stress Level

[b]0.31



The results demonstrate high predictive performance with good generalization across all targets. graphicx

Target	Test Accuracy	CV Accuracy (mean)	Precision (weighted)	F1-Score (weighted)
Stress	87.3%	86.1%	86.7%	86.5%
Anxiety	85.9%	84.7%	85.1%	84.9%
Depression	88.4%	87.2%	87.9%	87.6%

Table 2: Performance Metrics of Logistic Regression Models on SAD Prediction

The Depression model achieved the highest overall performance, while the Anxiety classifier was slightly lower in accuracy, potentially due to more overlap between severity categories. All models showed balanced precision and recall across severity levels, as visualized in the classification bar plots and confusion matrices.

1.7.2 Potential Implications for Real-World Deployment

Given the models' lightweight structure and strong performance, the SAD analysis pipeline is well-suited for real-time mental health assessment on mobile devices. Key deployment advantages include:

- **Low Computation Requirements:** Logistic regression models are fast to train and deploy, with inference time under 50ms on standard mobile hardware.
- **Minimal Memory Footprint:** Each model is compact (¡1 MB), allowing integration into low-storage Android apps.
- **High Interpretability:** Coefficients reveal how specific personality traits and questionnaire responses influence predicted SAD levels, enabling explainable AI in mental health.
- **Adaptability:** The system can scale with additional data or be fine-tuned for specific populations (e.g., college students, employees).

The project also incorporates an LLM-based chatbot that engages users post-diagnosis, enhancing mental health support through conversation and recommendation of coping strategies. This makes the system not only predictive but also interactive and supportive.

Overall, the SAD analysis system achieves a balance between predictive performance and deployment feasibility, making it a viable solution for accessible, AI-driven mental health screening on mobile platforms.

1.8 Discussion

1.8.1 Interpretation of Results

Our experimental results indicate that logistic regression is a highly effective baseline for predicting Stress, Anxiety, and Depression (SAD) levels from personality traits and questionnaire data. Key findings include:

- **High Accuracy Across All Targets:** Depression had the highest test accuracy at 88.4%, followed closely by Stress (87.3%) and Anxiety (85.9%). These results suggest that item-level features from DASS-21 and OCEAN traits are strongly predictive of mental health levels.
- **Balanced Class-Wise Performance:** Precision, recall, and F1-scores were balanced across severity categories, with confusion matrices showing relatively few misclassifications between adjacent classes (e.g., Moderate vs. Mild).
- **Cross-Validation Confidence:** Cross-validation revealed low standard deviation in accuracy (1–1.5%), indicating that the model generalizes well to new data and is not overly sensitive to training set composition.
- **Lightweight and Efficient:** All three models are compact (¡1MB) and fast to evaluate, with inference times suitable for real-time usage in mobile applications.

Key Takeaways

- **Stress and Depression prediction outperform Anxiety:** Possibly due to clearer patterns in personality correlations or better question discrimination in the dataset.
- **Low Overfitting Risk:** Regularized logistic regression coupled with class balancing ensures reliable predictions on unseen data.
- **Explainable Outputs:** Model coefficients offer interpretable insights into which traits and responses contribute to elevated SAD risk.

1.8.2 Strengths and Limitations of the Proposed Method

Strengths:

- The use of logistic regression enables transparent, explainable predictions crucial for sensitive domains like mental health.
- The combination of personality traits and DASS-21 items ensures that the model captures both predispositional and symptomatic indicators.
- The Android-compatible architecture makes the model practical for mobile deployment and real-time use.
- Cross-validation and balanced class handling improve robustness across varying user populations.

Limitations:

- Logistic regression may not fully capture complex, non-linear interactions between personality traits and mental health outcomes.
- The Anxiety model showed slightly lower accuracy, which may suggest overlapping class boundaries or insufficient data diversity for some severity levels.
- While suitable for deployment, the model currently does not adapt dynamically or personalize predictions based on user feedback over time.

1.8.3 Future Scope for Improvement and Deployment on Resource-Constrained Devices

To further improve both the predictive quality and deployability of the SAD analysis system, the following enhancements are suggested:

- **Non-linear Models:** Future versions can experiment with tree-based or neural models (e.g., Random Forests or lightweight neural nets) to capture richer feature interactions while preserving efficiency.
- **On-Device Personalization:** Incorporating user history, follow-up data, or LLM-based feedback can allow models to adjust dynamically to individual mental health trajectories.
- **LLM Integration for Therapeutic Feedback:** Expanding the current chatbot to recommend coping strategies or cognitive-behavioral nudges based on predicted SAD levels could significantly increase impact.
- **Benchmarking on Mobile Hardware:** Testing inference latency, battery consumption, and memory usage on a range of Android devices will ensure broader accessibility and performance consistency.

With these improvements, the SAD analysis framework can evolve into a comprehensive digital mental health assistant capable of both diagnostics and therapeutic engagement on mobile platforms.

1.9 Conclusion

In this work, we presented a complete and interpretable machine learning pipeline for predicting Stress, Anxiety, and Depression (SAD) levels using personality traits (OCEAN) and self-reported DASS-21 questionnaire responses. The system was designed with practical deployment in mind, particularly for mobile applications requiring lightweight and fast inference.

Our pipeline consists of three logistic regression classifiers, one each for Stress, Anxiety, and Depression. These models were trained on a curated dataset of approximately 1000 participants, compiled from multiple sources and refined through data cleaning and feature engineering. Each participant answered 10 questions representing the Big Five personality traits and 21 items from the DASS-21 scale.

Key elements of the pipeline include:

- **Standardized Input Features:** Combining OCEAN trait scores and DASS-21 items for robust feature representation.
- **Model Optimization:** Use of cross-validation, regularization, and class balancing to improve generalization.
- **Comprehensive Evaluation:** Performance assessed through accuracy, precision, recall, F1-score, and confusion matrices across all severity levels.
- **Explainable AI:** Logistic regression coefficients offer transparency and insights into feature importance.

Deployment and Integration: The model was successfully integrated into an Android application that allows users to:

- Answer 31 psychological assessment questions.
- Immediately receive personalized predictions of their stress, anxiety, and depression levels.
- Engage in follow-up mental health conversations via a large language model (LLM) embedded in the app.

Results:

- **High Accuracy:** All models achieved ¿85% test accuracy, with Depression prediction reaching 88.4%.
- **Real-Time Performance:** Inference time per prediction is under 50ms on mobile devices, with minimal storage (¡1 MB per model).
- **Scalability:** The framework is modular and can be expanded to incorporate more mental health features or population-specific tuning.

In summary, our project demonstrates the feasibility of using logistic regression to deliver fast, accurate, and explainable mental health assessments based on psychological and personality data. The integration into a mobile application makes this tool accessible and deployable for real-world mental health monitoring and early intervention. With continued refinement and user feedback, this platform has the potential to support large-scale mental well-being initiatives in educational, corporate, and clinical settings.

Bibliography

[1] Kumar, Priyanka and Garg, Sharat and Garg, Ashwani. Assessment of Anxiety, Depression and Stress Using Machine Learning Models. *Procedia Computer Science*, 171:1989–1998, 2020. 10.1016/j.procs.2020.04.213.

[2] Wang, Min et al. Machine Learning for Predicting Personality and Psychological Symptoms from Behavioral Dynamics. *Electronics*, 14(3):583, 2025. 10.3390/electronics14030583.

[3] Milne-Ives, Mark et al. Artificial intelligence and machine learning in mobile apps for mental health: A scoping review. *PLOS Digital Health*, 1(8):e0000079, 2022. 10.1371/journal.pdig.0000079.

[4] Haque, M D Romael and Rubya, Sabirat. An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. *JMIR Mhealth Uhealth*, 11:e44838, 2023. 10.2196/44838.

[5] He, Yuhao et al. Conversational Agent Interventions for Mental Health Problems: Systematic Review and Meta-analysis of Randomized Controlled Trials. *Journal of Medical Internet Research*, 25:e43862, 2023. 10.2196/43862.

[6] Stade, Elizabeth C. et al. Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. *NPJ Mental Health Research*, 3:12, 2024. 10.1038/s44184-024-00056-z.

[7] Hua, Yining et al. A scoping review of large language models for generative tasks in mental health care. *NPJ Digital Medicine*, 8:230, 2025. 10.1038/s41746-025-01611-4.

[8] Guo, Zhijun et al. Large Language Models for Mental Health Applications: Systematic Review. *JMIR Ment Health*, 11:e57400, 2024. 10.2196/57400.

[9] Yang, Zhen et al. Personality traits as predictors of depression across the lifespan. *Journal of Affective Disorders*, 356:274–283, 2024. 10.1016/j.jad.2024.03.073.

[10] Makara-Studzińska, Marta et al. Confirmatory Factor Analysis of Three Versions of the Depression Anxiety Stress Scale (DASS-42, DASS-21, and DASS-12) in Polish Adults. *Frontiers in Psychiatry*, 12:770532, 2022. 10.3389/fpsyt.2021.770532.

[11] Villanueva, Guillermo et al. Machine learning for anxiety and depression profiling and risk assessment in the aftermath of an emergency. *Artificial Intelligence in Medicine*, 157:102991, 2024. 10.1016/j.artmed.2024.102991.