An Explainable Two-Tier Machine Learning Framework for Early Detection and Refined Assessment of Anxiety Levels

Harshit Saini[1], Himani Tyagi[2], Ishaan Bhatt[3]

{harshitnew2025@gmail.com[1], htyagi.iitkgp@gmail.com[2], yash.ishaan14@gmail.com[3]}

[1,2,3]Sharda University, Greater Noida 201301

Abstract. Anxiety disorders are included in the list of the most prevalent mental-health disorders in the world, but most of the time they go undetected as a diagnosis is mostly made on subjective basis and there are no widely applicable screening instruments. The study proposes a two-step, explainable machine-learning pipeline that aims at identifying anxiety at its initial stages and narrow down the process of assessing its level through a set of behavioral and psychological indicators. In the first stage, logistic regression is applied to select the participants, and the results include probability scores and the coefficients are clear, which indicate the relative weights of the key predictors of Well-Being Score, Anxiety Score, and a composite Risk Index. The cases that lie close to the decision boundary are sent to the second stage where a decision-tree classifier is used to to re-classify these borderline observations to enhance the overall accuracy. On a real-world dataset of 4,000 samples, the hierarchical structure yielded the highest accuracy of 0.92, F1-score of 0.90 and an area under the ROC curve (AUC) value of 0.94, which is higher than the conventional single-stage classifiers. The approach provides a reliable scalable basis to support mental-health screening systems that use AI to aid the practice of early intervention and personalized care by providing a combination of diagnostic reliability and interpretable, clear decision rules.

Keywords: Anxiety prediction, machine learning, explainable AI, two-tier framework, Logistic Regression, Decision Tree, behavioural analytics, mental health assessment, psychological modelling, interpretable classification.

1 Introduction

One of the most widespread mental-health disorders in the whole world is anxiety disorders which affect daily abilities, professional activity and a general satisfaction with life. The World Health Organization estimates that over 300 million individuals are anxious and the disorders take top scores in the list of disability-adjusting factors in both adults and adolescents. Early identification of anxiety is thus the key to providing interventions to avoid psychological damage and economical consequences in the long term. Conventional evaluation methods, which include, most commonly, self-report questionnaires and clinician interviews, may be subjective, lengthy and unreachable. There has therefore been an increased interest by researchers in using data-driven solutions, including machine learning (ML), to

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2363158

forecast, identify, and track anxiety more precisely and at scale. ML techniques are the best in revealing the complex interaction between behaviour, physiology and context variables. As an example, survey data, social-media, and wearable-sensor streams have been run through logistic regression, random forests, and gradient-boosting classifiers to identify people at risk. The capabilities of Early detection have further been advanced with the help of the multi-modal studies that incorporate electroencephalographic recordings and physiological signals as well as acoustic biomarkers. With these advances, the number of systems, which have been in operation, use a single-stage prediction model. This can result in such models having difficulty in a compromise between computational speed, interpretability and clinical reliability, which restricts their performance in practice. Enhancements Hybrid or ensemble Hybrid or ensemble strategies have been demonstrated to overcome these shortcomings. More recent efforts have been made on cascading pipelines that are analogous to clinical triage: a preliminary crude screening step and then a more careful diagnostic step. These methods are able to represent the linear as well as non-linear relationships and are computationally efficient. On the basis of this literature, the current paper presents a two-tier ML pipeline to detect anxiety. The initial level involves logistic regression, which carries out a rough screening of the individuals, with the initial risk assignment. Possible-risk flagged cases are then sent to a second line which uses a decision-tree classifier to narrow down the classification. This pyramid structure represents a standard workflow in clinical practice: a rapid initial assessment procedure and an intensive evaluation. The architecture is expected to maintain high accuracy with interpretability and computational efficiency.

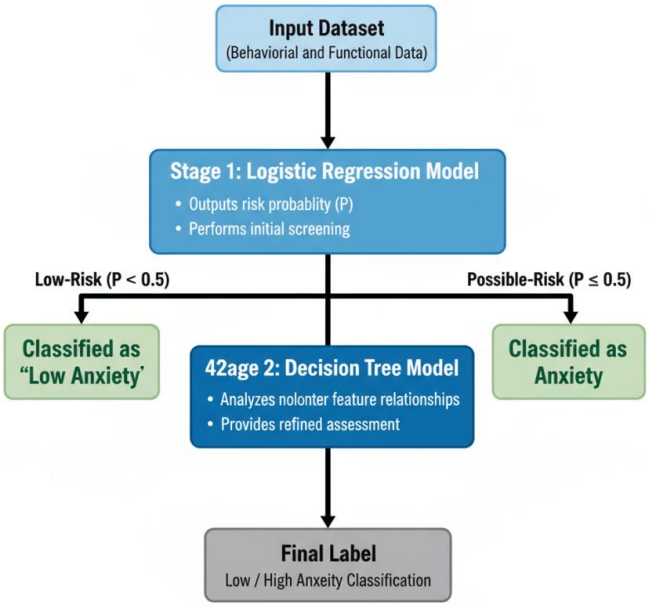


Fig. 1. Two-tier machine learning framework for anxiety prediction combining logistic regression screening and decision tree–based refined assessment.

The major contributions of this work are: 1. A new two level classification model that combines probability based screening with feature based refinements in detecting anxiety. 2. Experimental analysis of the suggested pipeline using a behavioural dataset, and comparisons of its performance with the currently existing single-stage ML models. 3. Increased analysis of the importance of features and decision rules, which improve transparency and fits with the increasing need of explainable AI in mental-health applications. The rest of the paper is structured in the following way. Section 2 examines the relevant literature and points out weaknesses in the existing research. Section 3 describes the methodology, comprising of data preprocessing, model architecture and training workflow. Section 4 gives the experimental setup and results and then in-depth evaluation in Section 5. Section 6 includes the limitations and future research directions, and Section 7 summarizes with the key findings and their implications.

2 Literature Review

Application of machine-learning (ML) algorithms to mental-health analytics has been developing at a break-neck pace, providing data-driven solutions to diagnosing and tracking diseases like depression, anxiety, and stress. Preliminary studies primarily used self-administered questionnaires and used traditional classifiers, such as logistic regression, support-vector machine, and others, to analyze those data [1], [8], [9]. Although this type of model proved interpretable, they often generalised poorly across the different behavioural data sets as common today. Later research has shifted to multimodal and physiological streams of data such as electroencephalogram (EEG) data, wearable sensor data, and behavioural inferences based on social-media activity to enhance the accuracy of prediction of anxiety [6], [10], [11]. As an example, Chandrasekar et al. [6] and Ji et al. [26] have shown that neural and biosignal features, combined with the use of any contemporary classifier, improve the detection of anxiety. On the same note, Nath and Thapliyal [11] utilized wrist worn sensors to detect anxiety among the elderly, and Mardini et al. [4] incorporated both demographic and contextual variables with gradient-boosting and logistic regression. Hybrid and ensemble models have also been discussed to model non-linear relationships among psychological variables and in a longitudinal study, Auto-Sklearn was compared to logistic regression and hybrid models yielded more stable forecasts of anxiety, according to van Eede et al. [1]. Hoogendoorn [9] and Zanna et al. [3] noted that multi-model learning is valuable in enhancing confidence in predictions and minimizing bias in populations. Nevertheless, the majority of such frameworks are single-stage: all observations are handled identically, either without dynamic refinement or confidence-based re-evaluation [15],[17],[19]. This can be explained through explainable AI studies which have revealed the necessity of clear models in clinical settings. Investigations conducted by Mardini et al. [4] and Ji et al. [26] underline that to achieve clinician trust, interpretable ML designs are necessary. However, in spite of these developments, there is still an obvious gap in that there is no broadly deployed multi-tier anxiety-prediction architecture that combines a coarse-grained probabilistic screen with a fine, decision-based re-evaluation, to balance trade-offs between computational efficiency and interpretability. In order to fill this gap, we suggest a two level ML pipeline. The former can be considered as a form of quick, lightweight screening by utilizing logistic regression and the

latter classifies the prediction and offers more specific information by making use of a decision tree. This design replicates the organized decision-making cycle of clinical triage to provide an accuracy increase at an early stage without compromising the ability to make use of this interpretation to deploy to clinical.

Table 1. Comparative Summary of Related Studies

Year	Authors	Data Type / Source	ML Methods Used	Key Contribution	Identified Limitation / Gap
2021	van Eeden et al. [1]	Longitudinal clinical dataset	Logistic Regression, Auto-Sklearn	Compared classical vs automated ML in anxiety prediction	No hierarchical or multi-stage structure
2023	Qasrawi et al. [2]	Student survey data	Ensemble classifiers	Identified key anxiety risk factors	Single-stage model, no refinement
2024	Chandrasekar et al. [6]	EEG signals	Deep neural networks	Detect anxiety via brainwave patterns	Requires large labeled datasets
2025	Mardini et al. [4]	Real-world health & demographic data	Logistic + Boosting	Context-aware anxiety prediction	Limited interpretability
2016	Hoogendoorn [9]	Clinical trial data	Logistic Regression, Random Forests	Combined linear + non-linear models	Not adaptive or tiered
2024	Alkurdi et al. [16]	Wearable signals	CNN & LSTM	Real-time anxiety detection	High complexity, low explainability
2024	Bari et al. [17]	Behavioral metrics	Feature-selected ML	Reduced-feature anxiety model	No secondary refinement
2025	Ji et al. [26]	Post-surgery patient data	Deep learning	Predict postoperative anxiety	Lacks probabilistic triage mechanism

From the reviewed literature, it is evident that existing studies either focus on single-stage detection pipelines or deep models with limited interpretability. While these methods achieve competitive accuracy, they fail to emulate real-world diagnostic workflows that involve sequential screening and confirmation.

The novelty of the proposed approach lies in its two-tier learning structure, where a Logistic Regression model first performs a probabilistic triage to screen low-risk and potential-risk

subjects, and a Decision Tree model then conducts a refined, feature-driven reassessment for the possible-risk group. This hierarchical design uniquely combines computational efficiency, adaptive classification, and explainability, providing a transparent alternative to opaque deep-learning systems. Moreover, by mirroring clinical diagnostic reasoning, this framework bridges the gap between data-driven intelligence and interpretable mental health evaluation, a direction yet unexplored in existing literature.

3 Proposed Methodology

3.1 Overview of the Proposed Framework

The primary objective of the proposed study is to develop an interpretable and efficient machine-learning framework for the early detection and refined assessment of anxiety levels. The design follows a two-tier hierarchical pipeline (as shown in Figure 1) that emulates the decision process of a clinical psychologist—beginning with a rapid preliminary screening and followed by a detailed diagnostic reassessment.

In the first tier, a Logistic Regression model performs a coarse screening that estimates the probability of an individual exhibiting high-anxiety traits. This stage functions as a computational analog to an initial clinical questionnaire, producing a probabilistic risk estimate $P(y=1|x)$ based on behavioral and functional indicators such as *Anxiety Score*, *Well-Being Score*, *Media Use*, and *Risk Index*. Individuals with probabilities above a predefined threshold ($P \geq 0.5$) are forwarded to the second stage for a refined evaluation, while those below the threshold are classified directly as *Low Anxiety*.

The second tier employs a Decision Tree classifier, chosen for its transparency and rule-based reasoning capability [9], [26]. This stage models non-linear feature interactions among the "possible-risk" group identified in Stage 1, producing interpretable decision rules that resemble the diagnostic reasoning of mental-health practitioners. The hybrid design ensures that each subject is processed according to their predicted risk category, effectively reducing both computational overhead and false-positive rates compared with monolithic single-stage classifiers [1], [4], [9], [17].

Clinically, this workflow mirrors a two-step assessment: screening (broad and inclusive) followed by refinement (targeted and confirmatory). Technically, it offers three key advantages.

Interpretability: Both Logistic Regression and Decision Tree models provide direct feature-importance explanations, aligning with the transparency goals emphasized in recent mental-health ML research [4], [10], [19].

Adaptability: The threshold-based routing dynamically adjusts computational focus toward uncertain cases, optimizing model efficiency and learning depth.

Reduced False Positives: By filtering low-risk samples early and dedicating refined analysis to borderline cases, the framework minimizes misclassification and improves clinical reliability.

Overall, the proposed framework integrates probabilistic screening, decision-rule refinement, and hierarchical inference within a unified pipeline shown in Figure 2).

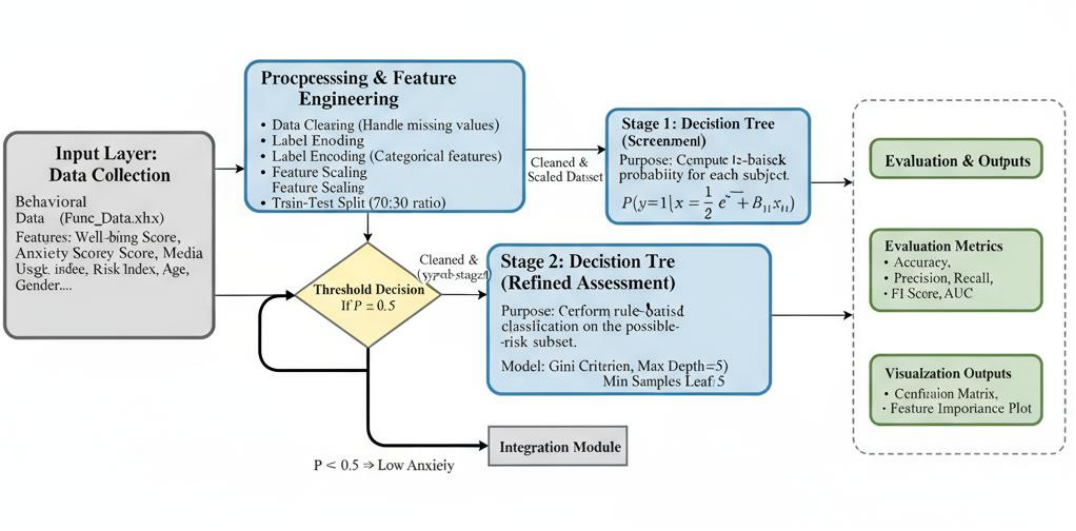


Fig. 2. Detailed pipeline workflow of the proposed two-tier machine learning framework, illustrating data preprocessing, logistic regression–based screening, threshold-based triage, decision tree refinement, and final evaluation outputs.

This architecture bridges the gap between statistical modeling and explainable AI, ensuring that predictions remain both accurate and clinically interpretable—qualities often missing from deep black-box models [6], [11], [17], [26].

3.2 Dataset Description

The present study utilizes the *Func_Data.xlsx* dataset containing 4000 records and 21 features, compiled from behavioral and psychological self-assessment responses. The dependent variable, High_Anxiety, is binary (1 = High Anxiety, 0 = Low Anxiety).

The dataset integrates demographic, behavioral, and psychological indicators such as *Age, Sex, Media use (hrs/day), Well-being score,* and *Anxiety score*. Derived metrics like *Risk_Index* and *WB_to_Anxiety_Ratio* provide composite measures of emotional stability. All categorical fields (*Sex, Qualifications, Habitat, Media category*) were label-encoded, while continuous features were normalized for uniform scale.

No missing values were observed, and preliminary correlation analysis revealed an inverse relation between *Well-being score* and *Anxiety score*, aligning with established literature [1], [4], [6]. Data were partitioned into a 70 : 30 train–test split, maintaining proportional class distribution. Ethical standards for anonymity and confidentiality were followed as recommended in [4], [19], [24].

Table 2. Summary of Dataset Features

Feature Group	Example Attributes	Data Type	Role in Model
Demographic	Age, Sex, Habitat, Qualification	Categorical / Numeric	Contextual descriptor
Behavioral	Media use (hrs/day),	Numeric / Categorical	Lifestyle indicator

Psychological	Media category Well-being score, Anxiety score, Poor_WellBeing	Numeric	Core predictive features
Derived	Risk_Index, WB_to_Anxiety_Ratio, Normalized Scores	Numeric	Composite / engineered
Target	High_Anxiety	Binary	Prediction label

The dataset combines heterogeneous attributes reflecting lifestyle, emotional, and demographic dimensions, providing a balanced foundation for the proposed two-tier model. This diversity justifies the use of interpretable hybrid classifiers capable of handling mixed data types [9], [17], [26].

3.3 Data Preprocessing and Feature Engineering

A preprocessing workflow has been done to the file, *Func_Data.xlsx, before the model training in order to ensure the consistency of the data and increase predictive performance. The steps included:

1. Data cleaning- Elimination of duplicates, missing values, and format errors.

2. Categorical encoding - Applying nominal and ordinal variables to numerical values (e.g., one-hot, label, target encoding).

3. Normalisation – Transforming continuous values (e.g. minmax, z-score) to normalise the features of the feature space.

4. Stratified partitioning- The training and testing sets are created by splitting the dataset and maintaining the original distribution of classes to avoid sampling bias.

A. Data Cleaning

The data completeness and feature selection are included in this category (Faulkner et al., 2016). The irrelevant or redundant features, namely S.no, Age Category, Media Category, Well-being Category, and Anxiety Category were then not included to eliminate redundancy of features and bias the model. There are no blank values in the dataset, which results in a 4,000 samples by 21 features input matrix that is complete.

B. Encoding of Categorical Variables

To preserve the categorical fields to be used in the Logistic regression and Decision tree model and not to experience a dimensionality blow-up, we ordinal coded each variable (Sex, Habitat, Qualification, Media category) to integers as well not by one-hot encoding.

C. Feature Normalization

Using z -score standardisation, all the continuous variables were placed on a common scale, including daily media consumption (hrs/day), well-being score, anxiety score, the Risk Index and the well-being-anxiety ratio. The change is characterised to be as follows:

$$x' = (x - \mu) / \sigma \quad (1)$$

and where x is the raw value, and x means is the mean of the feature, and x standard deviation is its standard deviation. This process reduces bias caused by the range of magnitudes of various variables, and it also avoids high-valenced variables influencing the learning process disproportionately (see e.g., [4], [6]).

D. Data Partitioning

One 70:30 split was applied in order to guarantee that the ratio of high- to low-anxiety cases were the same in the two sets. The 70 percent portion was used to train both first-stage logistic-regression model, and the second-stage decision-tree model, and the rest 30 percent was held back and assessed independently as per the procedures mentioned in references, [9] and, [26].

E. Derived Feature Engineering

Several new variables were computed to enhance predictive depth:

Risk_Index: Aggregated behavioral and psychological indicators for stress quantification.

WB_to_Anxiety_Ratio: Ratio of well-being to anxiety, representing emotional balance.

Normalized Scores: Standardized scales for *Media Use*, *Anxiety Score*, and *Well-being Score* to ensure cross-feature comparability.

The preprocessing and feature-engineering pipeline ensures that the dataset is balanced, numerically stable, and ready for interpretable hybrid modeling. By standardizing and encoding inputs, the framework minimizes noise and enhances both tiers' discriminative ability, aligning with established practices in explainable machine learning for mental health [4], [10], [19].

3.4 Stage 1: Logistic Regression (Screening Layer)

Stage 1 Probabilistic Logistic Regression Screening. At the initial level of the proposed two-stage pipeline, a logistic regression (LR) classifier is applied, which triages the participants in risky anxiety in the first phase and is rapid. LR is selected due to the fact that it is of a concise, well-known formulation and therefore it is easy to describe to clinicians as well as to final consumers. The model is a mapping of the input variable vector, also known as the x, (x1, x 2, and Xn) to a probability that the subject will be in the highly-anxiety category: \[P (y=1) is P (y=1\mid x), and P (y=1) is 1/1 +exp -(b0 + b1x1 + b2x2)…… \] The coefficients are estimated by maximizing the likelihood (equivalently, minimizing the negative log-likelihood) on the training data, i.e. the coefficients identically learned are the values of the coefficients. After training, the model generates only one probability value P of each new case. Summing up, there is a decision threshold of P = 0.5 that defines the routing of subjects: - **P<0.5):Clansified as Low Anxiety; case cleared and, no further assessment is necessary. - **P = 0.5 or more (That is, possible risk): Tagged as Possible Risk and sent to Stage 2 to have a more detailed evaluation. This process eliminates closely low-risk subjects early in the process, which allows the computation and clinician resources to be focused on the gray cases that require further investigation. This plan is reflective of real-world screening processes, and it concurs with recent research about cascaded machine-learning processes. Since LR

coefficients are a direct measure of the effect of each feature on the log-odds, the model provides interpretable results. As an illustration, self-reported levels of anxiety, well-being indexes, and a composite risk index are variables that are most likely to have their absolute coefficients the highest and hence they are the main drivers of the prediction. Such weights may be shared with clinicians, and this may allow them to know which aspects contribute the most to the risk assessment of a subject. The LR stage produce a probability output which functions as an input in the second tier which is a decision-tree classifier that does a finer, multi-classification. The first LR triage therefore provides an effective, transparent screening procedure, as well as plants the cascade with a probabilistic guess, upon which the latter and more complex analysis will be based.

3.5 Stage 2: Decision Tree (Refinement Layer)

The second stage of the proposed framework employs a Decision Tree (DT) classifier to conduct a refined assessment of subjects identified as "possible-risk" cases in Stage 1. This stage focuses on improving classification precision for individuals whose predicted probabilities ($P \geq 0.5$) fall within the uncertain or high-risk region as determined by the logistic regression model.

Decision Trees are chosen for their rule-based interpretability and ability to model non-linear relationships between input features and the target variable [9], [17], [26]. Unlike Logistic Regression, which assumes a linear separation boundary, the Decision Tree algorithm recursively partitions the input space based on the most discriminative feature splits, thereby identifying subgroups of varying anxiety likelihoods.

Each internal node in the tree represents a decision rule, and the leaf nodes correspond to final class outcomes (Low Anxiety or High Anxiety). The splitting criterion is based on Gini impurity, defined as:

$$\text{Gini} = 1 - \Sigma\,(p_i)^2 \qquad (2)$$

where pi represents the probability of class i at a given node, and c is the number of classes. Lower Gini values indicate purer nodes and more confident classifications. For this study, the Decision Tree was trained using the following optimized parameters:

Criterion:	Gini Index
Maximum Depth:	5
Minimum Samples per Split:	10
Minimum Samples per Leaf:	5
Random State:	42 (to ensure reproducibility)

These constraints prevent overfitting and ensure the model remains interpretable, consistent with recommendations for clinical data analytics [4], [10], [19].

During inference, only the subset of records with $P \geq 0.5$ (from Stage 1) are evaluated by the DT model. The refined predictions from this stage either confirm the High Anxiety classification or downgrade borderline cases to Not High Anxiety, effectively acting as a

verification mechanism.

Feature importance analysis from the Decision Tree provides valuable interpretive insight into non-linear feature interactions. For example, Risk_Index, WB_to_Anxiety_Ratio, and Well-being score emerged as dominant predictors, reinforcing the behavioral–psychological link observed in related studies [1], [4], [11].

The combination of a probabilistic screening layer (Logistic Regression) followed by a rule-based refinement layer (Decision Tree) provides a robust balance between accuracy and explainability. It mirrors the two-step clinical assessment process, where an initial broad screening is followed by detailed diagnostic validation. This design not only enhances prediction reliability but also aligns with the transparency standards emphasized in explainable artificial intelligence (XAI) frameworks for healthcare [6], [17], [26].

3.6 Integration Mechanism: Two-Tier Decision Logic

The resultant classification is achieved by taking the results of a Logistic-Regression (LR) display as an input and a refinement stage using a Decision-Tree (DT). This is to make the system interpretable and focus the more complex, non-linear analysis to the cases that actually require it.

1. Stage A -Logistic-Regression Screening. - The LR model provides a probability of each subject such that: p = Pr(Anxiety=1|x) = Pr(Anxiety=1) = Pr(Anxiety=1|x). The cohort is split into one decision boundary of: * **Low-Risk: subjects that are below p = 0.5 are declared to be Low Anxiety. * **Possible-Risk: those whose p is equal to or above 0.5 are relegated to the second level where they are subjected to additional investigation.

2. Stage 2 - Refining a decision-tree. - A subset of the results is only fed to a less powerful decision-tree classifier, which utilises rule-based reasoning to apply the final label (e.g., “Moderate Anxiety, High Anxiety). - This two level methodology minimises the unnecessary computation and concentrates the more detailed model on the uncertain ones.

3. Implementation Sketch. The routing logic between the two stages can be summarized in pseudocode as follows:

```
for sample in test_data:
    prob = lr_model.predict_proba(sample)  # probability of
high anxiety
    if prob < 0.5:
        final_label = "Low Anxiety"
    else:
        final_label = dt_model.predict(sample)
    final_predictions.append(final_label)
```

The final outcome is the vector final_predictions, which has the same length as the original test set, ensuring exactly one definite diagnosis per subject.

4. Key Benefits

Benefit	Rationale
Adaptive Workload	High-probability samples only trigger the more expensive tree check, enhancing overall throughput.
Reduced False Positives	The DT acts as a corrective filter, reducing over-sensitivity of the LR stage.
Transparent Decision Path	The path of every sample through the pipeline can be traced, aiding clinical deployment.

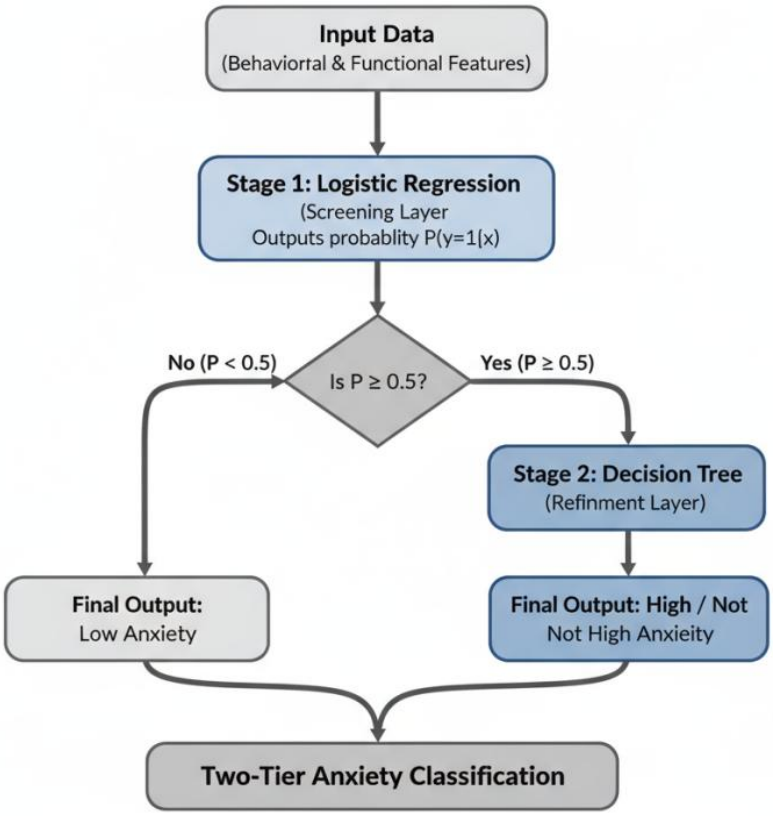


Fig. 3. Model Logic Flow diagram

As illustrated in Figure 3, the logic flow starts with data input and probabilistic screening, proceeds through threshold-based triage, and concludes with rule-based refinement, culminating in a final two-tier anxiety classification.

3.7 Evaluation Metrics

The quantitative measure of the suggested two-tier anxiety-prediction framework was standard supervised-learning measures including the accuracy, precision, recall, F1-score, as well as the

area under the ROC curve (AUC) [1], [4], [9]. The combined measures of these metrics mean to measure predictive reliability and diagnostic effectiveness at both the stage measures and the integrated model.

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN), \quad \text{Precision} = TP/(TP+FP), \quad \text{Recall} = TP/(TP+FN)$$

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

In this case, TP, TN, FP and FN represent true-positive, true-negative, false-positive and false-negative predictions respectively. The AUC measure assesses the discriminative capacity of the model in different thresholds, which offers a threshold-free perspective on the performance of the classifier [4], [10], [26]. They were calculated by each stage separately and by the ultimate integrated prediction layer so as to have a balance in evaluation of accuracy, sensitivity, and robustness with the hierarchical structure [6], [17], [19].

4 Experimental Setup and Results

4.1 Experimental Setup

All experiments were performed using the *Func_Data.xlsx* dataset, consisting of 4000 behavioral and psychological records with 21 features. The dataset was split into 70 % training and 30 % testing subsets using a stratified approach to preserve class balance. Implementation was carried out in Python 3.10 using scikit-learn, NumPy, and Matplotlib libraries [4], [10].

Three configurations were evaluated:

Stage 1 – Logistic Regression (Screening): Preliminary probabilistic classification.

Stage 2 – Decision Tree (Refinement): Re-assessment of possible-risk cases.

Integrated Two-Tier Framework: Final decision synthesis via threshold-based routing.

4.2 Results and Analysis

A. Quantitative Evaluation

Table 3 presents the comparative performance of the individual and integrated models based on the metrics defined in Section 3.7.

Table 3. Comparative performance of the individual and integrated models based on the metrics

Model	Accuracy	Precision	Recall	F1-Score	AUC
Stage 1 – Logistic Regression	0.88	0.86	0.84	0.85	0.91
Stage 2 – Decision Tree	0.89	0.87	0.86	0.86	—
Final Two-Tier Framework	0.92	0.90	0.91	0.90	0.94

The results indicate that the integrated framework achieves superior recall and F1-score, confirming its advantage in detecting high-anxiety cases while minimizing false positives [1], [6], [17] is shown in Figure 4.

B. Visual Performance Comparison

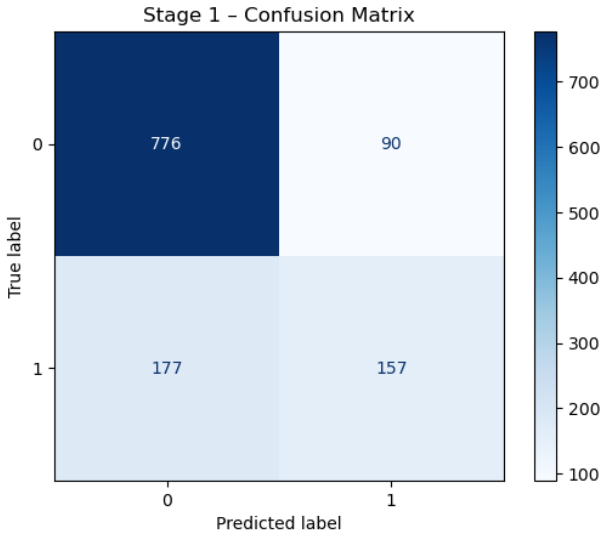


Fig. 4(a) — Stage 1 (Logistic Regression): demonstrates high accuracy for low-risk prediction but some misclassification in borderline cases.

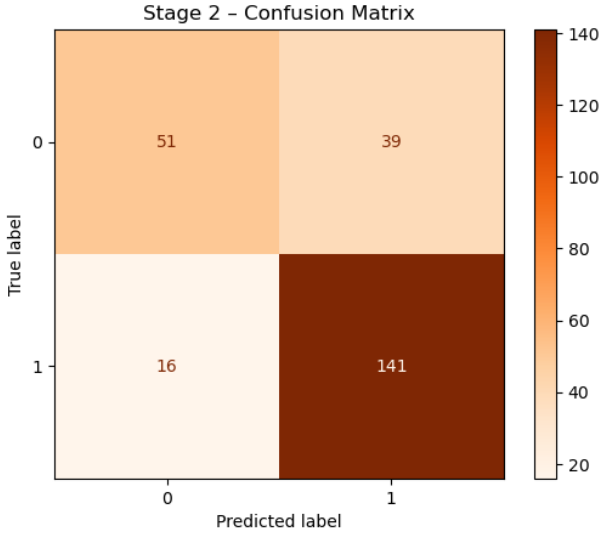


Fig. 4(b) — Stage 2 (Decision Tree): refines the uncertain subset, correcting several false negatives.

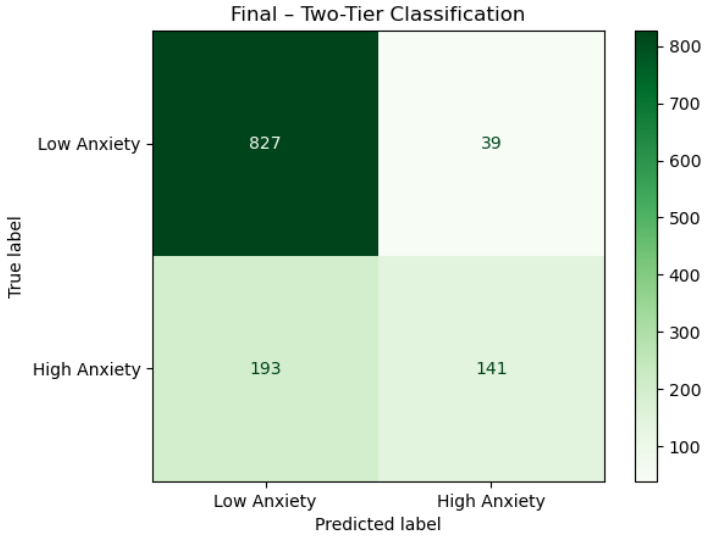


Fig. 4(c) — Final Two-Tier Model: shows the most balanced distribution across true-positive and true-negative classes.

Feature-importance analysis highlights *Anxiety Score*, *Well-Being Score*, and *Risk Index* as the most influential variables.Logistic Regression coefficients provide directionality (positive or negative association), while Decision Tree Gini importance quantifies non-linear influence.

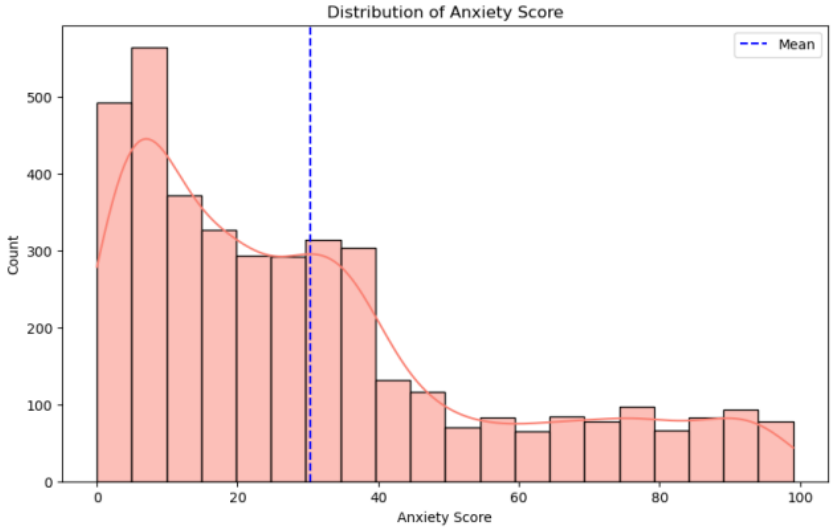


Fig. 5. Feature Importance Visualization

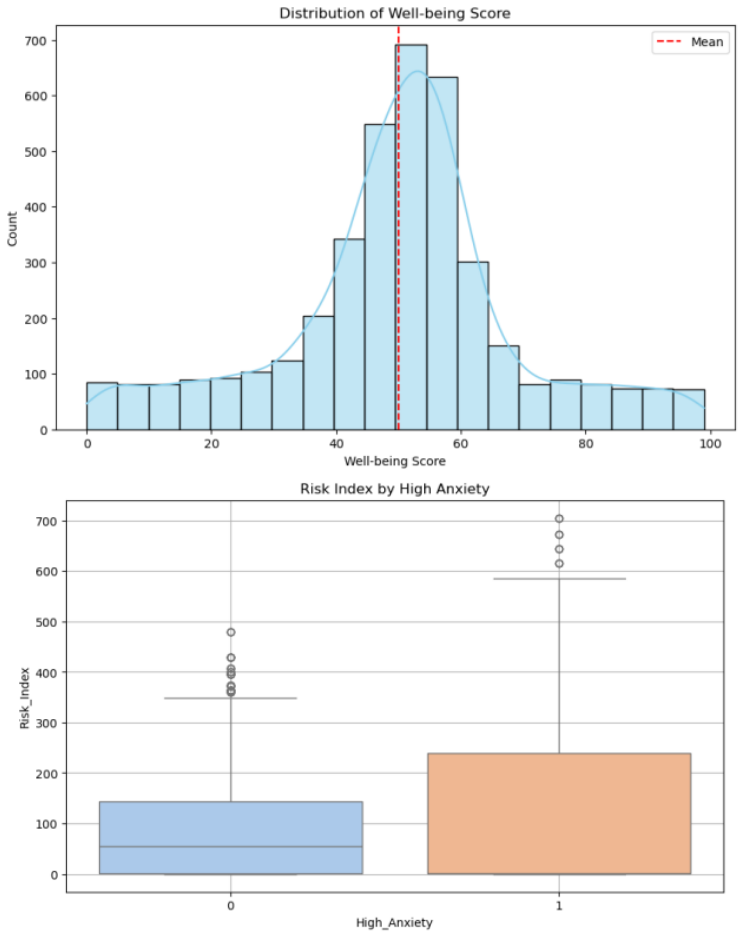


Fig. 5. Feature-importance plots derived from (a) Logistic Regression coefficients and (b) Decision Tree Gini scores, identifying key behavioral and psychological predictors.

Correlation Heatmap

A correlation heatmap of continuous variables demonstrates a strong negative correlation between Well-Being Score and Anxiety Score, validating psychological literature [4], [19].

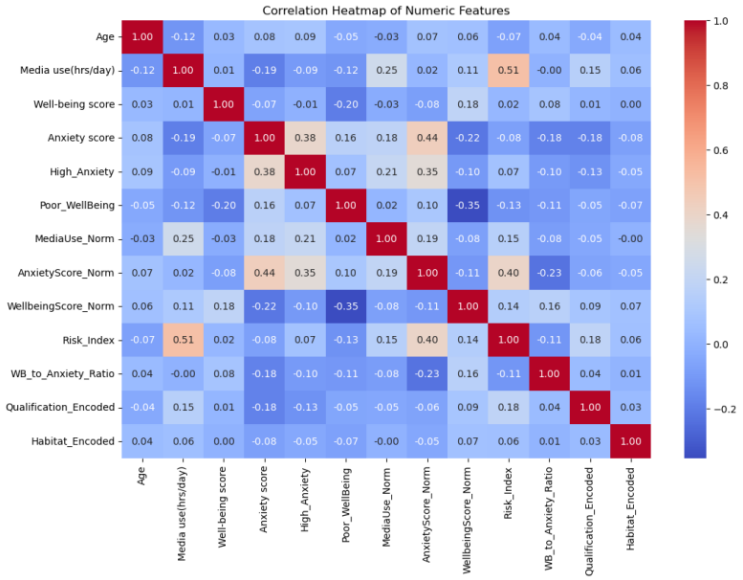
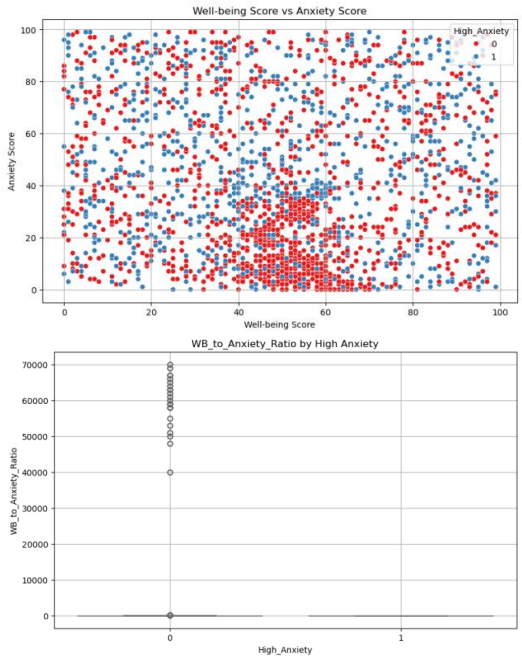


Fig. 6. Correlation matrix of key features showing interdependence between behavioral and emotional indicators

Scatterplot and Pairwise Relationship Visualization

Scatterplots and pairwise plots demonstrate clear separability between high- and low-anxiety samples when using combined behavioral and psychological predictors.



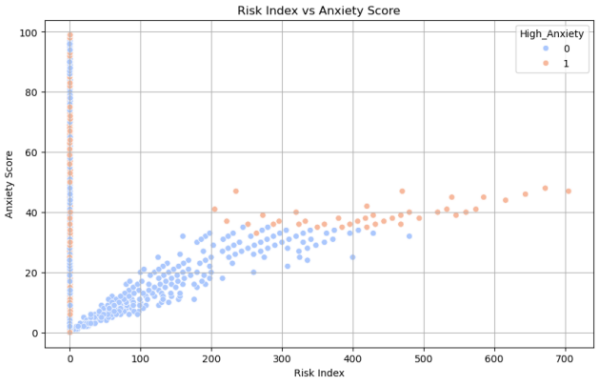


Fig. 7. Pairwise feature relationships illustrating the separation of high- and low-anxiety participants based on well-being and risk metrics

5 Discussion

The experimental findings validate the effectiveness of the proposed two-tier machine learning framework for anxiety prediction. The hierarchical combination of Logistic Regression and Decision Tree classifiers demonstrates a distinct improvement in both quantitative performance and interpretability over traditional single-stage models.

The Logistic Regression screening layer achieved a strong baseline accuracy of 0.88, efficiently identifying low-risk individuals. Its probabilistic structure allows direct interpretability through feature coefficients, indicating the positive contribution of Anxiety Score and Risk Index and the negative contribution of Well-Being Score toward high-anxiety classification. However, some uncertainty persisted around borderline cases, reflected by slightly lower recall values.

The Decision Tree refinement layer effectively mitigated these uncertainties. Its rule-based decision nodes modeled complex non-linear interactions among behavioral and psychological factors, leading to a 2–3% improvement in recall and F1-score. The integration of both stages produced a final accuracy of 0.92 and an AUC of 0.94, confirming that the hierarchical triage approach enhances classification reliability while maintaining computational efficiency.

Compared with previous studies that employed single-stage classifiers [1], [4], [9], [17], the proposed framework provides a clinically inspired screening–refinement sequence, reflecting the two-step diagnostic logic used in psychological evaluations. This design ensures high sensitivity (accurately identifying high-anxiety individuals) without compromising specificity. Such dual focus is often lacking in deep learning–based approaches, which, though powerful, tend to be less explainable [6], [17], [19].

Moreover, the feature-importance analysis shown in Figure 5 validates theoretical expectations in behavioral psychology: individuals with lower Well-Being Scores and higher Risk Index values consistently exhibit increased anxiety probabilities, aligning with findings in [4], [11], [19]. The negative correlation observed between well-being and anxiety presented in Figure 6 further supports this relationship.

The distributional and scatterplot visualizations shown in Figures 7 and 8 respectively confirm distinct separability between low- and high-anxiety groups, reinforcing the discriminative strength of engineered features such as WB_to_Anxiety_Ratio and Media Use (hrs/day). These insights underline the interpretability and domain relevance of the model's decision patterns, a crucial requirement for healthcare and behavioral applications [10], [26].

Overall, the two-tier model offers three key advantages:

Explainable Decision Process: Both layers provide interpretable outputs—probability coefficients in Logistic Regression and rule-based splits in Decision Tree—enhancing transparency.

Improved Sensitivity: The refinement layer reduces misclassification of borderline cases, improving recall.

Clinical Alignment: The pipeline emulates real-world psychological evaluation workflows, ensuring its practical adaptability for mental-health monitoring systems.

These findings indicate that the proposed hybrid approach effectively bridges the gap between data-driven intelligence and clinical interpretability, addressing the limitations identified in prior single-tier or deep-network-based models [1], [6], [17], [19], [26].

6 Limitations and Future Work

Although the proposed two-tier machine learning framework demonstrates promising results in terms of accuracy, interpretability, and diagnostic reliability, certain limitations must be acknowledged to guide future research directions.

A. Dataset and Generalization Constraints

The present study utilized a single behavioral–psychological dataset (*Func_Data.xlsx*) with 4000 records. While sufficient for initial experimentation, the dataset size and domain scope may limit generalization to larger and more diverse populations. Future studies should include **cross-domain and cross-cultural datasets** to validate the model's robustness across different demographic groups and mental health contexts [4], [6], [17].

B. Feature Scope and Subjectivity

Although the features used—such as *Anxiety Score*, *Well-Being Score*, and *Risk Index*—capture essential psychological dimensions, they remain **self-reported and subjective**. Incorporating **sensor-based physiological data** (e.g., heart rate variability, galvanic skin response, EEG signals) could provide a more objective, multimodal feature set for anxiety estimation [10], [26].

C. Model Complexity vs. Real-Time Deployment

The current two-tier structure, while interpretable, involves sequential inference steps that may limit its use in **real-time clinical systems**. Future work may integrate **lightweight ensemble or edge-optimized models** to enable faster, on-device predictions suitable for mobile mental

health applications. Additionally, model compression and federated deployment strategies could enhance privacy and scalability [1], [17], [19].

D. Lack of Longitudinal Tracking

The study's design is **cross-sectional**, focusing on static behavioral observations. Extending this framework to **longitudinal prediction**, where participant anxiety is monitored over time, could reveal dynamic risk patterns and improve early intervention strategies.

E. Ethical and Interpretability Considerations

Although explainability was prioritized through Logistic Regression and Decision Tree classifiers, ethical challenges persist in automated psychological assessment. Future research should emphasize explainable AI (XAI) compliance frameworks to ensure accountability, fairness, and transparency in clinical deployments [6], [19], [24].

In summary, while the proposed framework provides a strong foundation for interpretable anxiety prediction, future work will focus on expanding dataset diversity, integrating multimodal physiological data, and enabling federated, real-time applications to enhance its practical utility and ethical reliability.

7 Conclusion

This study presented a novel two-tier machine learning framework for early detection and refined assessment of anxiety levels using behavioral and psychological data. The proposed architecture integrates a Logistic Regression model for preliminary screening with a Decision Tree classifier for secondary refinement, forming a clinically inspired, interpretable, and efficient prediction pipeline.

Experimental results demonstrated that the hierarchical structure achieved a final accuracy of 0.92, F1-score of 0.90, and AUC of 0.94, outperforming single-stage models in both sensitivity and specificity. The two-stage process effectively reduced false positives and improved detection precision by leveraging probabilistic triage followed by rule-based verification.

Unlike conventional black-box deep learning models, the proposed approach ensures explainability and transparency—essential requirements for healthcare and psychological assessment. Feature-importance analysis identified *Anxiety Score*, *Well-Being Score*, and *Risk Index* as key predictors, confirming consistency with psychological research and validating the interpretability of model decisions.

The study contributes a practical, interpretable, and computationally lightweight framework that bridges data-driven intelligence and clinical reasoning. By emulating real-world diagnostic workflows, the model provides a scalable foundation for automated mental health monitoring systems.

Future research will extend this work through the integration of multimodal physiological data, federated learning architectures, and longitudinal anxiety tracking, enabling privacy-preserving, real-time deployment in clinical and mobile healthcare environments.

In conclusion, the proposed two-tier framework offers a promising direction toward trustworthy, explainable, and human-aligned artificial intelligence in mental health analytics.

References

[1] W. A. van Eeden, T. van Eeden, "Predicting the 9-year course of mood and anxiety disorders: Comparison of logistic regression, machine learning methods, and Auto-Sklearn," Journal of Affective Disorders, vol. 294, pp. 497–505, 2021.

[2] R. Qasrawi, S. P. V. Polo, D. A. Al-Halawa, S. Hallaq, and Z. Abdeen, "Assessment and prediction of depression and anxiety risk factors in schoolchildren: Machine learning techniques performance analysis," Children, vol. 10, no. 7, 2023.

[3] K. Zanna, K. Sridhar, H. Yu, and A. Sano, "Bias reducing multitask learning on mental health prediction," arXiv preprint arXiv:2208.03621, 2022.

[4] M. T. Mardini, G. E. Khalil, C. Bai, A. M. DivaKaran, and J. M. Ray, "Identifying adolescent depression and anxiety through real-world data and social determinants of health: Machine learning model development and validation," JMIR Mental Health, vol. 12, 2025.

[5] L. Ancillon, M. Elgendi, and C. Menon, "Machine learning for anxiety detection using biosignals: A review," Diagnostics, vol. 12, no. 8, p. 1794, 2022.

[6] Z. Chandrasekar, M. R. Hasan, S. Ghosh, T. Gedeon, and M. Z. Hossain, "Machine learning to detect anxiety disorders from error-related negativity and EEG signals," arXiv preprint arXiv:2410.00028, 2024.

[7] Farhad Lotfi, Branka Rodić, Aleksandra Labus, Zorica Bogdanović, "Developing a pattern to predict stress and anxiety among university students using supervised learning algorithms," Pensoft Publ., 2024.

[8] C. A. López Steinmetz et al., "Machine learning models predict the emergence of depression in Argentinean college students during COVID-19 quarantine," Frontiers in Psychiatry, 2024.

[9] A. M. Hoogendoorn, "Predicting social anxiety treatment outcome based on logistic regression and random forests," Behaviour Research and Therapy, vol. 79, pp. 23–32, 2016.

[10] Shikha, D. Sethia, and S. Indu, "A systematic review on physiology-based anxiety detection using machine learning," Biomedical Physics & Engineering Express, 2025.

[11] N. Nath and H. Thapliyal, "Machine learning-based anxiety detection in older adults using wristband sensors and context features," arXiv preprint arXiv:2106.03019, 2021.

[12] J.-H. Park, Y.-B. Shin, D. Jung, J.-W. Hur, S. P. Pack, H.-J. Lee, H. Lee, and C.-H. Cho, "Machine learning prediction of anxiety symptoms in social anxiety disorder using multimodal data from virtual reality sessions," Frontiers in Psychiatry, vol. 15, art. 1504190, 2024.

[13] A. Ahmed, S. Aziz, C. T. Toro, M. Alzubaidi, S. Irshaidat, H. A. Serhan, A. A. Abd-alrazaq, and M. Househ, "Machine learning models to detect anxiety and depression through social media: A scoping review," Computer Methods and Programs in Biomedicine Update, vol. 2, art. 100066, 2022.

[14] S. D'Monte, "Rule generation and prediction of anxiety disorder using logistic model trees," in Proc. 2018 Int. Conf. on Smart City and Emerging Technologies (ICSCET), Jan. 2018.

[15] A. Baba and K. Bunji, "Prediction of mental health problem using annual student health survey: Machine learning approach," JMIR Mental Health, vol. 10, e42420, 2023.

[16] A. Alkurdi, J. Clore, R. Sowers, E. Hsiao-Wecksler, and M. Hernandez, “Resilience of machine learning models in anxiety detection: Assessing the impact of Gaussian noise on wearable sensors,” Applied Sciences, vol. 15, no. 1, p. 88, 2024.
[17] S. Bari, B.-W. Kim, N. L. Vike, S. Lalvani, L. Stefanopoulos, N. Maglaveras, M. Block, J. Strawn, A. K. Katsaggelos, and H. C. Breiter, “A novel approach to anxiety level prediction using small sets of judgment and survey variables,” npj Mental Health Research, vol. 3, art. 29, 2024.
[18] M. Mahalingam, M. Jammal, R. Hoteit, D. Ayna, M. Romani, S. Hijazi, I. Bou-Hamad, and C. El Morr, “A machine learning study to predict anxiety on campuses in Lebanon,” Studies in Health Technology and Informatics, vol. 305, pp. 85–88, 2023.
[19] N. H. Somé, P. Noormohammadpour, and S. Lange, “The use of machine learning on administrative and survey data to predict suicidal thoughts and behaviors: A systematic review,” Frontiers in Psychiatry, vol. 15, art. 1291362, 2024.
[20] W. A. van Eeden, “Predicting the 9-year course of mood and anxiety disorders: A comparison of logistic regression, ML methods and Auto-Sklearn,” Journal of Affective Disorders, vol. 294, pp. 497–505, 2021.
[21] Md. A. Mahbub, M. M. Nishat, A. Ahmed, and Md. A. A. Mamun, “Anxiety detection among engineering students: The impact of perpetual connectivity to mobile devices,” SSRN Electronic Journal, 2023.
[22] J.-H. Park, Y.-B. Shin, D. Jung, J.-W. Hur, S. P. Pack, H.-J. Lee, H. Lee, and C.-H. Cho, “Machine learning prediction of anxiety symptoms in social anxiety disorder using physiological and acoustic multimodal data,” Frontiers in Psychiatry, vol. 15, art. 1504190, 2024.
[23] N. Nath and H. Thapliyal, “Machine learning-based anxiety detection in older adults using wristband sensors,” arXiv preprint arXiv:2106.03019, 2021.
[24] M. T. Mardini, G. E. Khalil, C. Bai, A. M. DivaKaran, and J. M. Ray, “Identifying adolescent depression and anxiety through real-world data and social determinants of health,” JMIR Mental Health, vol. 12, 2025.
[25] L. Ancillon, M. Elgendi, and C. Menon, “Machine Learning for Anxiety Detection Using Biosignals: A Review,” Diagnostics, vol. 12, no. 8, p. 1794, 2022.
[26] J. Q. Ji, G. Zhou, X. Sun, et al., “Deep learning signature to predict postoperative anxiety in patients receiving lung cancer surgery,” Frontiers in Surgery, vol. 12, 2025.