Journal of Information Technology and Computer Science
Volume 11, Number 1, April 2026, pp. 64-76
Journal Homepage: www.jitecs.ub.ac.id

Student Stress Level Prediction Based on DASS-42 Questionnaire Using XGBoost Algorithm: A Case Study of Undergraduate Information Technology Education Students

Moch Zoel Reinanda[1], Satrio Hadi Wijoyo*[2], Satrio Agung Wicaksono[3]
[123]Universitas Brawijaya, Malang
[1]reinanda59@gmail.com, {[2]satriohadi, [3]satrio}@ub.ac.id
*Corresponding Author

Received 08 December 2025 ; accepted xx month 21 May 2026

Abstract. Stress among university students negatively impacts their academic progress and mental health. Recent data showing that 42.9% of critical-status students in the faculty originate from the Information Technology Education program underscores the urgent need for early detection. This research designs and evaluates a machine learning model using the XGBoost algorithm to predict stress among undergraduate students of Information Technology Education at Universitas Brawijaya. Utilizing the CRISP-DM methodology, the study processes a dataset of 234 students derived from the DASS-42 questionnaire and academic records. The workflow included data pre-processing to handle missing values and class imbalance, followed by model training and evaluation. Six scenarios tested prediction targets (five-level multi-class and binary 'Normal' or 'Stress'), data handling, and hyperparameter tuning. Results indicate that the binary classification model was significantly superior. The best model, utilizing original data and default parameters, achieved an accuracy of 97.87%. Evaluation proved its reliability, achieving 100% recall for the 'Stress' class, ensuring no at-risk cases were missed (0 False Negatives). Feature importance analysis identified Mother's Education as the dominant predictor. The research output includes a functional dashboard prototype equipped with LIME interpretation for individual case analysis.

Keywords: machine learning, stress prediction, XGBoost, educational data mining, DASS-42

1 Introduction

Mental health constitutes a fundamental component of individual well-being and is inseparable from physical health. The World Health Organization (WHO) emphasizes that robust mental health enables individuals to realize their potential, work productively, and cope with daily life stresses [1]. In Indonesia, studies using the DASS-42 instrument on undergraduate populations indicate a high prevalence of mental health issues, with stress levels reaching 38.9%, of which 8.3% fall into the severe or extremely severe categories [2]. However, university students represent a vulnerable group experiencing significant psychological pressure during the transition from school to higher education. This phase involves profound academic, social, and lifestyle changes that can trigger psychological distress [3]. When unmanaged, this pressure leads to chronic stress, which research indicates correlates with decreased academic achievement and physical health disorders. Furthermore, unaddressed stress can create a “self-perpetuating cycle” that increases susceptibility to depression later in life [4].

This issue is particularly pronounced in the local context of the Faculty of Computer Science (FILKOM), Universitas Brawijaya. Internal faculty data indicates that 42.9% (9 out of 21 students) of the ‘critical status’ students originate from the Information Technology Education (PTI) program. Data from the counseling unit (UKLTKSP) underscores this urgency with a peak of 11 counseling requests in mid-2025. The researcher conducted preliminary interviews with five PTI students to investigate this phenomenon. The students reported a dual academic burden from the engineering and pedagogy curricula. This qualitative evidence establishes the unique stress context of the PTI students. The study exclusively prioritizes the high-risk PTI population to develop a targeted early warning model. Interviews with academic advisors reveal that while academic load is a factor, family background often acts as a root cause for these critical cases. This situation underscores the urgent need for a tailored early detection system to identify at-risk students before their conditions deteriorate.

Educational Data Mining (EDM) offers a promising solution by leveraging data to predict educational outcomes. Machine Learning (ML) algorithms have shown significant potential in predicting mental health issues in recent years. Previous studies have applied algorithms such as Support Vector Machine (SVM), Random Forest (RF), and Naive Bayes (NB) for stress classification with varying degrees of success [5], [6]. In accordance with the established scope of this research, the evaluation is strictly focused on the design and performance analysis of the eXtreme Gradient Boosting (XGBoost) algorithm, without conducting direct comparisons with other classification algorithms. This algorithm effectively manages missing values and models complex non-linear relationships, making it highly suitable for educational datasets [7].

Despite the established foundation of ML in stress prediction, several research gaps remain in the current literature. First, most studies focus on general student populations, overlooking the specific stressors of dual-competency programs like Information Technology Education. Comparative analysis with previous research reveals that while existing models focus heavily on psychological scores, they often neglect the integration of demographic backgrounds, such as parental education, which this study identifies as a significant predictor. Second, while many studies strive to maximize accuracy in multi-class classification, few have critically evaluated the practical trade-off between complex multi-level prediction versus robust binary screening for real-world early warning systems [8]. A binary approach may offer higher reliability, specifically regarding the Recall metric, ensuring no at-risk student is missed during initial screening.

The interpretability of such models remains a significant challenge for implementation in academic counseling. Academic advisors need to understand why a student is predicted as stressed to provide specific, actionable counseling. Techniques like Local Interpretable Model-agnostic Explanations (LIME) are rarely applied in this specific domain to bridge the gap between “black box” model predictions and human understanding [9]. The lack of interpretability often hinders the trust and adoption of ML models by non-technical educational staff.

This study aims to design a comprehensive stress prediction model using the XGBoost algorithm by integrating DASS-42 psychological data with academic records. The researcher rigorously compares multi-class and binary classification scenarios to determine the most effective approach for screening purposes. Furthermore, this research goes beyond prediction by employing Feature Importance and LIME to identify significant predictors at both population and individual levels. The ultimate goal is to provide a robust, interpretable basis for an early warning system that supports academic advisors in making data-driven interventions.

2 Method

This study employs a quantitative approach adopting the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. This framework ensures a structured workflow consisting of six iterative phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment [10]. The iterative nature of CRISP-DM allows for continuous refinement of the model based on evaluation results at each stage. By following this standard, the researcher ensures that the resulting model is not only statistically robust but also aligned with the practical needs of the academic environment. Figure 1 illustrates the complete stages of the research workflow adopted in this study.

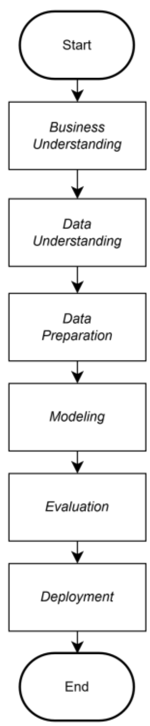

Figure 1. Research Framework

As depicted in Figure 1, the workflow begins with the Business Understanding phase to define the specific research objectives related to the “double burden” stress faced by PTI students. Subsequently, the Data Understanding and Preparation phases involve collecting DASS-42 responses, integrating them with academic records, and applying SMOTE to address class imbalance. The process moves to the Modeling phase, where the researcher trains the XGBoost algorithm under various scenarios to identify the optimal configuration. The Evaluation phase then validates the model using metrics such as Recall and interprets it using LIME. Finally, the study culminates in Deployment, where the researcher translates the best model into a functional dashboard prototype for academic advisors.

2.1 Literature Review

Previous research has demonstrated the effectiveness of applying machine learning algorithms to analyze student mental health data using the DASS instrument. Sulak and Koklu [5] tested various data mining methods on DASS-42 data and found that the Artificial Neural Network (ANN) algorithm achieved a high accuracy of 97.35% in predicting stress. Similarly, Rahunathan et al. [3] utilized the DASS-21 instrument, where the Naive Bayes algorithm proved reliable as an early detection system with an accuracy of 90.00%. These studies establish a high-performance benchmark and prove that stress patterns in questionnaire data can be classified effectively by computational models.

Regarding the specific algorithm used in this study, Du [11] applied XGBoost to predict student stress using student life data and successfully achieved an accuracy of 78.6%. The study highlighted the importance of addressing class imbalance using the Synthetic Minority Oversampling Technique (SMOTE), which significantly improved model performance. The relevance of handling imbalanced data is further supported by Priya, Garg, and Tigga [6], who recommended using holistic evaluation metrics such as F1-Score to avoid bias toward the majority class. Furthermore, Srinath et al. [1] demonstrated that systematic hyperparameter tuning could improve the performance of algorithms like Logistic Regression to 97.45% accuracy on large DASS-42 datasets. Based on these reviews, this study integrates XGBoost, SMOTE, and parameter optimization to build a stress prediction model specific to Information Technology Education students.

2.2 Data Acquisition

The researcher constructed the dataset by collecting data from two primary sources involving undergraduate students of the PTI at Universitas Brawijaya. Primary data was gathered via an online survey using the Depression Anxiety Stress Scales-42 (DASS-42) questionnaire, focusing specifically on the Stress subscale [12]. This subscale consists of 14 items measuring symptoms of tension and agitation, such as difficulty relaxing and over-reaction to situations. The instrument has been validated for its psychometric properties in the Indonesian context, ensuring its reliability for the local population [13]. Table 1 details the scoring categories for the stress subscale as defined in previous studies to provide clarity on the measurement standards used in this study [5].

Table 1. DASS-42 Assessment Categories

Category	Depression	Anxiety	Stress
Normal	0-9	0-7	0-14
Mild	10-13	8-9	15-18
Moderate	14-20	10-14	19-25
Severe	21-27	15-19	26-33
Very severe	>28	>20	>34

Secondary data was obtained directly from the faculty's academic database to complement the psychological profile with academic performance indicators. This dataset encompasses Semester Grade Point Average (IPS) for semesters 1-5, Cumulative Grade Point Average (IPK), and Total Credits (SKS). Additionally, the researcher included parental demographic data, specifically the father's and mother's occupation and education level, to analyze socio-economic influences. The final integrated dataset consists of 234 valid respondents from the cohorts of 2020, 2021, and 2022.

The researchers collected the primary data through a digital questionnaire form. The digital form included a data confidentiality notice for the participants. All students provided explicit informed consent via a mandatory checkbox prior to the data submission. The researchers obtained an official permission letter from the faculty administration to acquire the student academic records. The academic department released the academic data based on this official authorization. The researchers removed personal identifiers such as names and student identification numbers during the data preprocessing phase to guarantee participant anonymity. These comprehensive procedures fulfill the institutional ethical requirements for sensitive psychological data management.

2.3 Data Preparation

Comprehensive data preparation was performed to ensure the quality of the dataset before modeling, starting with data cleaning and handling missing values. Although the survey data was complete, some academic records contained missing values which were handled using median imputation to maintain data distribution integrity. Following this, the researcher transformed categorical features into numerical values using label encoding to ensure compatibility with the XGBoost algorithm. Specifically, parental education levels were mapped to a simplified ordinal scale to reduce cardinality while preserving the hierarchy of educational background.

The researcher engineered the target variable by calculating the stress score through the summation of the 14 stress-related items from DASS-42. Based on these scores, the study created two distinct target scenarios: a Multi-class Target classifying students into five levels, and a Binary Target classifying students into 'Normal' and 'Stress'. In the binary scenario, the 'Stress' class aggregates all categories from Mild to Extremely Severe, an approach designed to address class imbalance and simplify the output for practical screening [8]. Finally, the dataset was split into training and testing sets with an 80:20 ratio, and SMOTE was applied to the training data to address class imbalance [14]. This specific ratio is based on empirical experiments testing multiple proportions (from 50:50 up to 90:10). The 80:20 split yielded the highest accuracy, indicating that the model required a maximum portion of training data to effectively recognize complex patterns within this limited dataset. To ensure the robustness and reproducibility of the experiments, a fixed random seed (random_state=42) was implemented during the data splitting, SMOTE application, and model initialization processes. The SMOTE configuration utilized a default oversampling ratio to match the minority class population with the majority class exactly. The binary classification scenario calculated the nearest neighbors (k) dynamically with a maximum limit of 5. The multi-class scenario applied a similar dynamic calculation with a maximum limit of 3 based on the minimum class samples.

2.4 Modeling Strategy

The researcher selected XGBoost for the modeling phase due to its proven effectiveness in handling structured data. XGBoost optimizes a specific objective function, which combines a loss function to measure predictive error and a regularization term to control model complexity and prevent overfitting [15]. This algorithm constructs an ensemble of decision trees sequentially. The mathematical formulation of the objective function used in this study is presented in Equation (1).

$$L^{(M)}\big(F(x_i)\big) = \sum_{i=1}^{n} L\big(y_i, F(x_i)\big) + \sum_{m=1}^{M} \Omega(h_m) \quad (1)$$

In this equation, L denotes the differentiable convex loss function that measures the difference between the target y_i and the prediction $F(x_i)$, and Ω represents the regularization term for the m -th tree h_m to penalize complexity. The researcher conducted hyperparameter tuning using GridSearchCV across six experimental scenarios to ensure optimal performance. During this tuning process, a 3-fold cross-validation technique was integrated. This repetitive cross-validation approach provides a robust estimation of the model's generalization capability and significantly reduces bias in estimating model performance during the training process [16], [17]. The systematic search space included specific parameters such as n_estimators (100, 200), max_depth (3, 5, 7), learning_rate (0.1, 0.05), and subsample ratios (0.8, 1.0). This tuning process validated the model's stability and prevented overfitting, particularly when the model utilized synthetic data generated by SMOTE.

2.5 Evaluation Metrics

The researcher evaluated the models using the Confusion Matrix and standard classification metrics including Accuracy, Precision, Recall, and F1-Score [18]. The study prioritized Recall as the most critical metric in the context of an early warning system. High recall ensures that the model captures the maximum number of at-risk students, thereby minimizing the occurrence of False Negatives [19]. The metrics are calculated using the following equations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

TP denotes True Positives, TN denotes True Negatives, FP denotes False Positives, and FN denotes False Negatives. The confusion matrix offers a unified framework for understanding these classification errors [20]. Furthermore, the researcher employed Global Feature Importance and LIME to interpret the model's predictions. Global importance identifies which academic or demographic features contribute most significantly to the overall model structure [21]. LIME provides local interpretability for individual predictions, allowing academic advisors to understand the specific risk factors for a particular student [22].

3 Results and Discussion

3.1 Data Distribution Analysis

The researcher performed an exploratory analysis on the DASS-42 scores obtained from the survey to understand the baseline mental health status of the students. This analysis revealed a concerning prevalence of stress within the student population, particularly when grouped into binary categories. Figure 2 illustrates the visual distribution of the binary target class derived from the responses.

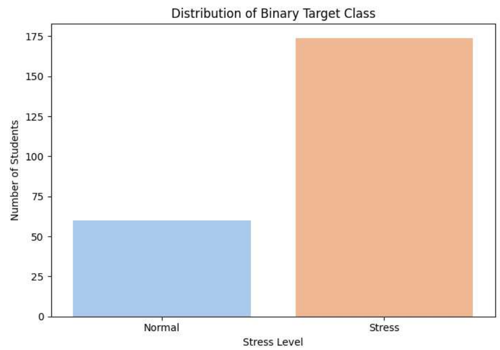

Figure 2. Distribution of Binary Target Class

The data distribution in Figure 2 indicates a significant class imbalance, where 174 students (74.4%) fall into the ‘Stress’ category, while only 60 students belong to the ‘Normal’ category. Furthermore, the detailed multi-class breakdown revealed that the ‘Moderate’ stress level constitutes the most populous class among the students. This skewed distribution strongly influenced the researcher’s decision to investigate SMOTE and binary classification strategies. Addressing this imbalance is crucial to prevent the model from developing a bias toward the majority class during the training phase [23].

3.2 Comparative Performance of Modeling Scenarios

The researcher executed six distinct modeling scenarios to determine the most effective predictive configuration. These scenarios compared variations in target classes, data handling, and parameter settings. Table 2 summarizes the performance metrics for each scenario.

Table 2. Performance Comparison of Six XGBoost Modeling Scenarios

No.	Prediction Target	Data Handling	Model Parameter	Acc. (%)	Prec. (%)	Rec. (%)	F1 (%)
1	Five-Class Stress Level	Without SMOTE	Default	68.09	58.22	51.98	50.14
2	**Two-Class Stress Level**	**Without SMOTE**	**Default**	**97.87**	**98.61**	**95.83**	**97.12**
3	Five-Class Stress Level	With SMOTE	Default	78.72	65.70	62.42	60.64
4	Two-Class Stress Level	With SMOTE	Default	95.74	94.40	94.40	94.40
5	Five-Class Stress Level	Without SMOTE	Tuned - GridSearchCV	65.96	38.33	49.84	43.27
6	Five-Class Stress Level	With SMOTE	Tuned - GridSearchCV	65.96	53.71	51.71	51.08

Table 2 summarizes the classification results across all experimental scenarios. The column headers represent the abbreviated forms of accuracy (Acc.) and the macro average for precision (Prec.), recall (Rec.), and F1-score (F1). The comparative results in Table 2 demonstrate a significant performance gap between binary and multi-class models. Scenario 2 emerged as the superior model, achieving an Accuracy of 97.87% and an F1-Score of 97.12%. This finding suggests that binary tasks yield higher accuracy than complex multi-class problems in this domain because the distinction between 'Normal' and 'Stressed' is computationally clearer than the fuzzy boundaries between specific stress levels [23]. Additionally, the application of SMOTE in the binary scenario resulted in a slight accuracy decrease, indicating that the XGBoost algorithm handles the inherent class imbalance effectively without the need for synthetic oversampling in this specific context [24].

3.3 In-Depth Evaluation of the Best Model

The researcher performed a detailed evaluation on the best model to validate its reliability for academic screening. Table 3 presents the detailed classification report, decomposing the performance by class.

Table 3. Detailed Classification Report of the Best Model

	Precision	Recall	F1-score	Support
Normal	1.00	0.92	0.96	12
Stress	0.97	1.00	0.99	35
Accuracy			0.98	47
Macro avg	0.99	0.96	0.97	47
Weighted avg	0.98	0.98	0.98	47

Figure 3 illustrates the Confusion Matrix of the model's predictions on the test dataset to provide a visual representation of these errors,

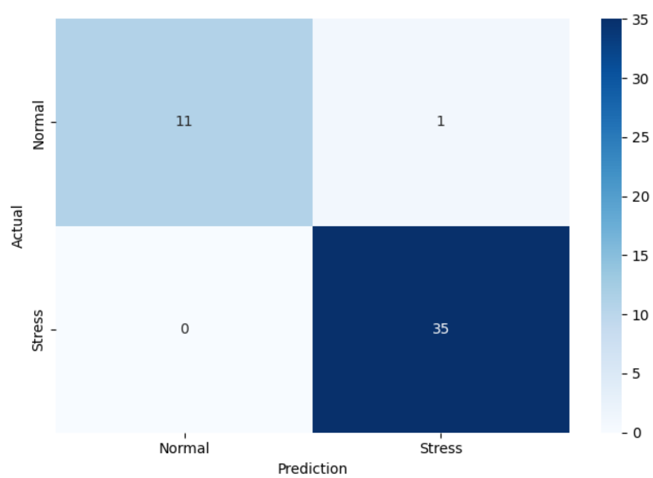

Figure 3. Confusion Matrix of the Best Model

The confusion matrix reveals a critical achievement regarding model sensitivity: the model achieved a recall of 1.00 for the 'stress' class. This outcome resulted in zero false negatives, meaning the system correctly identified all 35 actual stress cases without exception. Ideally, a mental health screening tool must avoid missing stressed students at all costs, as a false negative would imply a stressed student is labeled 'normal' and ignored by the support system. The single false positive observed is acceptable in an early warning context, as it merely results in a precautionary check-up rather than neglect [19].

The initial evaluation utilized an 80:20 holdout split on the testing data. The XGBoost model achieved a peak accuracy of 97.87% in this scenario. The researcher applied a stratified 5-fold cross-validation across the entire dataset $(N = 234)$ to test statistical stability. The researcher selected a value of $k = 5$ following the empirical guidelines of Abedin et al. [25]. Their study demonstrated an increase in fold-to-fold variance as the value of k grew from 3 to 20 [25]. The selection of small folds minimizes estimation variance on a dataset with a limited minority class. Table 4 displays the multi-metric evaluation results of this cross-validation process.

Table 4. Best Model Stability Evaluation via Stratified 5-Fold Cross-Validation

Evaluation Metric	Average Score (%)
Accuracy	91.88
Precision (Macro)	90.66
Recall (Macro)	88.55
F1-Score (Macro)	89.17

The model demonstrates a consistent predictive capability across all folds based on the data in Table 4. This validation process yielded an average accuracy of 91.88%. The low standard deviation value proves the strong generalization capability of the model. The XGBoost framework recorded a macro-average precision of 90.66% and a recall of 88.55%. These metric values mitigate statistical reliability concerns regarding the small holdout sample size. The alignment between the holdout accuracy and the cross-validation average confirms the robustness of the research methodology.

3.4 Analysis of Determinant Factors

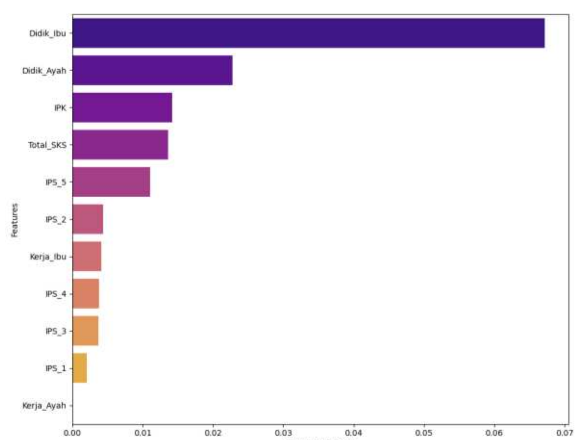

Figure 4. Feature Importance of Academic Variables

This section evaluates the predictive contribution of the academic and demographic features. The researcher excluded the items of the DASS-42 stress subscale from this feature importance analysis. These clinical items constitute the direct scoring formula for the target variable. Their inclusion creates target leakage in the predictive model. A model with these items predicts the stress levels solely based on the questionnaire scores. This study aims to identify early warning predictors from the non-clinical factors. Therefore, the analysis removes the clinical features to evaluate the true predictive power of the external determinants. Figure 4 displays the top predictors identified by the model.

Figure 4 displays the importance scores for each predictive feature. The visualization utilizes specific abbreviations for the student profile variables. The label ‘Didik_Ibu’ denotes the mother’s education level. The label ‘Didik_Ayah’ signifies the father’s education level. The abbreviation ‘IPK’ represents the cumulative grade point average. The term ‘Total_SKS’ states the total accumulated credits. The labels ‘IPS_1’ to ‘IPS_5’ represent the semester grade point averages from the first to the fifth semester. The code ‘Kerja_Ibu’ indicates the mother’s employment status. The code ‘Kerja_Ayah’ defines the father’s employment status.

The analysis revealed that Mother’s Education (Didik_Ibu) emerged as the most significant feature, obtaining an importance score of 0.067176, which is nearly three times higher than the second feature, Father’s Education (Didik_Ayah). This quantitative finding provides empirical support to qualitative insights gathered from faculty interviews, where academic advisors indicated that family background issues often cause student academic distress. Unlike purely academic stressors, parental background often dictates the socio-economic support system available at home. This outcome offers a significant contribution when explicitly connected to previous studies. While earlier research on student stress prediction [5], [6] predominantly emphasized academic workloads and psychological scores as primary features, our model reveals that for dual-competency programs, demographic factors and specifically Mother’s Education play a more critical role. Unlike models that exclude parental background, our findings imply that a student’s familial support system heavily dictates their resilience against academic pressure. The XGBoost algorithm computes feature importance scores based on empirical predictive contribution. This machine learning metric identifies key risk indicators within the specific dataset instead of traditional causal relationships. This mathematical approach highlights the strong predictive correlation of the mother’s education feature across the student profiles. Consequently, the practical implication of this study is that institutional interventions must be holistic. Universities should integrate psychosocial family counseling into their early warning systems, moving beyond mere academic tutoring [16].

3.5 Local Interpretability with LIME

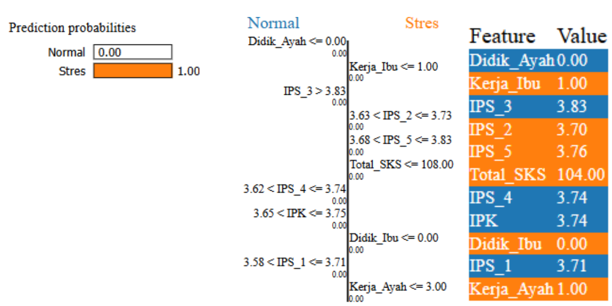

Figure 5. LIME Visualization for Student Index 0

This study utilizes the LIME method to provide individualized prediction explanations for the academic stakeholders. The local explanation mechanism identifies the specific stress factors for each student. The researcher present a single student instance (Index 0) as an illustrative example of this capability. Figure 5 displays the actual visualization and quantitative output for this specific instance.

The analysis reveals the academic performance in Semester 3 and 4 as protective factors for this student. The mother's employment (Kerja_Ibu) status overpowers these positive academic signs with a strong risk weight. This interpretability level demonstrates the capability of the model in identifying hidden high-risk students. The machine learning algorithm detects the emotional struggles from external pressures despite the good academic grades.

3.6 Implementation and Testing

The researcher implemented the best model into a web-based dashboard prototype to facilitate practical use by academic advisors. The dashboard processes the input and displays the prediction results alongside an interpretation of driving factors. The dashboard visualizes the "Dominant Stress Symptoms" and "Individual Driving Factors", allowing advisors to understand the specific context of the prediction. Figure 6 displays the actual visual prototype of this early warning dashboard.

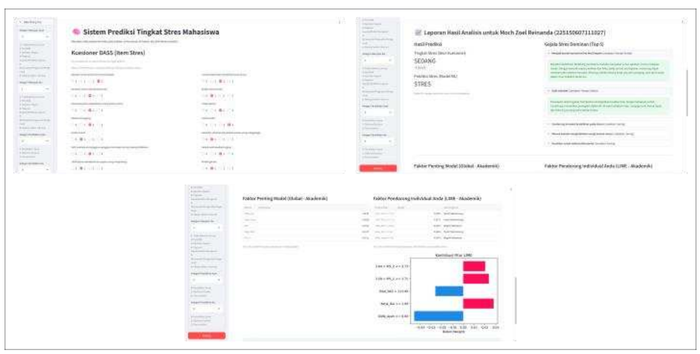

Figure 6. Web-based Dashboard Interface

Table 5. Dashboard Testing Results on Non-PTI Students

No.	Program	Score	Manual Result	Model Prediction	Result
1	Information Technology	21	Moderate	Stress	Valid
2	Information Technology	17	Mild	Stress	Valid
3	Computer Engineering	21	Moderate	Stress	Valid
4	Informatics Engineering	24	Moderate	Stress	Valid
5	Information Technology	19	Moderate	Stress	Valid
6	Information Systems	8	Normal	Normal	Valid

Furthermore, the researcher conducted functionality testing on six students from other study programs on the same faculty to validate the technical robustness of the system. Table 5 presents the comparison between model predictions and manual DASS-42 calculations for these external samples.

The results in Table 5 show 100% consistency between the model predictions and the actual labels derived from manual calculation. The model accurately classified students with high stress scores (Mild to Moderate) into the 'Stress' class and correctly identified students with low scores as 'Normal'. This success proves the model's generalization capability, indicating that the stress patterns learned from PTI students are relevant and applicable to students from other programs within the same faculty context. The researcher does not use this limited external sample to claim statistical cross-program comparisons. This specific evaluation purely confirms the technical robustness of the model against different academic backgrounds.

4 Conclusion

This study successfully designed a stress prediction model for Information Technology Education students using the XGBoost algorithm within the CRISP-DM framework. The researcher evaluated six experimental scenarios to determine the optimal configuration for an early warning system. The results established that the binary classification model using original data and default parameters delivered the superior performance, achieving an Accuracy of 97.87% on the test dataset.

The model demonstrated exceptional reliability for screening purposes. It achieved a perfect Recall of 100% for the 'Stress' class, ensuring that the system detects every at-risk student without exception (0 False Negatives). Furthermore, the analysis revealed that hyperparameter tuning using GridSearchCV did not improve performance compared to the robust default parameters of XGBoost, suggesting that the model captures the data patterns effectively without complex optimization.

Regarding determinant factors, the feature importance analysis highlighted that family background exerts a more significant influence on student stress than academic metrics alone. Specifically, Mother's Education emerged as the most dominant predictor, followed by Father's Education. This quantitative evidence aligns with qualitative insights from academic advisors, confirming that family-related issues often constitute the root cause of academic distress. Consequently, institutional interventions must extend beyond academic tutoring to include holistic psychosocial support.

Future work should focus on expanding the dataset to include students from diverse faculties to validate the model's generalizability across different academic disciplines. Additionally, researchers should extend the prediction scope to encompass Depression and Anxiety scales from the DASS-42 instrument, providing a more comprehensive mental health profile. Finally, the integration of this predictive model into a real-time academic information system would significantly enhance the capability of academic advisors to perform timely and data-driven interventions.

References

1. Srinath, K.S., Kiran, K., Pranavi, S., Amrutha, M., Shenoy, P.D., Venugopal, K.R.: Prediction of depression, anxiety and stress levels using DASS 42. In: *2022 IEEE 7th International Conference for Convergence in Technology (I2CT)*, pp. 1–6. Pune, India 2022.
2. E. Astutik, I. H. Perdana, D. S. Mariyati, and A. P. Ulfah, "Depression, anxiety, and stress among students in newly established remote university campus in Indonesia," *Malaysian Journal of Medicine and Health Sciences*, vol. 16, no. 1, pp. 270–277, 2020.
3. L. Rahunathan, G. Arun, P. Ragu, L. S. Satish, and V. N. Srisakthi, "Machine Learning Model for Student Stress Level Prediction," in *2025 4th OPJU International Technology Conference (OTCON)*, pp. 1-7, 2025.
4. C. Hammen, "Stress and depression," *Annu. Rev. Clin. Psychol*., vol. 1, no. 1, pp. 293-319, 2005.
5. S. A. Sulak and N. Koklu, "Analysis of Depression, Anxiety, Stress Scale (DASS-42) with

methods of data mining," *European Journal of Education*, vol. 59, no. 4, p. e12778, 2024.
6. A. Priya, S. Garg, and N. P. Tigga, "Predicting anxiety, depression and stress in modern life using machine learning algorithms," *Procedia Computer Science*, vol. 167, pp. 1258-1267, 2020.
7. I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, pp. 99129-99149, 2022.
8. L. Melaku, G. Bulcha, and D. Worku, "The prevalence and severity of depression, anxiety, and stress among medical undergraduate students of Arsi University and their association with substance use, Southeast Ethiopia," *Education Research International*, vol. 2021, no. 1, p. 9936157, 2021.
9. K. T. Sar and B. Kaya, "Diagnosis Heart Attack and Stroke Risk using Machine Learning Algorithms... with LIME," in *2025 15th International Conference on Advanced Computer Information Technologies (ACIT)*, pp. 824-827, 2025.
10. P. Chapman et al., CRISP-DM 1.0: Step-by-step data mining guide, SPSS Inc., 2000.
11. W. Du, "Application of improved SMOTE and XGBoost algorithm in the analysis of psychological stress test for college students," *Journal of Electrical and Computer Engineering*, vol. 2022, no. 1, p. 2760986, 2022.
12. P. F. Lovibond and S. H. Lovibond, "The structure of negative emotional states: Comparison of the DASS with the Beck Depression and Anxiety Inventories," *Behaviour Research and Therapy*, vol. 33, no. 3, pp. 335-343, 1995.
13. D. Muttaqin and S. Ripa, "Psychometric properties of the Indonesian version of the Depression Anxiety Stress Scale," *Psikohumaniora: Jurnal Penelitian Psikologi*, vol. 6, no. 1, pp. 61-76, 2021.
14. M. S. N. M. Danuri, R. Abd Rahman, I. Mohamed, and A. Amin, "The improvement of stress level detection in twitter: imbalance classification using SMOTE," in *2022 IEEE International Conference on Computing (ICOCO)*, pp. 294-298, 2022.
15. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794, 2016.
16. S. S. Shinde and A. S. Ghotkar, "From Questionnaires to Actionable Insights: Machine Learning for Mental Stress Detection," *Engineering, Technology & Applied Science Research*, vol. 15, no. 6, pp. 29240-29250, 2025.
17. A. Bozorgmehr and B. Weltermann, "Prediction of chronic stress and protective factors in adults: Development of an interpretable prediction model based on XGBoost and SHAP using national cross-sectional DEGS1 data," *JMIR AI*, vol. 2, p. e41868, 2023.
18. O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Scientific Reports*, vol. 14, no. 1, p. 6086, 2024.
19. D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Computers & Operations Research*, vol. 152, p. 106131, 2023.
20. J. Erbani, P. É. Portier, E. Egyed-Zsigmond, and D. Nurbakova, "Confusion matrices: A unified theory," *IEEE Access*, 2024.
21. Y. Takefuji, "Beyond XGBoost and SHAP: Unveiling true feature importance," *Journal of Hazardous Materials*, vol. 488, p. 137382, 2025.
22. A. Saha et al., "Towards Interpretable Diabetes Prediction: A Comparative Study of Machine Learning Models with Explainable AI using LIME and SHAP," in *2024 27th International Conference on Computer and Information Technology (ICCIT)*, 2024, pp. 1564-1569.
23. F. Rodrigues and H. Correia, "Semi-supervised and ensemble learning to predict work-related stress," *Journal of Intelligent Information Systems*, vol. 62, no. 1, pp. 77-90, 2024.
24. G. Vos et al., "Ensemble machine learning model trained on a new synthesized dataset generalizes well for stress prediction," *Journal of Biomedical Informatics*, vol. 148, p. 104556, 2023.
25. T. Abedin, H. Xu, and S. Uddin, "The impact of K selection in K-fold cross-validation on bias and variance in supervised learning models," *Scientific Reports*, 2026.