Nusantara Journal of Multidisciplinary Informatics

Volume 1 Issue 1
https://doi.org/10.56705/njomi.v1i1.3

Research Article

Plausible-Value-Sensitive Academic Resilience Among Socioeconomically Disadvantaged Indonesian Students: Evidence from PISA 2022

Riska[1]
[1] Politeknik Negeri Ujung Pandang, Jl. Politeknik, Tamalanrea Indah, Kec. Tamalanrea, Kota Makassar, Sulawesi Selatan 90245, riskaanasirr@gmail.com

Correspondence should be addressed to Riska; riskaanasirr@gmail.com

Received 26 February 2026 Accepted 12 April 2026; Published 30 April 2026

Abstract:
Introduction: Academic resilience refers to the ability of socioeconomically disadvantaged students to achieve strong academic outcomes despite structural constraints. However, resilience classifications derived from large-scale assessments may vary because student achievement is represented by multiple plausible values. Methods: This study analyzed PISA 2022 data from 13,439 Indonesian students across 410 schools. Socioeconomic disadvantage was defined using the weighted lowest quartile of the ESCS distribution, while academic resilience was identified separately for each of the ten mathematics plausible values using the weighted highest-quartile achievement threshold. Classification stability was examined across plausible values. Logistic Regression, Random Forest, and XGBoost were evaluated using five-fold school-grouped cross-validation, followed by predictor-domain ablation, SHAP-based importance analysis, and sensitivity testing. Results: An average of 15.20% of disadvantaged students were classified as academically resilient. However, 30.61% showed unstable classifications across plausible values, indicating substantial membership uncertainty. XGBoost achieved the best predictive performance with a weighted AUROC of 0.718 and AUPRC of 0.324. Psychological factors produced the largest incremental predictive gain beyond demographic and socioeconomic characteristics. Growth mindset, disciplinary climate, confidence in 21st-century mathematics skills, family support, and curiosity showed the most consistent predictive importance across plausible-value models. Conclusion: Academic resilience in Indonesia is measurable but not fully invariant across plausible values. Incorporating plausible-value sensitivity provides a more robust framework for identifying resilience and evaluating its associated protective factors.

Keywords: Academic Resilience; PISA 2022; Plausible Values; Socioeconomic Disadvantage; XGBoost; SHAP; Indonesia.

1. Introduction

Socioeconomic disadvantage is consistently associated with lower academic achievement, but this relationship is not deterministic. Some students from disadvantaged backgrounds are able to achieve strong academic outcomes despite limited socioeconomic resources. This phenomenon, commonly referred to as academic resilience, has become important in educational research because it shifts attention from disadvantage alone toward the individual and contextual factors that may support successful adaptation [1]–[3]. Previous studies have linked academic resilience with psychological characteristics, family support, school climate, engagement, and other potentially modifiable factors [1], [3], [4].

The Programme for International Student Assessment (PISA) provides an important framework for examining academic resilience across national education systems. In PISA 2022, socioeconomically disadvantaged students can be identified from the lowest national quarter of the index of economic, social and cultural status (ESCS), while resilient students are those who nevertheless perform in the highest quarter of national academic achievement [5]. In Indonesia, approximately 15% of disadvantaged students reached the highest national quarter of mathematics performance in PISA 2022 [5]. This indicates that a meaningful proportion of Indonesian students achieve relatively strong mathematics outcomes despite socioeconomic constraints.

Existing studies have examined factors related to academic resilience using conventional statistical approaches and, more recently, machine-learning techniques. Cheung et al. [6], for example, applied Random Forest and SHAP to PISA 2022 data from 79 countries and economies to identify characteristics associated with mathematics resilience. Indonesian PISA studies have also investigated relationships among socioeconomic conditions, growth mindset, school climate, teacher support, and mathematics achievement [7]–[9]. However, these studies generally treat student achievement or resilience classification as a fixed outcome. The stability of resilience membership itself has received substantially less attention.

Recent educational data-mining studies have further demonstrated the potential of supervised machine learning for analyzing student outcomes. Random Forest, XGBoost, and Logistic Regression have been applied to predict student pass/fail outcomes from learning-management-system activity, while Random Forest and neural networks have been evaluated for student grade classification [10], [11]. Support Vector Machine has also been used to classify study duration and graduation outcomes from academic and demographic characteristics [12]. Beyond academic performance, explainable machine-learning approaches have increasingly been applied to student and adolescent behavioral data to identify influential psychosocial and contextual predictors [13], [14]. These developments demonstrate the growing value of predictive modelling in student-related research while also reinforcing the importance of model validation, interpretability, and cautious interpretation of predictive associations.

This issue is important because PISA does not provide a single observed mathematics score for each student. Instead, proficiency is represented through multiple plausible values that are designed for population-level inference [15]. Using only one plausible value to construct a binary resilience outcome may therefore result in different students being classified as resilient depending on the plausible-value realization used. Previous research has already shown that academic-resilience classifications can vary considerably when definitions and achievement thresholds change [4], [16]. However, limited evidence exists on whether resilience classifications and their associated predictive factors remain stable across all mathematics plausible values, particularly for Indonesian students.

To address this gap, the present study develops a plausible-value-sensitive and school-aware analysis of academic resilience among Indonesian students using PISA 2022 data. Resilience membership is reconstructed separately across all ten mathematics plausible values to quantify classification stability. Logistic Regression, Random Forest, and XGBoost are then evaluated using school-grouped cross-validation, while SHAP-based analysis is used to examine whether predictor importance remains consistent across plausible-value-specific models. The study focuses on robust predictive associations rather than causal effects.

Accordingly, this study addresses three research questions: RQ1: How stable is the classification of academic resilience among socioeconomically disadvantaged Indonesian students across the ten mathematics plausible values in PISA 2022? RQ2: Which individual, family, and learning-environment factors consistently distinguish academically resilient from non-resilient disadvantaged students across plausible-value realizations? RQ3: How much incremental predictive information do psychological, family, and school-behavioral factors provide beyond demographic and socioeconomic characteristics? Through these questions, the study extends existing PISA-based resilience research by examining not only who is classified as resilient, but also how stable that classification and its predictive factors remain under achievement-measurement uncertainty.

2. Method

Data and Study Sample

This study used the PISA 2022 student-level public-use dataset provided by the OECD [17]. The analysis was restricted to Indonesia and included 13,439 students from 410 schools. The final student sampling weight (W_FSTUWT) was used for weighted percentile estimation, prevalence calculation, model training, and evaluation.

Students with valid ESCS data were used to determine the national socioeconomic threshold. Socioeconomic disadvantage was defined as membership in the weighted lowest 25% of the Indonesian ESCS distribution. This procedure identified 2,772 disadvantaged students, who formed the primary analytical sample.

Academic Resilience and Classification Stability

The weighted socioeconomic threshold was defined as:

$$T_{ESCS} = Q_{0.75}^{(w)}(ESCS) \tag{1}$$

where (Q_q^{(w)}) represents the weighted (q)-th quantile. The resulting Indonesian ESCS threshold was −2.3106.

Because PISA 2022 provides ten mathematics plausible values, the high-achievement threshold was calculated separately for each plausible value:

$$T_{MATHj} = Q_{0.75}^{(w)}\left(PV_j^{MATH}\right), \qquad j = 1,2,\dots,10 \tag{2}$$

Academic resilience was then independently determined for each plausible value:

$$R_{ij} = \mathbb{1}\left(ESC\,S_i \le T_{ESCS} \bigwedge PV_{ij}^{MATH} \le T_{MATHj}\right) \tag{3}$$

where (R_{ij}=1) indicates that student (i) satisfied the resilience criterion under plausible value (j).

Classification stability was measured from the number of plausible values under which a student was classified as resilient:

$$C_i = \sum_{j=1}^{10} R_{ij}, \qquad 0 \le C_i \le 10 \tag{4}$$

and standardized as:

$$S_i = \frac{C_i}{10} = \frac{1}{10}\sum_{j=1}^{10} R_{ij}, \qquad 0 \le S_i \le 1 \tag{5}$$

Students with (C_i=0) were considered consistently non-resilient, those with (C_i=10) consistently resilient, and those with (1\leq C_i\leq9) classification-unstable. Pairwise agreement, Cohen's kappa, and positive-class Jaccard similarity were also calculated to assess agreement across plausible values.

Predictor Variables and Preprocessing

Predictors were selected from four domains based on the academic-resilience literature and variable availability in the Indonesian PISA 2022 dataset.

Table 1. Characteristic

Domain	Variables
Demographic and socioeconomic	ST004D01T, AGE, GRADE, IMMIG, REPEAT, ESCS
Psychological	MATHEF21, CURIOAGR, GROSAGR
Family	FAMSUP
School and behavioral	BELONG, BULLIED, FEELSAFE, DISCLIM, TEACHSUP, SKIPPING, TARDYSD

Four initially considered variables, ANXMAT, MATHEFF, PERSEVAGR, and RELATST, were excluded because they contained 100% missing values in the Indonesian extract.

Missing predictor values were handled using median imputation fitted only on the training portion of each cross-validation fold. Continuous variables used by Logistic Regression were additionally standardized using training-fold parameters. A complete-case analysis was conducted as a robustness test.

Machine-Learning Models and Validation

Three binary classifiers were evaluated: Logistic Regression, Random Forest [18], and XGBoost [19]. Logistic Regression served as the linear baseline, while Random Forest and XGBoost were included to capture nonlinear relationships and interactions.

Comparative model evaluation and cross-validation have also been extensively applied in previous Indonesian machine-learning studies. Azis and collaborators evaluated K-Nearest Neighbor using cross-validation on binary health data [20], compared multiple classification algorithms under a common evaluation framework [21], and examined Support Vector Regression for continuous prediction problems [22]. Subsequent studies compared Decision Tree and Random Forest performance in multiclass image classification [23] and evaluated Random Forest using five-fold cross-validation for predictive modelling [24]. Collectively, these studies reinforce the importance of empirically comparing algorithms and validation strategies rather than assuming that a particular learning method will perform consistently across datasets and problem domains.

The principal XGBoost configuration used 220 trees, maximum depth of 3, learning rate of 0.04, row subsampling of 0.85, and feature subsampling of 0.85. Class imbalance was handled using the weighted ratio between negative and positive observations within the training data.

To avoid information leakage between students from the same school, model validation used five-fold GroupKFold cross-validation with CNTSCHID as the grouping variable. Consequently, students from the same school never appeared simultaneously in the training and test partitions.

Each model was evaluated separately for all ten plausible-value-specific resilience outcomes. Performance was assessed using survey-weighted AUROC and AUPRC, with AUPRC included because resilient students represented a minority class [26].

Ablation and SHAP Stability Analysis

An incremental ablation experiment was used to determine how much predictive information was added by different predictor domains. Four nested feature sets were compared:

- demographic and socioeconomic variables;
- demographic, socioeconomic, and psychological variables;
- the previous variables plus family support; and
- the full model including school and behavioral characteristics.

The strongest-performing model was subsequently interpreted using SHAP [28]. Global importance for each predictor was calculated from mean absolute SHAP values:

$$I_{fj} = \frac{1}{n_j}\sum_{i=1}^{n_j}\left|\phi_{ifj}\right| \tag{6}$$

where (I_{fj}) represents the importance of feature (f) under plausible value (j).

Predictor stability was evaluated using median SHAP rank across the ten models and the number of plausible values in which a feature appeared among the five most important predictors:

$$K_f^{(5)} = \sum_{j=1}^{10} \mathbb{I}\left(r_{fj} \leq 5\right),\ 0 \leq K_f^{(5)} \leq 10 \tag{7}$$

Higher values indicated more consistent predictor importance across plausible-value realizations.

Sensitivity Analysis

Two robustness analyses were conducted. First, the resilience definition was reconstructed using alternative percentile combinations of 20/80, 25/75, and 30/70 for socioeconomic disadvantage and mathematics achievement.

Second, the main XGBoost analysis was repeated using only complete cases to assess whether the findings depended on median imputation.

The complete analytical procedure was therefore:

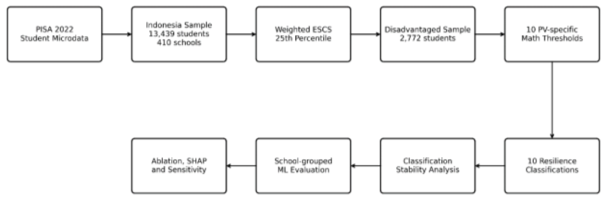

Figure 1. Analytical workflow for the plausible-value-sensitive academic resilience analysis.

3. Result and Discussion

Academic Resilience and Classification Stability

The weighted 25th percentile of ESCS was −2.3106, identifying 2,772 socioeconomically disadvantaged students. Across the ten mathematics plausible values, the weighted proportion classified as academically resilient ranged from 14.71% to 15.84%, with an average of 15.20%. This estimate closely corresponds to the approximately 15% reported for Indonesia in PISA 2022 [4], supporting the validity of the resilience-construction procedure.

However, resilience membership was not fully stable across plausible values. Of the 2,772 disadvantaged students, 1,762 were never classified as resilient, 144 were resilient under all ten plausible values, and 866 changed classification at least once. Survey-weighted analysis showed that approximately 30.61% of disadvantaged students belonged to this classification-unstable group.

Table 2. Distribution of resilience classification across ten plausible values

(C_i)	Students	Classification
0	1,762	Consistently non-resilient
1–9	866	Classification-unstable
10	144	Consistently resilient
Total	**2,772**	

Although pairwise agreement across plausible values reached approximately **88.9%**, Cohen's kappa was about **0.60** and positive-class Jaccard similarity was approximately **0.50**. These results indicate that high overall agreement partly reflects the dominance of non-resilient cases, while the identification of resilient students is less stable. This supports previous evidence that academic-resilience classification can be sensitive to measurement and operational choices [4], [16].

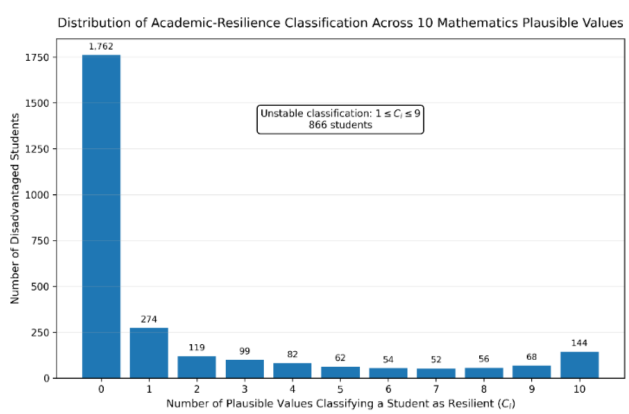

Figure 2. Distribution of disadvantaged Indonesian students according to the number of mathematics plausible values under which they were classified as academically resilient.

Machine-Learning Performance

Logistic Regression, Random Forest, and XGBoost were evaluated using five-fold school-grouped cross-validation across all ten resilience outcomes.

Table 3. Predictive performance across plausible values

Model	Weighted AUROC	Weighted AUPRC
Logistic Regression	0.675	0.287
Random Forest	0.705	0.315
XGBoost	**0.718**	**0.324**

XGBoost produced the strongest performance, with a mean weighted AUROC of 0.718 and AUPRC of 0.324. Its AUROC remained between approximately 0.692 and 0.741 across the ten plausible values, indicating that the predictive signal was not dependent on a single achievement realization. The performance remained moderate, which is reasonable because academic resilience is influenced by multiple factors that cannot be fully represented by cross-sectional questionnaire data.

The observed differences among Logistic Regression, Random Forest, and XGBoost are consistent with broader evidence that machine-learning performance is strongly dependent on the dataset, target definition, and available predictors. Educational data-mining studies have reported competitive performance from Random Forest in both LMS-based student outcome prediction and grade classification [10], [11], while SVM has shown strong performance for classification of study duration and graduation outcomes [12]. Previous comparative machine-learning studies have similarly emphasized that algorithm selection should be based on empirical validation across the specific data context rather than accuracy from a single modelling configuration [20], [21], [23].

Contribution of Predictor Domains

The ablation analysis showed that psychological and contextual characteristics added meaningful information beyond demographic and socioeconomic variables.

Table 4. Incremental contribution of predictor domains

Predictor set	AUROC	AUPRC
Demographic + socioeconomic	0.624	0.217
+ Psychological	0.69	0.298
+ Family support	0.708	**0.325**
+ School and behavioral	**0.718**	0.324

The largest improvement occurred when psychological variables were added, increasing AUROC from 0.624 to 0.690. Family support provided a further gain, while school and behavioral variables produced a smaller additional increase. These findings suggest that academic resilience cannot be explained by socioeconomic position alone and is also associated with students' psychological characteristics and contextual support [4].

Stability of Important Predictors

SHAP analysis was performed separately for each plausible-value-specific XGBoost model. Several predictors consistently remained among the most influential.

Table 5. Most stable predictors across ten plausible-value models

Predictor	Construct	Median rank	Top-5
GROSAGR	Growth mindset	1.5	10-Oct
MATHEF21	Mathematics confidence	3	10-Oct
DISCLIM	Disciplinary climate	3	10-Sep
FAMSUP	Family support	4	10-Jul

Predictor	Construct	Median rank	Top-5
CURIOAGR	Curiosity	5	10-Jun
ESCS	Socioeconomic status	5	10-May

Growth mindset was the most stable predictor, appearing among the Top-5 features in all ten models. Disciplinary climate, family support, and curiosity also showed relatively consistent importance. These patterns are consistent with resilience research emphasizing the combined roles of personal and contextual resources [5], [18].

MATHEF21 also appeared among the Top-5 predictors in all ten models, but its simple weighted association with resilience was not consistently positive. The importance of psychological, behavioral, and contextual variables also parallels recent applications of explainable machine learning to student and adolescent data. Halid et al. [13] showed that behavioral and well-being characteristics can provide substantial predictive information when modelling the impact of digital-platform use among students. Similarly, Setiawan et al. [14] found that psychosocial and behavioral characteristics contributed strongly to machine-learning predictions of depression risk among teenagers. Although these outcomes differ from academic resilience, they further demonstrate the potential value of integrating demographic, behavioral, and psychosocial information in predictive models involving young populations. Importantly, such feature-importance findings should remain predictive rather than causal interpretations. This finding demonstrates that predictive importance does not necessarily imply a simple positive or causal relationship. XGBoost may capture nonlinear or interaction effects that are not visible in univariate comparisons.

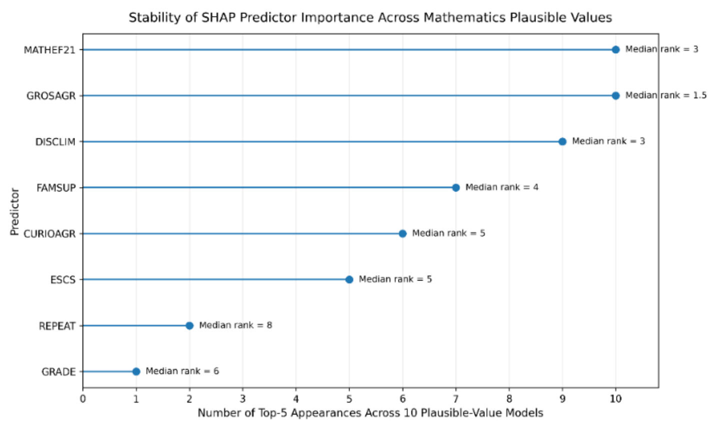

Figure 3. Stability of SHAP-based predictor importance across ten mathematics plausible-value-specific models.

Sensitivity and Robustness

Alternative resilience definitions produced the same general pattern.

Table 6. Sensitivity to alternative resilience thresholds

Definition	Unstable classification	XGBoost AUROC
20/80	22.37%	0.697
25/75	**30.61%**	**0.718**
30/70	37.17%	0.708

Classification instability remained present under all three definitions, indicating that the central finding was not limited to the 25/75 threshold. XGBoost performance also remained relatively stable.

The complete-case analysis involving 2,195 students produced an AUROC of 0.693 and AUPRC of 0.335, showing that the main predictive pattern was not solely dependent on median imputation.

Overall, the results demonstrate that academic resilience among disadvantaged Indonesian students is measurable but not completely invariant across plausible values. At the same time, growth mindset, disciplinary climate, family support, curiosity, and other student characteristics provided relatively stable predictive information. These findings support the use of all plausible values when resilience is constructed from threshold-based achievement measures and suggest that both individual and contextual factors should be considered when examining successful adaptation under socioeconomic disadvantage.

4. Conclusion

This study examined academic resilience among socioeconomically disadvantaged Indonesian students using PISA 2022 through a plausible-value-sensitive and school-aware framework. The average prevalence of academic resilience was approximately 15.20%, closely matching the official PISA 2022 estimate for Indonesia. However, resilience membership was not fully stable, as approximately 30.61% of disadvantaged students changed classification across the ten mathematics plausible values.

Among the evaluated models, XGBoost achieved the strongest performance with a weighted AUROC of 0.718 and AUPRC of 0.324. The ablation analysis showed that psychological factors provided the largest predictive improvement beyond demographic and socioeconomic characteristics, while family and school-related factors contributed additional information. Growth mindset, disciplinary climate, mathematics confidence, family support, and curiosity were among the most consistently influential predictors across plausible-value-specific models.

Sensitivity analyses using alternative resilience thresholds and complete-case data confirmed that the main findings were reasonably robust. These results indicate that academic resilience should not always be treated as a fixed binary characteristic when it is derived from threshold-based large-scale assessment scores. Considering all plausible values provides a more cautious assessment of resilience membership and the stability of its associated predictors.

Because PISA 2022 is cross-sectional, the identified relationships should be interpreted as predictive associations rather than causal effects. Future studies could extend this framework by incorporating replicate weights, additional PISA domains, longitudinal data, and cross-country comparisons.

References

[1] G. D. Borman and L. T. Overman, “Academic resilience in mathematics among poor and minority students,” *The Elementary School Journal*, vol. 104, no. 3, pp. 177–195, 2004, doi: https://doi.org/10.1086/499748.

[2] A. J. Martin and H. W. Marsh, “Academic resilience and its psychological and educational correlates: A construct validity approach,” *Psychology in the Schools*, vol. 43, no. 3, pp. 267–281, 2006, doi: https://doi.org/10.1002/pits.20149.

[3] T. Agasisti, F. Avvisati, F. Borgonovi, and S. Longobardi, “Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA,” *OECD Education Working Papers*, no. 167, OECD Publishing, Paris, France, 2018, doi: https://doi.org/10.1787/e22490ac-en.

[4] OECD, *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. Paris, France: OECD Publishing, 2023, doi: https://doi.org/10.1787/53f23881-en.

[5] W. Ye, N. Teig, and S. Blömeke, “Systematic review of protective factors related to academic resilience in children and adolescents: Unpacking the interplay of operationalization, data, and research method,” *Frontiers in Psychology*, vol. 15, Art. no. 1405786, 2024, doi: https://doi.org/10.3389/fpsyg.2024.1405786.

[6] Kismiantini, E. P. Setiawan, A. C. Pierewan, and O. A. Montesinos-López, “Growth mindset, school context, and mathematics achievement in Indonesia: A multilevel model,” *Journal on Mathematics Education*, vol. 12, no. 2, pp. 279–294, 2021, doi: https://doi.org/10.22342/jme.12.2.13690.279-294.

[7] R. Aditia and K. Széll, “Belonging matters: How context and inequalities shape student achievement in Indonesia,” *International Journal of Educational Research Open*, vol. 9, Art. no. 100512, 2025, doi: https://doi.org/10.1016/j.ijedro.2025.100512.

[8] A. I. Latifah and K. Kismiantini, "How school culture and climate mediated student's mathematics achievement: A path analysis of PISA 2022 Indonesia data," *Jurnal Pendidikan Progresif*, vol. 15, no. 3, pp. 1670–1687, 2025, doi: https://doi.org/10.23960/jpp.v15i3.pp1670-1687.

[9] W. Ye, R. Strietholt, and S. Blömeke, "Academic resilience: Underlying norms and validity of definitions," *Educational Assessment, Evaluation and Accountability*, vol. 33, pp. 169–202, 2021, doi: https://doi.org/10.1007/s11092-020-09351-7.

[10] P. Rachmadi, "A Pilot Study on Machine Learning Models for Predicting Student Pass/Fail Outcomes Using LMS Activity Logs and Interactive Dashboard Analytics," Indonesian Journal of Data and Science, vol. 7, no. 2, pp. 183–189, 2026, doi: https://doi.org/10.56705/ijodas.v7i2.415.

[11] S. Sivakumar and S. Venkataraman, "Evaluating Machine Learning Approaches: A Comparative Study of Random Forest and Neural Networks in Grade Classification," Indonesian Journal of Data and Science, vol. 6, no. 1, pp. 73–80, 2025, doi: https://doi.org/10.56705/ijodas.v6i1.240.

[12] Sumiyatun, Y. Cahyadi, and E. Faizal, "Implementation of Support Vector Machine Algorithm for Classification of Study Period and Graduation Predicate of Students," Indonesian Journal of Data and Science, vol. 6, no. 1, pp. 55–63, 2025, doi: https://doi.org/10.56705/ijodas.v6i1.214.

[13] A. Halid, D. A. Purnamasari, A. C. Saputra, and N. M. Setiohardjo, "Explainable Machine Learning for Predicting the Mental Health Impact of AI and Digital Platform Usage among Students," International Journal of Artificial Intelligence in Medical Issues, vol. 4, no. 1, pp. 1–18, 2026, doi: https://doi.org/10.56705/pxn6qg39.

[14] R. Setiawan, E. Najwaini, R. A. Azdy, and R. Rasyid, "Depression Risk Prediction Among Teenagers Using Explainable Machine Learning and Imbalanced Behavioral Data," International Journal of Artificial Intelligence in Medical Issues, vol. 4, no. 1, pp. 54–71, 2026, doi: https://doi.org/10.56705/0w9q4238.

[15] P. A. Jewsbury, Y. Jia, and E. J. Gonzalez, "Considerations for the use of plausible values in large-scale assessments," *Large-scale Assessments in Education*, vol. 12, Art. no. 24, 2024, doi: https://doi.org/10.1186/s40536-024-00213-y.

[16] K.-C. Cheung, P.-S. Sit, J.-Q. Zheng, C.-C. Lam, S.-K. Mak, and M.-K. Ieong, "A machine-learning model of academic resilience in the times of the COVID-19 pandemic: Evidence drawn from 79 countries/economies in the PISA 2022 mathematics study," *British Journal of Educational Psychology*, vol. 94, no. 4, pp. 1224–1244, 2024, doi: https://doi.org/10.1111/bjep.12715.

[17] OECD, *PISA 2022 Technical Report*. Paris, France: OECD Publishing, 2024, doi: https://doi.org/10.1787/01820d6d-en.

[18] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: https://doi.org/10.1023/A:1010933404324.

[19] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794, doi: https://doi.org/10.1145/2939672.2939785.

[20] F. Pedregosa *et al*., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[21] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, Art. no. e0118432, 2015, doi: https://doi.org/10.1371/journal.pone.0118432.

[22] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.

[23] OECD, *PISA 2022 Results (Volume V): Learning Strategies and Attitudes for Life*. Paris, France: OECD Publishing, 2024, doi: https://doi.org/10.1787/c2e44201-en.

[24] R. B. King, J. Li, and Y. Wang, "Both individual and peer growth mindsets matter for academic resilience," *npj Science of Learning*, vol. 11, Art. no. 17, 2026, doi: https://doi.org/10.1038/s41539-026-00403-z.

[25] H. Azis, Purnawansyah, F. Fattah, and I. P. Putri, "Performa Klasifikasi K-NN dan Cross-validation pada Data Pasien Pengidap Penyakit Jantung," ILKOM Jurnal Ilmiah, vol. 12, no. 2, pp. 81–86, 2020, doi: https://doi.org/10.33096/ilkom.v12i2.507.81-86.

[26] H. Azis, F. T. Admojo, and E. Susanti, "Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah," Techno.Com, vol. 19, no. 3, pp. 286–294, 2020, doi: https://doi.org/10.33633/tc.v19i3.3646.

[27] H. Azis, Purnawansyah, Nirwana, and F. A. Dwiyanto, "The Support Vector Regression Method Performance Analysis in Predicting National Staple Commodity Prices," ILKOM Jurnal Ilmiah, vol. 15, no. 2, pp. 390–397, 2023, doi: https://doi.org/10.33096/ilkom.v15i2.1686.390-397.

[28] H. Azis, Purnawansyah, and N. Alfiyyah, "Multiclass Classification on Nominal Value of Rupiah Banknotes Based on Image Processing," ILKOM Jurnal Ilmiah, vol. 16, no. 1, pp. 87–99, 2024, doi: https://doi.org/10.33096/ilkom.v16i1.1784.87-99.

[29] H. Azis and N. Rismayanti, "Prediksi Anemia dari Pixel Gambar dan Level Hemoglobin Menggunakan Random Forest Classifier," Jurnal Ilmiah NERO, vol. 9, no. 1, pp. 21–34, 2024, doi: https://doi.org/10.21107/nero.v9i1.27916.

[30] OECD, *PISA 2022 Results (Volume II): Learning During – and From – Disruption*. Paris, France: OECD Publishing, 2023, doi: https://doi.org/10.1787/a97db61c-en.