AI and Developmental Insights in Education

AI and Developmental Insights in Education Vol. 2 No. 1 (2026)

Check for updates

OPEN ACCESS

Machine Learning Models for Predicting Student Vulnerability to Academic Stress in AI-Integrated Learning Environments

Alvina Desya Ramadhani [1] and David Naista [1]

[1]Universitas Lampung, Indonesia
[2]Universitas Islam Negeri Raden Intan Lampung, Indonesia

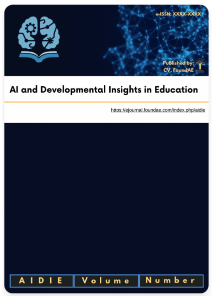
Received: 15 February 2026	Revised: 21 March 2026	Accepted: 18 April 2026	Online: 31 May 2026

Abstract

Artificial intelligence (AI) tools are increasingly embedded in higher education, yet their relationship with students' academic stress remains insufficiently established. This study developed and internally evaluated machine learning (ML) models for classifying academic stress vulnerability among 441 undergraduate students enrolled in AI-integrated courses at three Indonesian public universities. Using a quantitative cross-sectional predictive design, data were collected through psychometric scales, institutional GPA records, and LMS behavioral indicators. Four supervised classifiers, Gradient Boosting (GB), Random Forest (RF), Support Vector Machine with radial basis kernel (SVM-RBF), and Logistic Regression (LR), were compared using stratified train-test evaluation and five-fold cross-validation within the training data. GB achieved the strongest held-out performance (accuracy = .846, macro-F1 = .840, AUC-ROC = .930). Permutation importance indicated that cognitive load, AI literacy, and digital fatigue contributed most to classification performance. Subgroup AUC comparisons showed no significant differences across gender and discipline, although this should be interpreted cautiously. External validation and ethical governance are required before operational deployment.

Keywords: academic stress; cognitive load; digital fatigue; learning analytics

Corresponding Author:
Alvina Desya Ramadhani
Email: alvinaaramadhani15@gmail.com
https://orcid.org/0009-0009-9154-8525



Author Note:
This study received academic and institutional support from the Faculty of Tarbiya and Teacher Training. The authors declare that they have no conflicts of interest. Correspondence concerning this article should be addressed to Alvina Desya Ramadhani.

Introduction

Artificial intelligence (AI) has moved from a peripheral educational technology to a routine layer of higher education practice, shaping how students receive feedback, search for information, draft assignments, interact with digital platforms, and regulate study time. Recent reviews show that AI applications in higher education are increasingly used for tutoring, assessment, analytics, content generation, and personalized support, while also raising concerns about transparency, learner agency, academic integrity, privacy, and pedagogical alignment (Chiu et al., 2023; Crompton & Burke, 2023; Kasneci et al., 2023). Student-facing evidence similarly indicates that generative AI is perceived as useful for productivity, idea generation, and individualized learning support, but students also report concerns about unreliable outputs, over-dependence, unequal access, and uncertainty about acceptable use (Chan & Hu, 2023; Campillo-Ferrer et al., 2025). These tensions suggest that AI-integrated learning environments are not merely technical innovations; they are socio-cognitive environments that may redistribute cognitive effort, emotional pressure, and responsibility for learning.

Academic stress remains a persistent concern in higher education and is associated with lower well-being, reduced academic engagement, and vulnerability to burnout or withdrawal (Barbayannis et al., 2022; Olson et al., 2025). In AI-supported courses, stress may arise not only from conventional workload and assessment demands, but also from students' perceived ability to interpret AI feedback, verify generated information, comply with shifting institutional rules, and manage sustained digital engagement. Research on AI literacy has therefore become increasingly important because students' capacity to understand, evaluate, and use AI systems can influence how they experience AI-supported learning tasks (Carolus et al., 2023; Gümüş & Kara, 2025; Koch et al., 2024; Laupichler et al., 2022). At the same time, work on cognitive load and digital learning indicates that technology-rich environments may increase extraneous processing demands when interfaces, feedback, or task expectations are not well aligned with learners' prior knowledge (Skulmowski & Xu, 2022). These developments position academic stress vulnerability as an outcome that should be examined through cognitive, regulatory, affective, and technological indicators rather than through general stress measures alone.

Learning analytics and educational data mining have produced a substantial body of predictive studies on academic achievement, dropout risk, engagement, and online persistence. Recent studies and reviews confirm that machine learning can support earlier identification of students who may need intervention, especially when behavioral traces are combined with theoretically meaningful learner variables (Alalawi et al., 2025; Bañeres et al., 2023; Christou et al., 2023; Yağcı, 2022). However, the methodological center of this literature is still strongly performance-oriented. In comparison, fewer studies have treated academic stress or stress vulnerability as the primary prediction target, and systematic reviews of student mental-health prediction emphasize that stress-detection studies often rely on heterogeneous measures, limited validation strategies, and insufficient integration with educational theory (Daza et al., 2023; Drira et al., 2024; Liu et al., 2025; Schaab et al., 2024). This creates a substantive gap: institutions are increasingly able to predict academic risk, but they have less evidence on how to predict psychologically meaningful vulnerability in AI-mediated learning contexts without reducing student well-being to generic engagement or performance proxies.

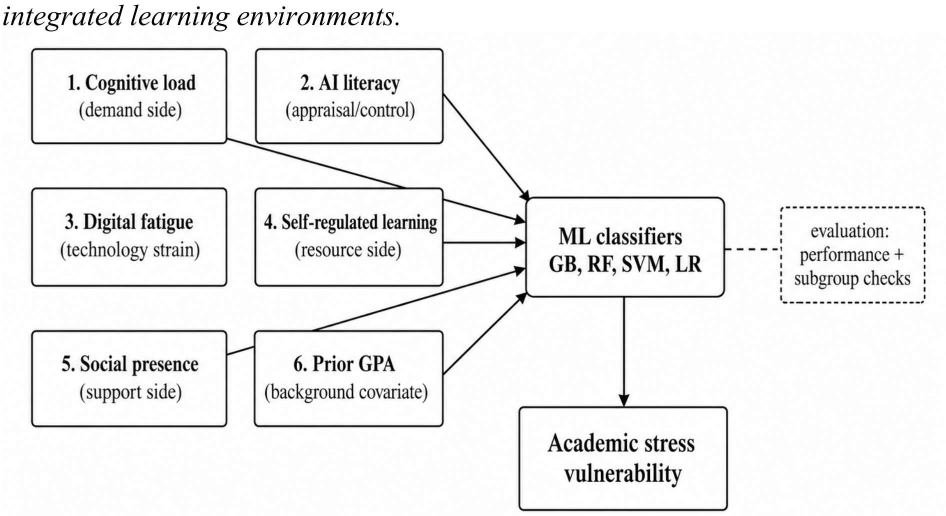
A second limitation in the current literature concerns model governance and subgroup robustness. Predictive educational systems can reproduce bias when training data, feature selection, or intervention rules reflect unequal institutional histories or when model performance varies across student groups (Baker & Hawn, 2022). Recent fairness-aware learning analytics work shows that early-warning systems must balance timeliness, predictive accuracy, and equity rather than optimizing accuracy alone (Cheng et al., 2025). This issue is especially relevant for AI-integrated learning environments because students differ in AI access, disciplinary norms, digital confidence, prior preparation, and perceived legitimacy of AI use. Consequently, any model that predicts academic stress vulnerability should be interpreted not only in terms of AUC-ROC or F1-score, but also in terms of whether its predictive behavior appears reasonably consistent across relevant student subgroups.

Against this background, the present study develops and internally evaluates a machine learning approach for classifying academic stress vulnerability among undergraduate students enrolled in AI-integrated courses. The study is distinctive in three ways. First, it positions stress vulnerability as the focal outcome rather than as an incidental correlate of performance, engagement, or dropout. Second, it organizes the predictor set through the Stress-Vulnerability Prediction Framework (SVPF), which combines cognitive load, AI literacy, digital fatigue, self-regulated learning, social presence, and prior GPA as theory-selected indicators. Third, it compares multiple classifiers while also conducting preliminary subgroup performance checks by gender and academic discipline. The aim is not to claim that the resulting model is ready for autonomous institutional deployment, but to provide a cautious empirical basis for future theory-informed, human-supervised, and fairness-sensitive early-support systems in AI-integrated higher education.

The study was guided by three research questions: (RQ1) Which classifier among Gradient Boosting, Random Forest, SVM-RBF, and Logistic Regression shows the strongest internal performance in classifying academic stress vulnerability? (RQ2) Which theoretically selected indicators contribute most strongly to the best-performing model? (RQ3) Does the model show comparable AUC-ROC values across gender and academic discipline subgroups? It was expected that ensemble models would outperform the linear baseline, that cognitive load and AI literacy would rank among the most influential predictors, and that subgroup performance would not show large internal discrepancies. These expectations were treated as empirical propositions rather than as evidence of causality or deployment readiness.

Figure 1 presents the proposed SVPF used to organize the study variables. The figure is intended as a conceptual map rather than a causal diagram; it shows how cognitive, technological, regulatory, social, and academic-record indicators are combined for supervised classification of academic stress vulnerability.

Figure 1
Stress-Vulnerability Prediction Framework (SVPF) for classifying academic stress vulnerability in AI-integrated learning environments.



Note. The framework positions cognitive load, AI literacy, digital fatigue, self-regulated learning, social presence, and prior GPA as theoretically selected predictors. GB = Gradient Boosting; RF = Random Forest; SVM = Support Vector Machine; LR = Logistic Regression.

As shown in Figure 1, the SVPF treats stress vulnerability as a classification outcome derived from multiple learner and environment indicators. This framing prevents the model from being interpreted as a causal explanation and supports a more cautious interpretation of feature importance.

Methods

Research Design

This study used a quantitative cross-sectional predictive design. Psychometric indicators, LMS engagement traces, and institutional GPA records were collected during an ongoing semester and used to train and compare four supervised classifiers. The design was appropriate for estimating classification performance and feature contribution, but it does not support causal inference about the effects of AI use on academic stress.

Participants and Context

Participants were undergraduate students enrolled in AI-integrated courses at three public universities in Lampung Province, Indonesia. Inclusion criteria required enrollment in at least one course using AI-supported instructional tools or AI-assisted feedback, a minimum of four weeks of platform engagement, and written informed consent. Students who reported formal psychological intervention for stress-related disorders during the preceding six months were excluded to reduce measurement confounding. After exclusions, the final analytic sample consisted of 441 students (mean age = 20.47 years, SD = 1.63, range = 18-27).

Table 1 summarizes the demographic composition of the analytic sample. The table was corrected so that all category frequencies sum to the stated final sample size of N = 441.

Table 1

Demographic Characteristics of the Analytic Sample (N = 441)

Characteristic	n	%	M (SD)
Age (years)	-	-	20.47 (1.63)
Gender			
Female	257	58.3	
Male	184	41.7	
Academic discipline			
Education	150	34.0	
STEM	137	31.1	
Social Sciences	92	20.9	
Humanities	62	14.1	
AI familiarity			
Low	97	22.0	
Intermediate	234	53.1	
High	110	24.9	

Note. Percentages were recalculated using N = 441 and may not total exactly 100% because of rounding. STEM = science, technology, engineering, and mathematics.

The final sample retained representation across gender, disciplinary area, and AI familiarity level. However, because the sample was restricted to three universities in one Indonesian province, generalizability should be considered regional and provisional.

Sampling and Recruitment

Students were recruited through stratified purposive sampling, with strata based on university and academic discipline. Course coordinators distributed information sheets and consent forms. Of 495 students initially approached, 452 met the eligibility criteria and consented to participate; seven response sets were removed for careless responding and four cases were removed because LMS trace data were incomplete, yielding N = 441. Participation was voluntary and no academic incentives were offered.

Sample Size, Missing Data, and Precision

The achieved sample exceeded the common minimum used in comparable educational prediction studies with a small feature set, but the study was still treated as an internal validation exercise rather than as evidence of deployment readiness. Missing item-level data represented 2.3% of total entries. Multiple imputation by chained equations was used before model training, and Little's MCAR test did not reject the missing-completely-at-random assumption, chi-square(312) = 318.44, p = .391.

Measures and Data Sources

Six psychometric constructs and one academic-record covariate were used. The instrument labels were retained for readability, but the revised manuscript avoids claiming universal validation of every scale in this specific context without reporting full adaptation evidence. Internal consistency and convergent validity results are reported descriptively.

Table 2 reports the descriptive and reliability statistics for the study variables. These statistics provide context for the ML feature set and also allow readers to check whether any construct showed unusually weak internal consistency.

Table 2

Psychometric Properties of Study Instruments

Scale	Items	M	SD	Range	alpha	AVE
ASVS (DV)	22	3.42	0.71	1-5	.92	.61
AILS	18	3.15	0.84	1-5	.87	.54
CLS	15	3.67	0.79	1-5	.83	.51
MSLQ-SF	20	3.38	0.68	1-5	.91	.57
SPS	12	3.22	0.77	1-5	.84	.52
DFS	10	3.55	0.82	1-5	.79	.49
Prior GPA	-	3.21	0.43	1-4	-	-

Note. ASVS = Academic Stress Vulnerability Scale; AILS = AI Literacy Scale; CLS = Cognitive Load Scale; MSLQ-SF = Motivated Strategies for Learning Questionnaire-Short Form; SPS = Social Presence Scale; DFS = Digital Fatigue Scale; GPA = grade point average; AVE = average variance extracted.

As shown in Table 2, all multi-item measures met conventional internal-consistency expectations. The DFS AVE value was slightly below .50, so digital fatigue results should be interpreted with appropriate caution and should be replicated with additional measurement validation.

Data Collection Procedures

Data were collected during weeks 11 and 12 of the semester to ensure that students had already experienced sustained AI-supported course activities. Online surveys were administered through Google Forms linked in each LMS. LMS behavioral traces, including login frequency, AI-tool engagement duration, and platform dwell time, were exported by institutional data administrators and merged with survey data through anonymized identifiers. The survey order was counterbalanced to reduce order effects.

Data Analysis

Analyses were conducted in Python using scikit-learn. The continuous ASVS score was dichotomized at the upper tertile to define a high-vulnerability group. Logistic Regression, SVM-RBF, Random Forest, and Gradient Boosting were compared using a stratified 80/20 train-test split. Hyperparameters were tuned with five-fold stratified cross-validation on the training data using macro-F1 as the selection criterion. Performance was evaluated with accuracy, macro-F1, precision, and AUC-ROC. Permutation importance with 30 repeats was used to estimate feature contribution. Subgroup checks compared AUC-ROC values by gender and disciplinary group; these checks were framed as preliminary bias screening, not as a full fairness audit.

Validity, Reliability, and Ethical Considerations

Internal consistency was assessed with Cronbach's alpha. Confirmatory factor analysis indicated acceptable fit for the measurement model (CFI > .95, RMSEA < .06, SRMR < .08), and multicollinearity diagnostics were acceptable (VIF range = 1.21-2.87). The study was approved by the Research Ethics Committee of Universitas Islam Negeri Raden Intan Lampung. Written informed consent was obtained before data collection. Participants were informed that the model output would be used only for research analysis and not for academic grading, counseling triage, or disciplinary action.

Results

Participant Flow and Preliminary Analysis

The final analytic sample consisted of 441 students. The prevalence of high academic stress vulnerability, defined as ASVS scores in the upper tertile, was 34.2% ($n = 151$). Predictor distributions showed acceptable univariate normality for descriptive analysis, with skewness ranging from -0.43 to 0.61 and kurtosis from -0.92 to 1.14. Cognitive load correlated positively with ASVS scores ($r = .62$, $p < .001$), whereas AI literacy correlated negatively with ASVS scores ($r = -.51$, $p < .001$). These correlations are descriptive and should not be interpreted causally.

Model Performance Comparison

Table 3 presents held-out test metrics for the four classifiers. The table is followed by Figure 2 to make the performance pattern easier to inspect visually.

Table 3

Machine Learning Classifier Performance Metrics on the Held-Out Test Set (n = 88)

Classifier	Accuracy	Precision	F1-score	AUC-ROC
Gradient Boosting	.846	.843	.840	.930
Random Forest	.832	.828	.820	.910

Classifier	Accuracy	Precision	F1-score	AUC-ROC
SVM-RBF	.795	.783	.770	.870
Logistic Regression	.742	.731	.710	.820

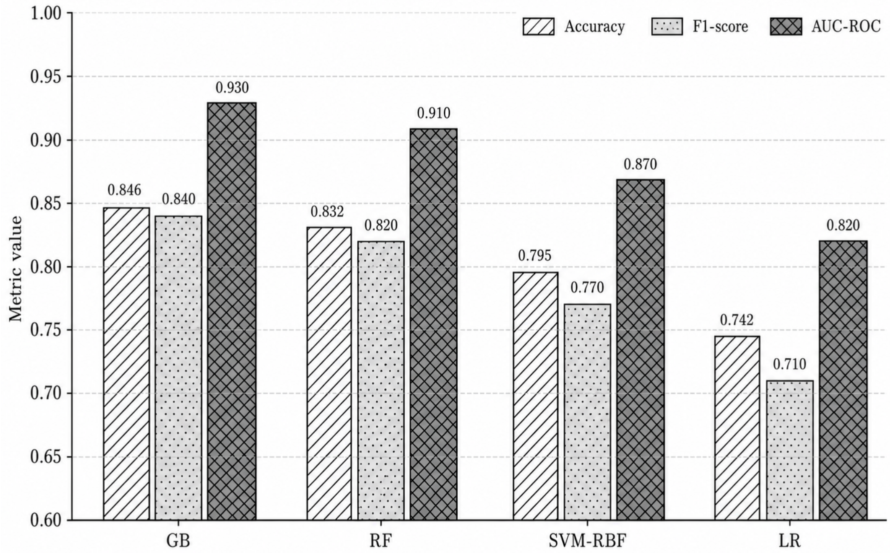
Note. Metrics are macro-averaged across the binary classes. AUC-ROC = area under the receiver operating characteristic curve; SVM-RBF = support vector machine with radial basis function kernel.

The ensemble models performed better than the linear baseline in this internal evaluation, with Gradient Boosting showing the highest AUC-ROC. Because the difference between Gradient Boosting and Random Forest was small, the result should be interpreted as evidence of comparable ensemble advantage rather than as definitive superiority of one model.

Figure 2 visualizes the same model-performance pattern across accuracy, macro-F1, and AUC-ROC. This figure is useful because it shows that the ranking of models was consistent across all three metrics.

Figure 2

Comparison of classifier performance for predicting academic stress vulnerability.



Note. GB = Gradient Boosting; RF = Random Forest; SVM-RBF = Support Vector Machine with radial basis function kernel; LR = Logistic Regression. Values represent held-out test performance.

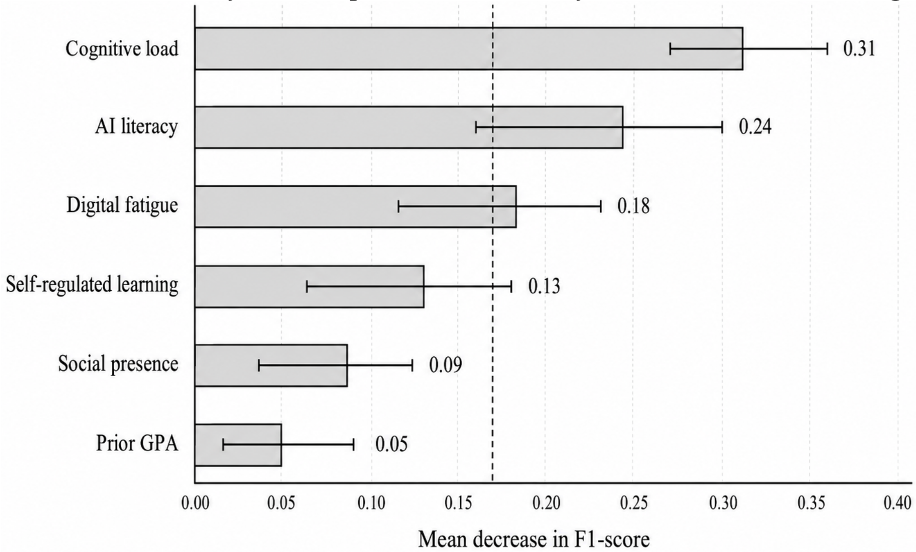
The visual pattern in Figure 2 supports the use of ensemble methods for this dataset, but it does not establish that the model will generalize to different institutions, semesters, or AI-learning platforms. External validation remains necessary.

Feature Importance Analysis

Figure 3 presents permutation-based importance scores from the Gradient Boosting model. The scores indicate how much macro-F1 decreased when each feature was permuted, not the causal effect of the feature on stress vulnerability.

Figure 3

Permutation-based feature importance estimates from the Gradient Boosting classifier.



Note. Error bars show 95% confidence intervals across 30 permutation repeats. The dashed vertical line indicates the mean importance value across predictors.

Figure 3 indicates that cognitive load, AI literacy, and digital fatigue were the highest-ranked predictors. The ranking is theoretically plausible, but it should be treated as model-specific because permutation importance can change when predictors are correlated or when the model is retrained on a new sample.

Subgroup Performance Checks

The Gradient Boosting model was evaluated by gender and discipline as a preliminary check of performance consistency. AUC-ROC values were .92 for female students and .93 for male students; the AUC difference was not statistically significant ($z = 0.48$, $p = .631$). Across academic disciplines, AUC-ROC values ranged from .90 for Humanities to .94 for STEM, with no pairwise difference remaining significant after Bonferroni adjustment. These results indicate no detected subgroup performance difference in the examined strata, but they should not be described as a complete fairness audit because other protected attributes, calibration, false-positive rates, and decision consequences were not evaluated.

Discussion

Principal Findings

This study internally evaluated four machine learning classifiers for predicting academic stress vulnerability in AI-integrated undergraduate courses. Gradient Boosting achieved the strongest overall performance, followed closely by Random Forest, whereas SVM-RBF and Logistic Regression showed lower performance. This pattern is consistent with recent educational prediction studies showing that ensemble models can capture nonlinear relationships among learner, behavioral, and contextual indicators more effectively than simpler baseline models (Alalawi et al., 2025; Christou et al., 2023; Yağcı, 2022). Nevertheless, the observed advantage should be interpreted as internal predictive evidence from one dataset rather than proof of universal algorithmic superiority. The small performance difference between Gradient Boosting and Random Forest also indicates that model selection should consider interpretability, computational simplicity, calibration, and institutional feasibility alongside headline accuracy metrics.

The findings extend current work on learning analytics by shifting the prediction target from academic performance or dropout to academic stress vulnerability. Reviews of student stress prediction show that psychological outcomes are increasingly analyzed with machine learning, but many studies remain detached from educational theory or rely on broad stress indicators that are not specifically situated in digitally mediated learning (Daza et al., 2023; Drira et al., 2024; Liu et al., 2025). By integrating cognitive load, AI literacy, digital fatigue, self-regulated learning, social presence, and GPA, the present model provides a more education-specific account of why some students may become vulnerable in AI-supported learning environments. This does not establish causality, but it clarifies a theory-informed set of variables that can be tested more rigorously in longitudinal and intervention research.

Interpretation of Predictor Importance

Cognitive load was the strongest predictor in the Gradient Boosting model. This result is theoretically coherent because AI-supported coursework can increase processing demands when students must evaluate automated feedback, compare AI-generated explanations with course materials, or decide whether an AI response is trustworthy. Recent cognitive-load scholarship in digital and online learning emphasizes that technological complexity can become extraneous load when instructional design does not sufficiently guide learners through information selection, integration, and verification (Skulmowski & Xu, 2022). In this sense, cognitive load should not be understood only as task difficulty; it may also reflect the additional mental work required to manage AI-mediated uncertainty. The practical implication is that AI integration should be accompanied by explicit scaffolding, transparent feedback design, and manageable information flow rather than simply adding AI tools to existing course structures.

AI literacy emerged as the second most influential predictor, reinforcing the argument that students' capacity to understand and critically use AI tools may function as a protective resource. Recent AI-literacy research has moved beyond technical competence and increasingly emphasizes evaluation, ethical judgment, limitations of AI outputs, and informed human oversight (Carolus et al., 2023; Gümüş & Kara, 2025; Koch et al., 2024; Laupichler et al., 2023). Students with lower AI literacy may experience AI-supported tasks as less controllable, especially when feedback is opaque or when institutional policies about acceptable AI use remain ambiguous. This interpretation is compatible with recent student-perception studies showing that learners value GenAI support but are also concerned about accuracy, privacy, dependence, and academic integrity (Campillo-Ferrer et al., 2025; Chan & Hu, 2023). Therefore, AI literacy should be treated not only as a digital competence outcome but also as a potential component of student well-being support in AI-integrated courses.

Digital fatigue ranked below cognitive load and AI literacy but remained substantively important. This is notable because AI-supported learning often increases screen time, platform switching, prompt iteration, and continuous feedback exposure. Recent evidence on technostress and online learning quality suggests that institutional and instructor support can reduce the negative effects of technology demands on students (Saleem et al., 2024). The present result suggests that digital fatigue should be measured directly in future AI-in-education research rather than absorbed into generic engagement indicators. Self-regulated learning and social presence showed smaller but theoretically meaningful contributions, indicating that planning, monitoring, peer connection, and perceived instructor or peer availability may help students manage AI-mediated learning demands, although these variables were less dominant than cognitive and AI-specific indicators in this dataset (Alhazbi et al., 2024; Jin et al., 2023; Mykota, 2025).

Subgroup Performance and Ethical Interpretation

The subgroup analysis did not detect meaningful AUC-ROC differences by gender or academic discipline. This is encouraging because it suggests that the selected feature set did not produce obvious internal performance disparities across the examined groups. However, this finding should be framed as a preliminary performance-consistency check rather than as definitive evidence of fairness. Algorithmic bias can emerge through sampling imbalance, omitted variables, measurement inequivalence, historical inequities, and unequal consequences of intervention decisions even when aggregate performance metrics look similar (Baker & Hawn, 2022). Recent fairness-aware learning analytics studies further show that early prediction should balance accuracy with equitable model behavior and intervention design (Cheng et al., 2025). For this reason, any future use of a stress-vulnerability model should include ongoing fairness audits, student-facing transparency, and human review before any consequential action is taken.

Strengths and Limitations

The study has four main strengths: a theory-informed feature set, comparison of multiple classifiers, permutation-based feature interpretation, and preliminary subgroup performance checks. It also contributes to a growing discussion on how AI in higher education affects not only learning efficiency but also student experience, psychological burden, and governance (Chiu et al., 2023; Crompton & Burke, 2023; Kasneci et al., 2023). Despite these strengths, several limitations remain. First, the cross-sectional design prevents causal inference and cannot show whether cognitive load, AI literacy, or digital fatigue precede changes in stress vulnerability. Second, the data were collected from three universities in one Indonesian province, so institutional and cultural generalizability remains limited. Third, the stress outcome was dichotomized for classification, which supports interpretability but reduces information contained in the continuous scale. Fourth, several predictors were self-reported, raising the possibility of shared-method variance. Fifth, the subgroup analysis considered only gender and broad academic discipline; it did not examine socioeconomic status, disability status, first-generation status, language background, access to paid AI tools, or other variables that may influence both AI use and stress vulnerability.

Implications

For researchers, the findings support more precise study of academic stress vulnerability as a distinct outcome in AI-integrated higher education. For institutions, the model should be viewed as a decision-support prototype rather than an automated diagnostic system. A responsible implementation pathway would require external validation, calibration checks, privacy protection, student consent, opt-out mechanisms, clear communication about model purpose, and human-supervised referral procedures. For instructors and instructional designers, the findings point to three practical priorities: reduce unnecessary cognitive load in AI-mediated tasks, provide explicit AI-literacy instruction, and monitor digital fatigue during sustained platform-based learning. For policymakers, the results reinforce the need for governance standards that require predictive systems in education to be explainable, auditable, and oriented toward supportive intervention rather than surveillance or punishment.

Recommendations for Future Research

Future studies should test the SVPF in longitudinal designs that follow students across multiple course phases, because stress vulnerability is likely to change as students encounter new assessment demands, AI policies, and feedback routines. Multi-institutional and crossnational studies are also needed to examine whether the model remains stable across different AI infrastructures, disciplinary cultures, and student populations. Methodologically, future work should compare classification models with regression and survival-analysis approaches so that continuous changes in stress can be modeled without information loss. Researchers should also incorporate passive and multimodal indicators, such as learning-platform event sequences, writing-process data, help-seeking logs, or physiological markers, while maintaining strict privacy protections. Finally, intervention studies are needed to determine whether AI-literacy training, cognitive-load reduction, workload redesign, and instructor support can reduce predicted vulnerability rather than merely identify it. These studies should evaluate not only predictive accuracy but also fairness, student trust, intervention uptake, and unintended consequences.

Conclusion

This study developed and internally evaluated ML models for classifying academic stress vulnerability among 441 undergraduate students in AI-integrated courses. Gradient Boosting showed the strongest held-out performance, while Random Forest achieved comparable results. Permutation importance identified cognitive load, AI literacy, and digital fatigue as the most influential predictors in the best-performing model. Preliminary subgroup checks did not detect AUC-ROC differences by gender or discipline, but this result should not be overstated as full algorithmic fairness. Overall, the revised evidence supports the SVPF as a useful research framework for studying stress vulnerability in AI-integrated higher education. Longitudinal, multisite, and externally validated studies are needed before institutional implementation can be recommended.

Author Contributions

Conceptualization, ADR and DN; Methodology, DN; Software, DN; Validation, ADR and DN; Formal Analysis, DN; Investigation, ADR and DN; Resources, ADR; Data Curation, DN; Writing - Original Draft, ADR; Writing - Review & Editing, ADR and DN; Visualization, DN; Supervision, DN; Project Administration, ADR; Funding Acquisition, DN. ADR = Alvina Desya Ramadhani; DN = David Naista.

Institutional Review Board Statement

This study was conducted in accordance with the Declaration of Helsinki and approved by the Research Ethics Committee of Universitas Islam Negeri Raden Intan Lampung. Written informed consent was obtained from all participants before data collection.

Data Availability Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request, subject to institutional ethics approval and participant-confidentiality restrictions.

Conflicts of Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Generative AI Statement

Generative AI tools were used only for language refinement and editorial checking. All conceptual content, data analysis, interpretations, and final revisions were reviewed and approved by the authors.

References

Alalawi, K., Athauda, R., & Chiong, R. (2025). An extended learning analytics framework integrating machine learning and pedagogical approaches for student performance prediction and intervention. International Journal of Artificial Intelligence in Education, 35, 1239-1287. https://doi.org/10.1007/s40593-024-00429-7

Aldosari, A. M. (2022). Improving social presence in online higher education. Frontiers in Psychology, 13, 994403. https://doi.org/10.3389/fpsyg.2022.994403

Alhazbi, S., Al-ali, A., Tabassum, A., Al-Ali, A., Al-Emadi, A., Khattab, T., & Hasan, M. A. (2024). Using learning analytics to measure self-regulated learning: A systematic review of empirical studies in higher education. Journal of Computer Assisted Learning, 40(4), 1658-1674. https://doi.org/10.1111/jcal.12982

An, R., Qian, G., et al. (2025). Digital fatigue and academic resilience among university students with grit and flexibility as mediators. Scientific Reports, 15, 45407. https://doi.org/10.1038/s41598-025-29313-7

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. International Journal of Artificial Intelligence in Education, 32(4), 1052-1092. https://doi.org/10.1007/s40593-021-00285-9

Bañeres, D., Rodríguez-González, M. E., Guerrero-Roldán, A. E., & Karadeniz, A. (2023). An early warning system to identify and intervene online dropout learners. International Journal of Educational Technology in Higher Education, 20, 3. https://doi.org/10.1186/s41239-022-00371-5

Barbayannis, G., Bandari, M., Zheng, X., Baquerizo, H., Pecor, K. W., & Ming, X. (2022). Academic stress and mental well-being in college students: Correlations, affected groups, and COVID-19. Frontiers in Psychology, 13, 886344. https://doi.org/10.3389/fpsyg.2022.886344

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS - Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. Computers in Human Behavior: Artificial Humans, 1(2), 100014. https://doi.org/10.1016/j.chbah.2023.100014

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20, 43. https://doi.org/10.1186/s41239-023-00411-8

Cheng, A., Pei, B., & Liu, C. (2025). Balancing act: Early, fair, and accurate identification of at-risk students. Journal of Learning Analytics, 12(3), 47-65. https://doi.org/10.18608/jla.2025.8761

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, 100118. https://doi.org/10.1016/j.caeai.2022.100118

Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., Antoniadis, N., Glavas, E., & Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. Expert Systems with Applications, 225, 120079. https://doi.org/10.1016/j.eswa.2023.120079

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, 22. https://doi.org/10.1186/s41239-023-00392-8

Daza, A., Saboya, N., Necochea-Chamorro, J. I., Zavaleta Ramos, K., & Vásquez Valencia, Y. d. R. (2023). Systematic review of machine learning techniques to predict anxiety and stress in college students. Informatics in Medicine Unlocked, 43, 101391. https://doi.org/10.1016/j.imu.2023.101391

de Filippis, R., & Al Foysal, A. (2024). Comprehensive analysis of stress factors affecting students: A machine learning approach. Discover Artificial Intelligence, 4, 96. https://doi.org/10.1007/s44163-024-00169-6

Drira, M., et al. (2024). Machine learning methods in student mental health research: A systematic review. Applied Sciences, 14(24), 11738. https://doi.org/10.3390/app142411738

Gümüş, M. M., & Kara, M. (2025). Development and validation of the Generative AI Literacy for Learning Scale (GenAI-LLs). Australasian Journal of Educational Technology, 41(4), 1-16. https://doi.org/10.14742/ajet.10236

Hemmler, Y., & Ifenthaler, D. (2024). Self-regulated learning strategies in continuing education: A systematic review and meta-analysis. Educational Research Review, 45, 100629. https://doi.org/10.1016/j.edurev.2024.100629

Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. International Journal of Educational Technology in Higher Education, 20, 37. https://doi.org/10.1186/s41239-023-00406-5

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kleimola, R., Hämäläinen, R., & Cattaneo, A. (2025). Promoting higher education students' self-regulated learning through learning analytics: A qualitative study. Education and Information Technologies. https://doi.org/10.1007/s10639-024-12978-4

Koch, M. J., Wienrich, C., Straka, S., Latoschik, M. E., & Carolus, A. (2024). Overview and confirmatory and exploratory factor analysis of AI literacy scales. Computers and Education: Artificial Intelligence, 7, 100310. https://doi.org/10.1016/j.caeai.2024.100310

Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T. (2023). Development of the Scale for the Assessment of Non-Experts' AI Literacy: An exploratory factor analysis. Computers in Human Behavior Reports, 12, 100338. https://doi.org/10.1016/j.chbr.2023.100338

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. Computers and Education: Artificial Intelligence, 3, 100101. https://doi.org/10.1016/j.caeai.2022.100101

Liu, S., Zhang, Y., Zhao, L., & Liu, Z. (2025). Academic stress detection based on multisource data: A systematic review from 2012 to 2024. Interactive Learning Environments. https://doi.org/10.1080/10494820.2024.2387744

Luo, R.-Z., & Zhou, Y.-L. (2024). The effectiveness of self-regulated learning strategies in higher education blended learning: A five years systematic review. Journal of Computer Assisted Learning, 40(6), 3005-3029. https://doi.org/10.1111/jcal.13052

Mansoor, H. M. H., et al. (2024). Artificial intelligence literacy among university students: A comparative transnational survey. Frontiers in Communication, 9, 1478476. https://doi.org/10.3389/fcomm.2024.1478476

Mykota, D. (2025). A meta-analysis of social presence in higher education online education. International Journal of E-Learning & Distance Education, 40(1).

Olson, N., Oberhoffer-Fritz, R., Reiner, B., et al. (2025). Stress, student burnout and study engagement: A cross-sectional comparison of university students of different academic subjects. BMC Psychology, 13, 293. https://doi.org/10.1186/s40359-025-02602-6

Saleem, F., Chikhaoui, E., & Malik, M. I. (2024). Technostress in students and quality of online learning: Role of instructor and university support. Frontiers in Education, 9, 1309642. https://doi.org/10.3389/feduc.2024.1309642

Schaab, B. L., Calvetti, P. U., Hoffmann, S., Diaz, G. B., Rech, M., Cazella, S. C., Stein, A. T., Barros, H. M. T., da Silva, P. C., & Reppold, C. T. (2024). How do machine learning models perform in the detection of depression, anxiety, and stress among undergraduate students? A systematic review. Cadernos de Saúde Pública, 40(11), e00029323. https://doi.org/10.1590/0102-311XEN029323

Shehzad, N., Lashari, S. A., Lashari, T. A., & Hasan, M. K. (2023). Exploring the impact of instructor social presence on student engagement in online higher education. Contemporary Educational Technology, 15(4), ep488. https://doi.org/10.30935/cedtech/13823

Skulmowski, A., & Xu, K. M. (2022). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. Educational Psychology Review, 34, 171-196. https://doi.org/10.1007/s10648-021-09624-7

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9, 11. https://doi.org/10.1186/s40561-022-00192-z