Journal Website

Article history:
Received 23 September 2025
Revised 02 December 2025
Accepted 09 December 2025
Published Online 10 February 2026

Journal of Adolescent and Youth Psychological Studies

Volume 7, Issue 2, pp 1-11

E-ISSN: 2981-2526

Explainable Machine Learning Prediction of Dropout Risk Using Psychosocial and Cognitive Variables

Eirik. Solheim[1], Florian. Reimann[2*], Andrés. Mejía[3]

[1] Department of Clinical Psychology, University of Bergen, Bergen, Norway
[2] Department of Neuropsychology, Heidelberg University, Heidelberg, Germany
[3] Department of Social Psychology, Universidad del Valle, Cali, Colombia

*** Corresponding author email address**: florian.reimann@uni-heidelberg.de

A r t i c l e I n f o

Article type:
Original Research

How to cite this article:
Solheim, E., Reimann, F., & Mejía, A. (2026). Explainable Machine Learning Prediction of Dropout Risk Using Psychosocial and Cognitive Variables. *Journal of Adolescent and Youth Psychological Studies, 7*(2), 1-11.
http://dx.doi.org/10.61838/kman.jayps.5089

A B S T R A C T

Objective: The present study aimed to develop and validate an explainable machine learning model capable of accurately predicting school dropout risk by integrating psychosocial and cognitive variables.

Methods and Materials: A cross-sectional predictive design was employed with a sample of 1.172 secondary school students from three federal states in Germany. Standardized instruments were used to assess psychosocial variables including depressive symptoms, academic self-efficacy, school belonging, teacher and peer support, self-regulation, and academic motivation. Cognitive performance was measured through computerized tasks assessing working memory, processing speed, and fluid reasoning. Socioeconomic status and migration background were included as contextual covariates. Data preprocessing involved multiple imputation, normalization, and class imbalance correction using SMOTE. The dataset was partitioned into training, validation, and independent test subsets. Multiple supervised learning algorithms—logistic regression with elastic net regularization, support vector machines, random forest, and gradient boosting—were trained and compared using cross-validated hyperparameter optimization. Model performance was evaluated using AUC, balanced accuracy, F1-score, and calibration indices. Explainability was ensured through SHAP-based global and local feature attribution analyses.

Findings: Gradient boosting achieved the highest predictive performance (AUC = 0.92; balanced accuracy = 0.86), significantly outperforming linear models. Psychosocial variables demonstrated stronger predictive power than cognitive variables alone, yet the integration of both domains significantly improved overall model accuracy. Depressive symptoms, academic self-efficacy, and school belonging emerged as the most influential predictors, while processing speed and working memory provided incremental predictive validity. Nonlinear threshold effects were observed, indicating that elevated emotional distress and reduced cognitive efficiency substantially increased dropout probability.

Conclusion: The findings demonstrate that explainable machine learning models integrating psychosocial and cognitive indicators can reliably predict dropout risk while preserving interpretability.

Keywords: *School dropout; explainable artificial intelligence; machine learning; cognitive functioning; early warning systems; educational risk modeling.*

1. Introduction

School dropout remains one of the most pressing educational and public health challenges globally, given its profound implications for lifetime socioeconomic outcomes, mental health trajectories, and civic participation. Beyond immediate academic disengagement, dropout is associated with increased risk for unemployment, psychiatric vulnerability, health-compromising behaviors, and reduced life expectancy. Contemporary research increasingly conceptualizes dropout not as a singular academic failure but as a multidimensional developmental outcome shaped by complex interactions among psychosocial, cognitive, neurobiological, and contextual determinants. This integrative understanding aligns with systemic risk frameworks that emphasize cumulative vulnerability across domains of functioning (Nielsen et al., 2023; Ustinova et al., 2022).

Recent empirical work has underscored the salience of psychosocial factors—including emotional well-being, resilience, social support, coping styles, and academic adjustment—in shaping educational persistence. Resilience, in particular, has emerged as a protective construct moderating the effects of adversity on developmental outcomes. Studies examining resilience under conditions of war-related stress and socio-political instability demonstrate that perceived social support and adaptive coping significantly buffer psychological distress (Яновська & Perelygina, 2025). Similarly, longitudinal investigations of resilience markers in youth exposed to chronic medical adversity highlight the predictive value of emotional regulation and adaptive engagement in sustaining functional outcomes over time (Sirois et al., 2022).

Within educational settings, psychosocial adjustment has been closely linked to academic persistence. Exposure to post-traumatic stress symptoms predicts difficulties in academic adjustment and school engagement, emphasizing the need to integrate mental health variables into predictive educational models (Muhammad & Abdullahi, 2025). Moreover, patterns of resilient functioning in early life show that distinct psychosocial profiles correspond to differential academic and emotional outcomes (Cahill et al., 2023). These findings collectively indicate that dropout risk must be interpreted within broader psychological and social ecologies rather than narrowly defined academic metrics.

Cognitive functioning represents another critical pillar in understanding dropout vulnerability. Neurocognitive outcomes are deeply intertwined with emotional regulation, environmental stress, and learning processes. Developmental neuro-oncology research demonstrates that emotional and cognitive outcomes are reciprocally influenced across developmental stages, reinforcing the necessity of multidimensional assessment in youth populations (Christou, 2025). Research exploring psychiatric traits and dynamic functional connectivity further reveals that cognitive deficits and emotional dysregulation share overlapping neural substrates (Yan et al., 2023). Such findings support the proposition that dropout risk models must incorporate both psychosocial and neurocognitive dimensions to achieve meaningful explanatory power.

Executive function and self-regulatory capacity are especially central to academic persistence. Early childhood research linking smartphone addiction to executive dysfunction and emotional dysregulation illustrates how behavioral regulation challenges can disrupt adaptive learning pathways (Warmansyah et al., 2023). Similarly, parenting practices and emotional intelligence have been shown to correlate with school readiness and cognitive preparedness (Mashar & Astuti, 2022). The interplay between executive control and psychosocial context is further supported by investigations demonstrating that trait resilience interacts with neurocognitive functioning to influence vulnerability in youth at risk for mental disorders (Mętel et al., 2020).

Beyond individual traits, social determinants of health substantially influence cognitive and academic trajectories. Neighborhood opportunity structures and socioeconomic context are strongly associated with neurocognitive and psychological outcomes in pediatric populations (Nielsen et al., 2023). Long-term follow-up studies among lower-educated adults reveal that risk and protective factors for cognitive decline are heavily shaped by contextual resources and social engagement (Ribeiro & Leist, 2023). Emotional and instrumental support also function as protective moderators in cognitive aging, illustrating the broader principle that psychosocial resources influence cognitive functioning across the lifespan (Morris et al., 2023). These insights reinforce the importance of integrating contextual predictors into dropout modeling frameworks.

Advances in artificial intelligence (AI) and machine learning (ML) offer unprecedented opportunities to model complex, nonlinear interactions among multidimensional predictors. AI-based modeling of mood, coping, work engagement, and social factors has demonstrated robust predictive capacity in mental health outcomes, highlighting

JAYPS
Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526

the suitability of machine learning approaches for psychosocial forecasting (Kundu, 2025). Deep learning models identifying transdiagnostic biotypes in youth with ADHD and anxiety disorders further illustrate how computational frameworks can uncover latent dimensions underlying behavioral outcomes (Jiao, 2025). Similarly, EEG-based paradigms have shown promise in discriminating psychiatric conditions using neurophysiological signals, underscoring the growing convergence between computational neuroscience and applied prediction models (Yang et al., 2025).

Despite these methodological advances, concerns regarding algorithmic opacity and interpretability persist. The integration of Theory of Mind (ToM) and metacognitive dimensions into AI systems has been proposed as a pathway toward more transparent and human-aligned models (Bamicha & Drigas, 2024). The need for explainability is especially critical in educational contexts, where predictive systems must inform intervention rather than merely classify risk. Ethical implementation requires interpretable algorithms capable of identifying modifiable psychosocial and cognitive drivers of vulnerability.

Recent interdisciplinary workplace health initiatives demonstrate how complex modeling frameworks can inform preventive strategies by integrating psychosocial, biological, and environmental data (Moortel et al., 2025). Analogously, mentoring interventions grounded in social justice frameworks highlight the transformative potential of targeted psychosocial support in high-risk populations (Miner-Romanoff & Greenawalt, 2024). These applied models emphasize that predictive analytics must ultimately translate into actionable, context-sensitive interventions.

Furthermore, biological and lifestyle factors intersect with cognitive and psychosocial functioning. Multilingualism has been associated with cognitive benefits that may enhance academic resilience (Daud, 2024). Household physical activity is positively linked to neural structural integrity, illustrating lifestyle-cognition interdependencies (Koblinsky et al., 2021). Conversely, depression and hormonal imbalances have been associated with cognitive impairment, underscoring the relevance of mental health assessment in predictive educational models (Vornyk, 2021).

The COVID-19 pandemic has further amplified the urgency of integrated predictive frameworks. Longitudinal studies reveal that pandemic-related stress altered health behaviors, psychosocial functioning, and cognitive performance across populations (Hausman et al., 2022; Slade et al., 2022). Interventions combining meditation and aerobic exercise improved mental health outcomes among educators during pandemic conditions, suggesting that psychosocial modulation can influence cognitive and emotional trajectories (Demmin et al., 2022). Such findings reinforce the dynamic and modifiable nature of psychosocial-cognitive systems relevant to dropout prevention.

Importantly, AI-driven modeling approaches must be grounded in validated measurement and biomarker research to ensure methodological rigor. Protocol-based natural history studies establishing biomarkers and clinical endpoints illustrate how structured, longitudinal data collection enhances predictive validity (Mul, 2025). Systematic analyses integrating biological, sociodemographic, and psychosocial contributors across life stages provide a blueprint for multidimensional risk modeling (Gajewski et al., 2022, 2023).

In educational neuroscience, early music education has been shown to enhance cognitive, emotional, and social development, further supporting the integration of neurocognitive enrichment variables into academic persistence models (Sanchez, 2025). Biofeedback and game-based learning approaches in special education similarly illustrate how technology can be leveraged to improve cognitive regulation and emotional engagement (Karageorgopoulos et al., 2025). Meanwhile, research exploring the "Will to Exist, Live and Survive/Fight" construct demonstrates how motivational resilience predicts adaptive functioning across cultures (Kira et al., 2023).

Collectively, the literature converges on three critical insights: first, dropout risk emerges from complex interactions between psychosocial and cognitive variables; second, contextual and social determinants significantly moderate these interactions; and third, advanced machine learning methods are uniquely positioned to model such complexity, provided that interpretability and ethical transparency are prioritized. Although prior studies have independently examined psychosocial predictors (Cablaida & Delfino, 2025; Lupini et al., 2023), neurocognitive outcomes (Rehan & Phillips, 2023), and environmental risk parameters (Galitskaya et al., 2024), there remains a critical gap in integrative, explainable machine learning models that simultaneously incorporate these domains within educational dropout prediction.

Therefore, the present study aims to develop and validate an explainable machine learning model for predicting school dropout risk using integrated psychosocial and cognitive

JAYPS
Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526

variables, while ensuring algorithmic transparency and interpretability suitable for preventive educational intervention.

2. Methods and Materials

2.1. Study Design and Participants

This study employed a cross-sectional predictive design integrating psychometric assessment with explainable machine learning modeling to identify and interpret dropout risk among secondary school students in Germany. The target population consisted of students enrolled in public lower and upper secondary schools across three federal states: North Rhine–Westphalia, Bavaria, and Berlin. Using a stratified cluster sampling strategy to ensure representation across urban and semi-urban districts, 1,248 students were initially invited to participate. After excluding incomplete responses and cases with excessive missing data exceeding 10% per individual, the final analytic sample comprised 1,172 students. Participants ranged in age from 14 to 18 years, with a mean age of 16.1 years (SD = 1.12). The sample included 594 female students (50.7%) and 578 male students (49.3%). Approximately 27.4% of participants had a migration background as defined by at least one parent born outside Germany. Dropout risk status was operationalized using a composite indicator combining administrative records of chronic absenteeism (greater than 20% unexcused absence in the previous semester), grade retention history, and official at-risk classification by school counseling services.

2.2. Measures

Data collection was conducted during the spring academic term using standardized, validated German-language instruments administered in classroom settings under supervised conditions. Psychosocial variables included academic motivation, assessed using the German adaptation of the Academic Self-Regulation Questionnaire; school belonging, measured through the Psychological Sense of School Membership scale; perceived teacher support and peer support, assessed via subscales of the Classroom Life Measure; and academic self-efficacy, evaluated using the General Academic Self-Efficacy Scale. Emotional functioning variables included depressive symptoms measured by the Patient Health Questionnaire–Adolescent version and test anxiety assessed with the German Test Anxiety Inventory. Behavioral engagement and self-regulation were measured using the Self-Regulation Questionnaire–Academic domain. Cognitive variables included working memory capacity assessed via a computerized n-back task, processing speed measured through a digit-symbol substitution task, and fluid reasoning evaluated using a short form of Raven's Advanced Progressive Matrices. Socio-demographic covariates included parental education level, socioeconomic status indexed by the Family Affluence Scale, and migration background. All psychometric instruments demonstrated acceptable internal consistency in the present sample, with Cronbach's alpha coefficients ranging from .78 to .91. Cognitive tasks were administered using standardized computerized procedures in school computer laboratories to ensure uniform testing conditions.

2.3. Data Analysis

Data analysis followed a multi-stage machine learning pipeline designed to maximize predictive performance while preserving interpretability. Preprocessing included outlier screening, multiple imputation for item-level missing values using predictive mean matching, z-standardization of continuous predictors, and one-hot encoding of categorical variables. The dataset was randomly partitioned into training (70%), validation (15%), and test (15%) subsets using stratified sampling to preserve the proportion of at-risk students across splits. To address class imbalance, given that 18.6% of the final sample met criteria for high dropout risk, Synthetic Minority Over-sampling Technique (SMOTE) was applied exclusively within the training set.

Several supervised classification algorithms were implemented, including logistic regression with elastic net regularization, random forest, gradient boosting (XGBoost), and support vector machines with radial basis kernels. Hyperparameter optimization was conducted using five-fold cross-validated grid search within the training set, with model selection based on area under the receiver operating characteristic curve (AUC), F1-score, and balanced accuracy. Final model performance was evaluated on the held-out test set to prevent overfitting.

To enhance model transparency and facilitate psychological interpretability, explainability techniques were systematically applied. Global feature importance was examined using permutation importance and mean decrease in impurity for tree-based models. Local interpretable model-agnostic explanations (LIME) were generated to analyze individual-level prediction patterns. Additionally,

E-ISSN: 2981-2526

SHapley Additive exPlanations (SHAP) values were computed to quantify the marginal contribution of each psychosocial and cognitive predictor to model output, both globally and for high-risk individuals. Partial dependence plots were generated to visualize nonlinear relationships between key predictors and dropout probability.

Model calibration was assessed using calibration curves and Brier scores to ensure probabilistic reliability. To evaluate robustness, sensitivity analyses were conducted by re-estimating models without cognitive variables and separately without psychosocial variables, allowing for comparative assessment of incremental predictive validity. All analyses were performed using Python (scikit-learn, XGBoost, SHAP libraries) in a reproducible computational environment, with statistical significance interpreted at an alpha level of .05 where applicable for baseline regression comparisons.

3. Findings and Results

Descriptive statistics for all psychosocial, cognitive, and demographic variables included in the predictive modeling are presented in Table 1. The table provides means, standard deviations, and observed ranges for continuous variables, as well as frequencies and percentages for categorical indicators. These descriptive results offer an overview of the distributional properties of the predictors and allow preliminary comparison between students classified as high dropout risk and those classified as low risk prior to machine learning modeling.

Table 1

Descriptive Statistics of Psychosocial, Cognitive, and Demographic Variables (N = 1.172)

Variable	Total Sample Mean (SD) / n (%)	Low Risk (n = 954) Mean (SD) / n (%)	High Risk (n = 218) Mean (SD) / n (%)
Academic Motivation	3.62 (0.74)	3.78 (0.69)	2.94 (0.81)
School Belonging	3.71 (0.68)	3.85 (0.61)	3.05 (0.73)
Academic Self-Efficacy	3.55 (0.72)	3.69 (0.66)	2.98 (0.78)
Teacher Support	3.66 (0.70)	3.79 (0.63)	3.10 (0.77)
Peer Support	3.81 (0.65)	3.89 (0.60)	3.45 (0.76)
Depressive Symptoms	9.42 (5.31)	8.21 (4.76)	14.63 (5.92)
Test Anxiety	2.97 (0.81)	2.84 (0.75)	3.52 (0.86)
Self-Regulation	3.49 (0.69)	3.62 (0.63)	2.92 (0.74)
Working Memory (n-back accuracy %)	73.84 (9.22)	75.12 (8.61)	68.17 (10.33)
Processing Speed (standard score)	101.45 (12.34)	103.18 (11.76)	93.91 (12.85)
Fluid Reasoning (Raven score)	24.63 (4.88)	25.32 (4.51)	21.45 (5.11)
Female	594 (50.7%)	472 (49.5%)	122 (56.0%)
Migration Background	321 (27.4%)	231 (24.2%)	90 (41.3%)
Low SES	284 (24.2%)	188 (19.7%)	96 (44.0%)

As shown in Table 1, students classified as high dropout risk exhibited consistently lower scores across core psychosocial variables, including academic motivation, school belonging, academic self-efficacy, teacher support, peer support, and self-regulation. The largest mean differences were observed for depressive symptoms and self-efficacy, suggesting substantial emotional and motivational disparities between groups. Cognitive indicators also showed systematic differences, with high-risk students demonstrating lower working memory accuracy, slower processing speed, and reduced fluid reasoning scores. Socio-demographic disparities were also notable: migration background and low socioeconomic status were substantially overrepresented among high-risk students. These descriptive findings indicate multidimensional vulnerability profiles spanning emotional, motivational, cognitive, and contextual domains.

JAYPS Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526

Table 2

Predictive Performance of Machine Learning Models on Test Set

Model	AUC	Balanced Accuracy	Precision	Recall	F1-Score	Brier Score
Logistic Regression (Elastic Net)	0.83	0.76	0.69	0.71	0.70	0.146
Support Vector Machine (RBF)	0.85	0.78	0.72	0.73	0.72	0.139
Random Forest	0.89	0.83	0.78	0.81	0.79	0.121
Gradient Boosting (XGBoost)	0.92	0.86	0.82	0.85	0.83	0.108

Table 2 demonstrates that ensemble tree-based models outperformed linear and kernel-based models. Gradient boosting achieved the highest discriminative capacity, with an AUC of 0.92 and balanced accuracy of 0.86, indicating strong capacity to differentiate between high-risk and low-risk students. The random forest model also showed robust performance, with only slightly lower metrics. Logistic regression, while more interpretable, demonstrated comparatively reduced predictive power. Calibration analysis revealed that gradient boosting also achieved the lowest Brier score, indicating superior probabilistic reliability.

Table 3

Global Feature Importance Based on SHAP Values (Gradient Boosting Model)

Predictor	Mean Absolute SHAP Value	Relative Importance (%)
Depressive Symptoms	0.184	17.6
Academic Self-Efficacy	0.169	16.2
School Belonging	0.141	13.5
Processing Speed	0.116	11.1
Self-Regulation	0.109	10.4
Academic Motivation	0.098	9.4
Migration Background	0.072	6.9
Low SES	0.067	6.4
Working Memory	0.054	5.2
Teacher Support	0.041	3.9
Fluid Reasoning	0.038	3.6
Peer Support	0.021	2.0

Table 3 indicates that depressive symptoms emerged as the most influential predictor of dropout risk, followed closely by academic self-efficacy and school belonging. Notably, processing speed ranked among the top five predictors, suggesting that cognitive processing efficiency plays a nontrivial role in dropout vulnerability. Socioeconomic status and migration background also contributed meaningfully, though to a lesser extent than core psychosocial factors. Peer support demonstrated comparatively lower global influence, although local explanations revealed heterogeneity in its contribution across individuals.

To further clarify the incremental predictive value of psychosocial versus cognitive variables, a comparative model analysis is presented in Table 4.

Table 4

Comparative Model Performance by Predictor Domain (Gradient Boosting)

Predictor Set	AUC	Balanced Accuracy	F1-Score
Psychosocial Variables Only	0.89	0.83	0.80
Cognitive Variables Only	0.77	0.71	0.68
Combined Psychosocial + Cognitive	0.92	0.86	0.83

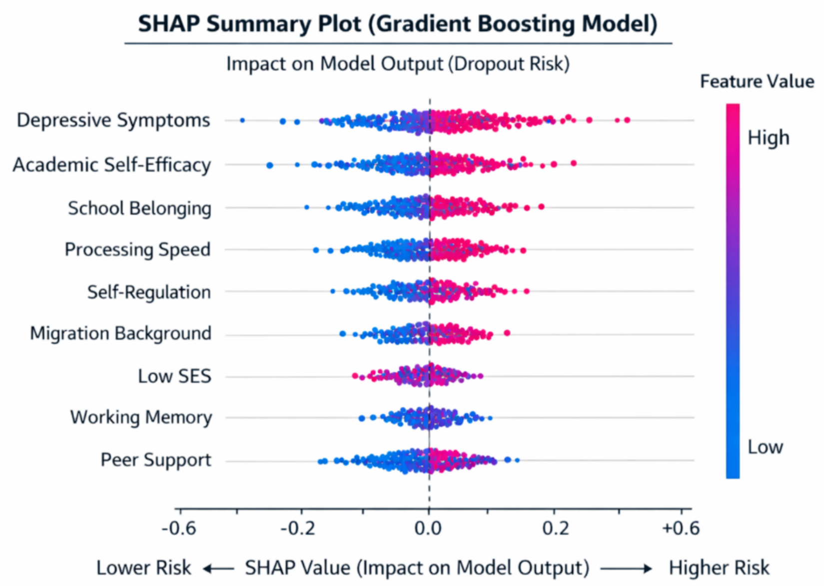
As displayed in Table 4, psychosocial variables alone achieved strong predictive accuracy, outperforming models based solely on cognitive indicators. However, combining both domains yielded the highest overall performance,

JAYPS Adolescent and Youth Psychological Studies E-ISSN: 2981-2526

indicating incremental predictive validity of cognitive measures beyond psychosocial risk factors. This finding supports the multidimensional nature of dropout vulnerability and underscores the added value of integrating cognitive assessments into early warning systems.

Figure 1

SHAP Summary Plot of Gradient Boosting Model Showing Direction and Magnitude of Predictor Effects on Dropout Risk Probability



The SHAP summary visualization revealed that higher depressive symptoms, lower academic self-efficacy, reduced school belonging, slower processing speed, and diminished self-regulation were associated with increased predicted dropout probability. The plot also demonstrated nonlinear threshold effects, particularly for depressive symptoms and processing speed, where risk probability increased sharply beyond moderate symptom severity and below-average cognitive performance. Local explanation analysis further showed individualized risk constellations, indicating that dropout vulnerability may arise from distinct combinations of psychosocial and cognitive factors rather than a uniform risk pathway.

4. Discussion

The present study sought to develop and validate an explainable machine learning model for predicting school dropout risk using integrated psychosocial and cognitive variables. The findings demonstrated that ensemble-based algorithms, particularly gradient boosting, achieved high predictive accuracy while maintaining interpretability through SHAP-based feature attribution. Importantly, psychosocial variables—especially depressive symptoms, academic self-efficacy, and school belonging—emerged as the most influential predictors, while cognitive indicators such as processing speed and working memory contributed incremental predictive validity. These results reinforce the conceptualization of dropout as a multidimensional phenomenon shaped by interacting emotional, cognitive, and contextual determinants.

The prominence of depressive symptoms in the model aligns with extensive literature linking emotional distress to academic disengagement and maladaptive adjustment. Youth exposed to post-traumatic stress and insecurity show marked challenges in academic adjustment, indicating that emotional dysregulation undermines persistence in educational settings (Muhammad & Abdullahi, 2025). Similarly, investigations of resilience during large-scale crises demonstrate that emotional coping capacity and social support significantly buffer stress-related impairments (Яновська & Perelygina, 2025). The current findings extend these insights by demonstrating that depressive symptom severity exerts not only theoretical relevance but also quantifiable predictive impact within computational models of dropout risk.

Academic self-efficacy and school belonging were also central predictors. These findings resonate with research showing that psychosocial resources—particularly resilience and perceived support—are strongly associated with adaptive functioning across diverse populations (Cahill et al., 2023; Sirois et al., 2022). Emotional and instrumental support have been identified as protective factors for cognitive and psychological outcomes in vulnerable groups (Morris et al., 2023). Within educational environments, such support structures may function as stabilizing forces that mitigate dropout trajectories. The model's identification of self-efficacy as a key driver further supports motivational frameworks emphasizing perceived competence as a determinant of sustained academic engagement.

Cognitive variables demonstrated meaningful, though comparatively smaller, contributions. Processing speed and working memory independently improved predictive performance beyond psychosocial measures. These findings align with neurocognitive research indicating that executive and attentional processes are foundational to adaptive functioning (Yan et al., 2023). Studies examining executive dysfunction in relation to emotional dysregulation further illustrate that deficits in cognitive control exacerbate vulnerability in youth populations (Warmansyah et al., 2023). Moreover, research exploring the interplay between trait resilience and neurocognitive functioning in at-risk mental states demonstrates that cognitive resources modulate psychological vulnerability (Mętel et al., 2020). The present findings support the proposition that dropout risk involves both affective distress and diminished cognitive efficiency.

Socioeconomic and contextual factors, including low SES and migration background, also contributed to predictive outcomes. These results are consistent with evidence demonstrating that neighborhood opportunity structures and social determinants of health significantly shape neurocognitive and psychological development (Nielsen et al., 2023). Longitudinal investigations among lower-educated populations reveal cumulative risk mechanisms linking socioeconomic disadvantage to cognitive decline and reduced functional outcomes (Ribeiro & Leist, 2023). The integration of contextual predictors within machine learning models therefore reflects ecological validity and aligns with multidimensional risk frameworks.

The comparative analysis further demonstrated that psychosocial variables alone yielded strong predictive capacity, yet the inclusion of cognitive indicators significantly improved overall model performance. This finding mirrors interdisciplinary modeling approaches in mental health research, where AI systems integrating mood, coping, and social variables outperform single-domain models (Kundu, 2025). Deep learning investigations identifying transdiagnostic youth biotypes similarly highlight the importance of multidimensional data integration (Jiao, 2025). In educational contexts, such integrative modeling approaches are essential for capturing the nonlinear, interactive pathways leading to dropout.

The use of explainable AI techniques represents a major methodological contribution. By employing SHAP-based explanations, the model provided transparent quantification of individual predictor contributions. This approach aligns with calls for human-centered AI systems incorporating metacognitive and Theory of Mind dimensions to enhance interpretability and ethical alignment (Bamicha & Drigas, 2024). In sensitive domains such as education, interpretability is not optional but essential, as predictive outputs must guide interventions rather than merely categorize risk. The current findings demonstrate that high predictive accuracy and algorithmic transparency are not mutually exclusive.

The strong predictive influence of depressive symptoms also aligns with research linking mental health disruptions to altered cognitive functioning during periods of crisis, including the COVID-19 pandemic (Hausman et al., 2022; Slade et al., 2022). Interventions combining mindfulness and aerobic exercise have been shown to improve mental health and cognitive outcomes in educational professionals, suggesting modifiable pathways linking emotional well-being and functional performance (Demmin et al., 2022). Similarly, structured emotional support interventions in multicultural educational contexts have demonstrated preventive value against psychosocial maladjustment (Liu, 2025). These findings suggest that early identification of emotional distress via predictive modeling could facilitate timely psychosocial interventions.

The role of cognitive enrichment variables further supports preventive approaches. Research on early music education demonstrates that structured cognitive engagement enhances emotional and social development (Sanchez, 2025). Biofeedback and game-based learning in special education contexts also highlight how cognitive regulation can be strengthened through targeted interventions (Karageorgopoulos et al., 2025). These findings suggest that cognitive deficits identified in predictive models may represent modifiable risk factors rather than fixed vulnerabilities.

Broader evidence indicates that lifestyle and biological variables intersect with psychosocial functioning. Multilingualism has been associated with cognitive benefits

JAYPS Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526

that may indirectly enhance academic resilience (Daud, 2024). Physical activity correlates with neural integrity and cognitive performance, underscoring the systemic nature of protective factors (Koblinsky et al., 2021). Conversely, depressive states and hormonal dysregulation have been linked to cognitive impairment, reinforcing the bidirectional interplay between emotional health and cognitive functioning (Vornyk, 2021). The present findings are consistent with this biopsychosocial framework.

The results also align with systematic analyses demonstrating that biological, sociodemographic, and psychosocial factors jointly influence functional outcomes across the lifespan (Gajewski et al., 2022, 2023). Similarly, longitudinal biomarker studies emphasize the importance of structured, multidimensional measurement protocols for enhancing predictive validity (Mul, 2025). By integrating psychosocial and cognitive indicators within a unified modeling framework, the present study contributes to this evolving paradigm.

Mentoring and social justice-oriented interventions have demonstrated significant impact on resilience and educational persistence in high-risk populations (Miner-Romanoff & Greenawalt, 2024). Patterns of resilient functioning indicate that adaptive psychosocial profiles can offset early adversity (Cahill et al., 2023). Moreover, motivational constructs such as the "Will to Exist, Live and Survive/Fight" have shown cross-cultural predictive validity for adaptive outcomes (Kira et al., 2023). These insights underscore that dropout risk is not deterministic but modifiable through targeted psychosocial reinforcement.

Furthermore, studies exploring behavioral health outcomes in chronic pediatric conditions demonstrate how social determinants shape cognitive and emotional trajectories (Lupini et al., 2023). Research on cognitive predictors in addiction services also reveals that executive functioning deficits are associated with functional impairment (Gooden et al., 2022). These parallels highlight that dropout vulnerability shares structural similarities with broader mental health risk models.

Finally, the identification of nonlinear risk thresholds within SHAP analyses supports emerging perspectives on dynamic risk modeling. Computational neuroscience research utilizing EEG paradigms to distinguish psychiatric states demonstrates that complex, nonlinear signatures underpin behavioral outcomes (Yang et al., 2025). The current findings extend this logic to educational prediction, illustrating that dropout risk arises from interactive, nonlinear patterns across domains.

5. Conclusion

In sum, the present study confirms that explainable machine learning models integrating psychosocial and cognitive variables can achieve high predictive accuracy while preserving interpretability. The findings substantiate multidimensional risk frameworks and provide empirical support for integrative, transparent predictive analytics in educational contexts.

6. Limitations & Suggestions

Despite its methodological strengths, several limitations warrant consideration. First, the cross-sectional design restricts causal inference and limits conclusions regarding developmental trajectories leading to dropout. Longitudinal data would provide stronger evidence for temporal ordering among psychosocial, cognitive, and contextual predictors. Second, although the sample was sizeable and demographically diverse, generalizability may be constrained to similar educational contexts. Third, while explainability techniques were employed, model outputs remain probabilistic and may not capture unmeasured contextual nuances such as family dynamics or teacher–student relationships. Finally, reliance on self-report measures for psychosocial variables introduces potential response biases that could influence predictive estimates.

Future research should prioritize longitudinal modeling frameworks to examine dynamic risk trajectories and causal mechanisms. Incorporating multimodal data sources—including neurophysiological measures, ecological momentary assessments, and digital behavioral indicators—could enhance predictive granularity. Comparative studies across cultural and educational systems would clarify contextual moderators of model performance. Additionally, intervention-based validation studies are needed to test whether early identification through explainable machine learning leads to measurable reductions in dropout rates. The integration of reinforcement learning approaches may further enable adaptive, individualized risk monitoring systems.

From a practical standpoint, the findings underscore the necessity of embedding mental health screening and psychosocial support within school-based early warning systems. Educational institutions should consider integrating interpretable AI tools that identify modifiable risk factors rather than solely flagging at-risk students. Targeted interventions aimed at strengthening academic self-efficacy, emotional regulation, and executive functioning may yield meaningful reductions in dropout vulnerability. Collaboration among educators, school psychologists, and data scientists is essential to ensure ethical deployment, transparency, and continuous monitoring of predictive systems. Ultimately, predictive analytics should function not as labeling mechanisms but as supportive instruments guiding preventive, student-centered educational practice.

Acknowledgments

We would like to express our appreciation and gratitude to all those who cooperated in carrying out this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

The study protocol adhered to the principles outlined in the Helsinki Declaration, which provides guidelines for ethical research involving human participants.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

Authors' Contributions

All authors equally contributed to this article.

References

Bamicha, V., & Drigas, A. (2024). Strengthening AI via ToM and MC Dimensions. *Scientific Electronic Archives*, *17*(3). https://doi.org/10.36560/17320241939

Cablaida, C. L., & Delfino, A. A. (2025). Building Resilience: Its Implication to SARDO's Learning Recovery. *International Journal of Social Science Humanity & Management Research*, *04*(06). https://doi.org/10.58806/ijsshmr.2025.v4i6n24

Cahill, S., Hager, R., & Shryane, N. (2023). Patterns of Resilient Functioning in Early Life: Identifying Distinct Groups and Associated Factors. *Development and Psychopathology*, *36*(4), 1789-1809. https://doi.org/10.1017/s0954579423001165

Christou, A. I. (2025). Neurocognitive and Emotional Outcomes in Childhood Cancer: A Developmental Perspective. *Current Oncology*, *32*(11), 611. https://doi.org/10.3390/curroncol32110611

Daud, R. (2024). The Cognitive Benefits of Speaking Multiple Languages. *European Journal of Linguistics*, *3*(4), 1-16. https://doi.org/10.47941/ejl.2055

Demmin, D., Silverstein, S. M., & Shors, T. J. (2022). Mental and Physical Training With Meditation and Aerobic Exercise Improved Mental Health and Well-Being in Teachers During the COVID-19 Pandemic. *Frontiers in human neuroscience*, *16*. https://doi.org/10.3389/fnhum.2022.847301

Gajewski, P. D., Rieker, J. A., Athanassiou, G., Bröde, P., Claus, M., Golka, K., Hengstler, J. G., Kleinsorge, T., Nitsche, M. A., Reinders, J., Tisch, A., Watzl, C., Wascher, E., & Getzmann, S. (2022). A Systematic Analysis of Biological, Sociodemographic, Psychosocial, and Lifestyle Factors Contributing to Work Ability Across the Working Life Span: Cross-Sectional Study (Preprint). https://doi.org/10.2196/preprints.40818

Gajewski, P. D., Rieker, J. A., Athanassiou, G., Bröde, P., Claus, M., Golka, K., Hengstler, J. G., Kleinsorge, T., Nitsche, M. A., Reinders, J., Tisch, A., Watzl, C., Wascher, E., & Getzmann, S. (2023). A Systematic Analysis of Biological, Sociodemographic, Psychosocial, and Lifestyle Factors Contributing to Work Ability Across the Working Life Span: Cross-Sectional Study. *Jmir Formative Research*, *7*, e40818. https://doi.org/10.2196/40818

Galitskaya, V., Drigas, A., & Αντωνίου, Α.-Σ. (2024). Understanding the Challenges of Learning Disabilities and Psychosocial Disorders. *Scientific Electronic Archives*, *17*(6). https://doi.org/10.36560/17620242008

Gooden, J. R., Cox, C. A., Petersen, V., Curtis, A., Sanfilippo, P. G., Manning, V., Bolt, G., & Lubman, D. I. (2022). Predictors of Cognitive Functioning in Presentations to a Community-Based Specialist Addiction Neuropsychology Service. *Brain Impairment*, *24*(1), 54-68. https://doi.org/10.1017/brimp.2021.38

Hausman, H. K., Dai, Y., O'Shea, A., Dominguez, V., Fillingim, M., Calfee, K., Carballo, D., Hernandez, C., Perryman, S., Kraft, J. N., Evangelista, N. D., Etten, E. J. V., Smith, S. G., Bharadwaj, P. K., Song, H., Porges, E. C., DeKosky, S. T., Hishaw, G. A., Marsiske, M., . . . Woods, A. J. (2022). The Longitudinal Impact of the COVID-19 Pandemic on Health Behaviors, Psychosocial Factors, and Cognitive Functioning in Older Adults. *Frontiers in Aging Neuroscience*, *14*. https://doi.org/10.3389/fnagi.2022.999107

Jiao, Y. (2025). Deep Learning of Brain-Behavior Dimensions Identifies Transdiagnostic Biotypes in Youth With ADHD and Anxiety Disorders. https://doi.org/10.1101/2025.10.13.682243

Karageorgopoulos, I., Ανδρικόπουλος, A., Georgiou, E. S., & Koutsojannis, C. (2025). The Impact and Integration of Biofeedback and Game-Based Learning in Special Education: A Global Perspective on Early Childhood Intervention. https://doi.org/10.20944/preprints202507.1879.v1

Kira, I. A., Özcan, N. A., Shuwiekh, H., & Alhuwailah, A. (2023). The "Will to Exist, Live and Survive/Fight" (WTELS-F) Scale Initial Short Version: Cross-Cultural Validation of Its Reliability, Structural, Predictive and Incremental Validity. *Psychology*, *14*(03), 470-497. https://doi.org/10.4236/psych.2023.143026

Koblinsky, N., Meusel, L.-A., Greenwood, C. E., & Anderson, N. D. (2021). Household Physical Activity Is Positively Associated With Gray Matter Volume in Older Adults. *BMC*

Geriatrics, *21*(1). https://doi.org/10.1186/s12877-021-02054-8

Kundu, P. (2025). Artificial Intelligence Modeling of Mood, Coping, Work Engagement and Social Factors in Predicting Mental Health Outcomes. *Bio Web of Conferences*, *204*, 01013. https://doi.org/10.1051/bioconf/202520401013

Liu, Y. (2025). Directions for Emotional Support in Multicultural Education From a School Violence Prevention Perspective: An Integrative Approach Combining Mindfulness and Counseling. *Humanities Meditation Res Inst*, *3*(2), 41-53. https://doi.org/10.60140/mche.2025.3.2.41

Lupini, F., Rubinstein, T. B., Mackey, E. R., & Sule, S. (2023). Behavioral Health Outcomes and Social Determinants of Health in Children With Diabetes and Juvenile Arthritis. https://doi.org/10.21203/rs.3.rs-3610878/v1

Mashar, R., & Astuti, F. P. (2022). Correlation Between Parenting Skills, Children's Emotional and Intelligence Quotient With School Readiness. *Jpud - Jurnal Pendidikan Usia Dini*, *16*(2), 215-223. https://doi.org/10.21009/jpud.162.02

Mętel, D., Arciszewska, A., Frydecka, D., Cechnicki, A., & Gawęda, Ł. (2020). M65. Interplay of Trait Resilience With Neurocognitive Functioning and Personality Characteristics in Young People With Putative 'At Risk Mental State'. *Schizophrenia Bulletin*, *46*(Supplement_1), S159-S160. https://doi.org/10.1093/schbul/sbaa030.377

Miner-Romanoff, K., & Greenawalt, J. (2024). Evaluation of the Arthur Project: Evidence-Based Mentoring in a Social Work Framework With a Social Justice Approach. *Societies*, *14*(7), 123. https://doi.org/10.3390/soc14070123

Moortel, D. D., Turner, M. C., Arensman, E., Alex Binh Vinh Duc, N., & González, V. M. (2025). Improving Health-Promoting Workplaces Through Interdisciplinary Approaches. The Example of WISEWORK-C, a Cluster of Five Work and Health Projects Within Horizon-Europe. *Scandinavian Journal of Work Environment & Health*, *51*(4), 259-264. https://doi.org/10.5271/sjweh.4238

Morris, E. P., Kraal, A. Z., Levy, S. A., Arias, F., Chen, R., Šeblová, D., Jimenez, M. P., Farina, M., Zlatar, Z. Z., Chanti-Ketterl, M., Lor, Y., Fletcher, E., Manly, J. J., & Glymour, M. M. (2023). 17 Emotional and Instrumental Support as Protective Factors in Cognitive Aging Among Black and Hispanic/Latinx Older Adults. *Journal of the International Neuropsychological Society*, *29*(s1), 330-331. https://doi.org/10.1017/s135561772300454x

Muhammad, S., & Abdullahi, T. (2025). Post-Traumatic Stress and Academic Adjustment Among Children Affected by Insecurity in Northern Nigeria: Strategies for Psychosocial Intervention in Schools. *Federal University Gusau faculty of Education Journal*, *5*(4), 81-88. https://doi.org/10.64348/zije.202592

Mul, K. (2025). Establishing Biomarkers and Clinical Endpoints in Myotonic Dystrophy Type 1 (END-DM1): Protocol of an International Natural History Study. *PLoS One*, *20*(12), e0331163. https://doi.org/10.1371/journal.pone.0331163

Nielsen, J., Sharkey, C., Hardy, K. K., & Walsh, K. S. (2023). 11 Social Determinants of Health in Pediatric Brain Tumor Survivors: Associations Between Neighborhood Opportunity and Neurocognitive and Psychological Outcomes. *Journal of the International Neuropsychological Society*, *29*(s1), 13-14. https://doi.org/10.1017/s1355617723000942

Rehan, S., & Phillips, N. A. (2023). Social Engagement and Cognitive Decline: Exploring Relationships Between Psychosocial Functioning and Cognitive Performance in Individuals With Mild Cognitive Impairment. *Alzheimer S & Dementia*, *19*(S19). https://doi.org/10.1002/alz.078068

Ribeiro, F. S., & Leist, A. (2023). Risk and Protective Factors for Cognitive Decline in Lower Educated Older Adults With 15-year Follow-up. *Alzheimer S & Dementia*, *19*(S8). https://doi.org/10.1002/alz.067868

Sanchez, M. (2025). Neuroscientific Foundations of Early Music Education: Enhancing Cognitive, Emotional, and Social Development in Primary Schools. *International Journal of Scientific Research and Management*, *13*(10), 4383-4404. https://doi.org/10.18535/ijsrm/v13i10.el03

Sirois, P. A., Huo, Y., Nozyce, M. L., Garvie, P. A., Harris, L., Malee, K., McEvoy, R. E., Mellins, C. A., Nichols, S., Smith, R., & Tassiopoulos, K. (2022). Ageing With HIV: A Longitudinal Study of Markers of Resilience in Young Adults With Perinatal Exposure to HIV, With or Without Perinatally Acquired HIV. *Journal of the International AIDS Society*, *25*(S4). https://doi.org/10.1002/jia2.25982

Slade, K., Davies, R., Pennington, C. R., Plack, C. J., & Nuttall, H. E. (2022). The Impact of Age and Psychosocial Factors on Cognitive and Auditory Outcomes During the COVID-19 Pandemic. https://doi.org/10.31234/osf.io/rqfjy

Ustinova, O. Y., Zaitseva, N. V., & Eisfeld, D. A. (2022). Substantiating Optimal Parameters of Risk Factors Existing in the Educational Environment for Schoolchildren as Per Indicators of Physical, Mental and Somatic Health. *Health Risk Analysis*(2), 48-63. https://doi.org/10.21668/health.risk/2022.2.05

Vornyk, B. (2021). Testosterone, Depression, and Cognitive Impairment in Men: An Attempt at Practical Analysis. *Health of Man*(4), 86-89. https://doi.org/10.30841/2307-5090.4.2021.252406

Warmansyah, J., Ismandela, A., Nabila, D. F., Wulandari, R. T., Wahyu, W. P., Khairunnisa, K., putri, A., Komalasari, E., Sari, M., & Yuningsih, R. (2023). Smartphone Addiction, Executive Function, and Mother-Child Relationships in Early Childhood Emotion Dysregulation. *Jpud - Jurnal Pendidikan Usia Dini*, *17*(2), 241-266. https://doi.org/10.21009/jpud.172.05

Yan, X., Feng, Z., Zhang, H., Zhou, T., Yu, X., & Zhang, Y. (2023). Dynamic Functional Connectivity Changes Associated With Psychiatric Traits and Cognitive Deficits in Cushing's Disease. https://doi.org/10.21203/rs.3.rs-2489663/v1

Yang, C., Pi, Y., Wang, W., Huang, Y., Tang, N., Wang, H., & Wen, S. (2025). Evaluating the Efficacy of Three Classical EEG Paradigms in the Discrimination of Bipolar Depression. *Frontiers in Psychiatry*, *16*. https://doi.org/10.3389/fpsyt.2025.1545132

Яновська, С. Г., & Perelygina, L. (2025). Peculiarities of Resilience and Coping Strategies of the Ukrainian Civilian Population During the Full-Scale War: The Role of Social Support. *Visnyk of v N Karazin Kharkiv National University a Series of Psychology*(78), 46-52. https://doi.org/10.26565/2225-7756-2025-78-07

JAYPS Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526