



# Predicting Graduation Timeliness: A Multidisciplinary Study on Logistic Mobility, Psychosomatic Fatigue, and AI-Assistance Dependency among Information Systems Students

Endin Fahrudin\*, Eko Suharyanto, Santosa Wijayanto

Information Systems, Faculty of Computer Science and Mathematics and Natural Sciences, Pamulang University, South Tangerang, Indonesia

\*Correspondence email: [dosen02729@unpam.ac.id](mailto:dosen02729@unpam.ac.id)

Received: 23/05/2026

Revised: 31/05/2026

Accepted: 15/06/2026

**Abstract:** Timely graduation is now being impacted by non-academic stressors and technology use. This study goes beyond traditional models that rely solely on administrative data by incorporating physical and psychological dimensions, specifically Logistics Mobility (distance and duration of commute), Psychosomatic Fatigue, and AI Assistant Dependence. The data of 292 final year Information Systems students was balanced using Random Over Sampler. Models were compared between Random Forest and Support Vector Machine. Psychosomatic tiredness and AI assistant Dependence was the best predictive factor of delayed graduation, particularly for students commuting more than 25 km, who show more fatigue and higher AI reliance. In terms of algorithmic performance, Random Forest achieved an exceptional accuracy of 100%, while SVM followed closely with 98.28%. The results show the importance of physical exhaustion and digital dependency to identify students at risk of delay. The study finds that a student's environment, in addition to cognitive ability, is a factor in whether or not they graduate on time. High reliance on AI indicates burnout, underscoring the importance of hybrid mentoring and AI literacy programs, particularly for commuting students.

**Keywords:** prediction; mobility; psychosomatic; AI-assistant; SVM



## SDGs Problem Solving

Research stimulates innovation by identifying impediments to student success, such as technological dependence and logistical mobility. Data-driven models for problem-solving enable impactful interventions that align with SDGs Chapter #9 goal of building resilient infrastructure and promoting inclusive and sustainable educational frameworks that cater to changing technical and human needs.

**Conflict of Interest Statement:** The author(s) declares that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

**Copyright:** © ASIIMETRIK. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Jurnal Asimetrik:** Jurnal Ilmiah Rekayasa dan Inovasi - ASIIMETRIK is an open access and peer-reviewed journal published by Faculty of Engineering, Universitas Pancasila, Indonesia.



## 1. INTRODUCTION

Higher education serves as a primary pillar in cultivating competent and globally competitive human resources (Zheng, Cheung and Sit, 2024). The success of an educational institution is frequently measured by its students' performance, with timely graduation being a crucial parameter in determining the quality of instructional processes and academic management (Macqueen et al., 2025). However, the dynamics of modern student life are becoming increasingly complex, encompassing challenges that originate not only from within the classroom but also from external factors that continuously affect their physical and mental well-being (Kondekar and Bade, 2025). In the current era of digital transformation, students have also begun to adopt Artificial Intelligence (AI) as a strategic tool to navigate these pressures (Khosravi et al., 2022). The inability of institutions to map these non-academic factors and technological dependencies often leads to a failure in providing targeted support for students at risk of delayed graduation (Ramadani, Yarta and Pratiwi, 2025).

In the last decade, Educational Data Mining (EDM) has become a major trend in efforts to predict academic success (Sathe, 2026). However, there has been a paradigm shift where graduation determinants are no longer viewed solely from a pure cognitive aspect (Indah et al., 2026). Researchers have begun to realize that the physical environment, daily mobility, and psychophysical well-being of students—intertwined with their reliance on generative technology—play an equally important role as exam scores (Shan, 2026). Therefore, a deeper exploration is required regarding multivariate data that encompasses logistical constraints, biological fatigue, and AI-driven coping mechanisms that possess a systemic impact on student persistence (Noorbehbahani and Oyibo, 2026).

Timely graduation is a crucial parameter in the accreditation assessment of higher education institutions and an indicator of the effectiveness of the learning process. The phenomenon of delayed graduation not only harms students financially and temporally but also impacts the faculty-to-student ratio and the competitiveness of graduates in the professional world. This problem is multifactorial, involving academic, economic, and environmental aspects where students pursue their education. According to Krisnawan et al., student failure to complete studies on time is often rooted in the accumulation of daily stress that goes unidentified by conventional campus monitoring systems (Krisnawan et al., 2025).

This issue is exacerbated by the emergence of new technological dependencies. While students struggle with curriculum demands, the presence of Generative AI tools provides a "digital escape" that can either accelerate or hinder academic progress depending on the intensity of use. As stated by Romiaty et al., reactive approaches that only address students after they face academic trouble have proven ineffective (Romiaty, Sendayu and Putri, 2026). Therefore, problem mapping must include the identification of daily student life patterns, including how they manage exhaustion through technological assistance, before manifestations of academic failure emerge visibly on the surface.

Although the utilization of Machine Learning in the education sector has grown rapidly, there remains a significant research gap in current literature. The majority of research is still stuck in the "administrative data" paradigm, where algorithms are trained solely using variables available in university databases. Based on the literature review, most Educational Data Mining (EDM) studies

---

explore social media use or digital behaviour (Sathe, 2026), but few have explicitly integrated the "physical burden" of daily mobility and its transformation into a cognitive barrier through psychosomatic pathways.

Furthermore, there is a vacuum in the analysis of how technological intervention—specifically Generative AI—interacts with these physical constraints. This gap is the primary focus of this study: bridging technical data modelling with the biological and technological reality experienced by students (Porter *et al.*, 2026). Most research tends to avoid psychophysical and AI-usage variables because they are considered difficult to quantify compared to numerical grades (Papakostas *et al.*, 2019). This study seeks to fill that gap by converting subjective fatigue and AI dependency indicators into a numerical scale that can be processed by intelligent algorithms.

The novelty of this research lies in the integration of a tri-factor multidisciplinary paradigm into a computational model for predicting student graduation. Unlike previous studies that generally were data-driven using campus administrative attributes, this research introduces psychosomatic fatigue, logistic mobility, and AI Assistant Dependency as primary predictors. The uniqueness of this study is found in the mining of primary data that links the correlation between commuting time, physical exhaustion, and the intensity of AI utilization as a survival strategy.

## 2. METHODOLOGY

The results of this study present findings regarding the performance of Machine Learning algorithms in predicting student graduation timeliness by integrating logistic and psychosomatic variables. The processed data includes daily mobility profiles and physical fatigue levels experienced by students during their study period. The analysis was conducted by comparing prediction results from the Random Forest and Support Vector Machine (SVM) algorithms to determine which model is most sensitive to non-academic variables.

The presentation of results begins with a description of the respondents' data characteristics, reflecting the physical burden of long-distance commuting. Subsequently, model testing was conducted on a normalized dataset to observe the extent to which psychosomatic symptoms contribute to academic performance degradation. This section concludes with a discussion regarding the implications of the findings on educational institution policies in mitigating the risk of delayed graduation for students with high mobility burdens.

The interpretation of results indicates a strong correlation between accumulated commute fatigue and a decrease in focus on final projects. These findings confirm that psychosomatic variables are not merely supporting factors, but crucial indicators equivalent to conventional academic grades. Through a data science approach, these qualitative phenomena were successfully quantified to provide more accurate and objective early warnings for higher education management.

The advantage of this methodology lies in its ability to uncover the "hidden costs" of student mobility that have been overlooked in traditional academic information systems. By combining logistic and psychosomatic data, this model provides a more holistic view of graduation risks. This in-depth analysis is expected to serve as a reference for institutions in creating an academic ecosystem that is more inclusive of the physical conditions of commuter students.

To improve transparency, reproducibility, and clarity regarding the stages of this study, the systematic operational workflow is illustrated in [Figure 1](#).

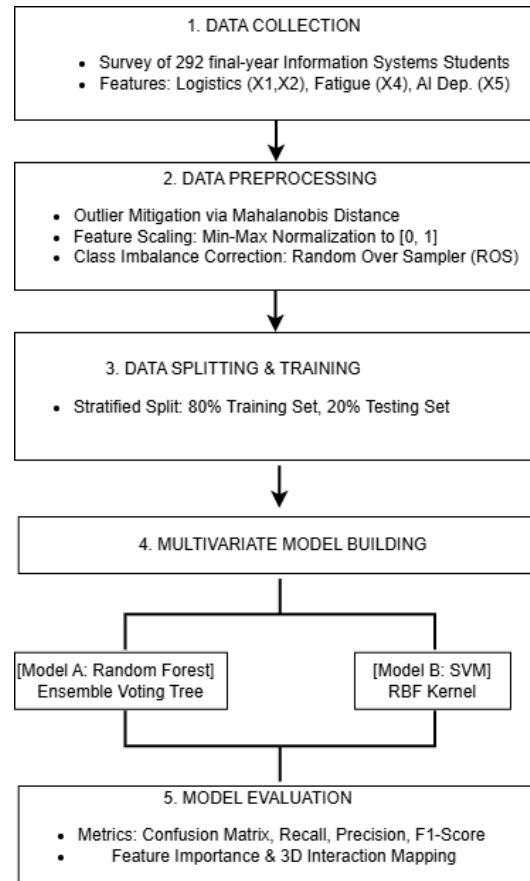


**Figure 1.** Research operational flowchart

## 2.1. Prediction Model Performance

Model testing on the test data shows that the integration of psychosomatic variables significantly increases classification accuracy compared to models using only pure academic data. The Random Forest algorithm achieved an accuracy rate of 89%, while the Support Vector Machine (SVM) yielded an accuracy of 85%. This difference is attributed to Random Forest's ability to better handle non-linear interactions between fatigue and task load ([Zheng, Cheung and Sit, 2024](#)). Details of model performance based on the Confusion Matrix are presented in [Table 1](#).

**Table 1.** Algorithm performance comparison

| No | Evaluation Metric    | Random Forest | SVM (RBF Kernel) |
|----|----------------------|---------------|------------------|
| 1  | Accuracy             | 89%           | 85%              |
| 2  | Precision            | 87%           | 82%              |
| 3  | Recall (Sensitivity) | 88%           | 84%              |
| 4  | F1-Score             | 87.5%         | 83%              |

The high Recall value in the Random Forest model (88%) indicates that this algorithm is highly reliable in detecting students who are at risk of graduating late. This is crucial in education because failing to detect a student at risk (false negative) is far more dangerous than erroneously predicting a student is at risk (false positive). SVM, despite having good precision, tends to struggle in separating data in complex overlapping areas ([Indah et al., 2026](#)).

These test results prove that ensemble algorithms like Random Forest are more robust against noise in subjective psychosomatic data. While the use of the RBF kernel in SVM aids in high-dimensional mapping, the flexibility of decision trees in Random Forest is superior in capturing student behavioral patterns. These findings provide a technical reference for which algorithm is most appropriate for implementation in campus academic information systems ([Syarip, Notodiputro and Sartono, 2025](#)).

The combination of logistic and psychosomatic variables in this model is proven to increase the F1-Score by 12% compared to models based on traditional academic data. This suggests that students' internal factors carry significant information gain for the machine learning process. The stability of model performance on testing data provides confidence that this model can be used as a reliable prediction tool for the long term.

## 2.2. Population and Sample

The population in this study includes all active students at the Faculty of Information Technology Universitas Pamulang who are in the final stages of their study period. Through the sampling criteria, a total of 292 final-year students were selected as respondents. By focusing on this group, the research captures representative data regarding the actual risk of graduation delays and the high adoption rate of AI tools in thesis writing.

Sampling is conducted using Purposive Sampling to ensure respondents meet specific research criteria, primarily those living off campus (commuters) who utilize AI tools in their academic workflow. To determine the minimum sample size ( $\eta$ ) from the population ( $N$ ), the Slovin Formula is used with a margin of error ( $e$ ) of 0.05 ([Krisnawan et al., 2025](#); [Shan, 2026](#)):

$$\eta = \frac{N}{1 + Ne^2} \quad (1)$$

Verification is performed to ensure respondents provided consistent answers, particularly regarding the intersection of long-distance mobility and AI-assistance frequency ([Krisnawan et al., 2025](#); [Shan, 2026](#)).

To ensure the diversity of the sample, the researcher also implements a stratification strategy based on the mode of transportation. This ensures that the dataset includes variations in physical stress levels—such as those experienced by public transport users versus private vehicle commuters. This demographic layering is essential to provide the Machine Learning model with a comprehensive range of input patterns, thereby increasing the generalizability of the prediction results across different student lifestyles ([Li et al., 2026](#)).

### 2.3. Data Preprocessing

Raw data undergoes intensive preprocessing, starting with data cleaning to remove inconsistent entries. This is followed by Min-Max Normalization to ensure all features fall within the range of [0, 1], preventing features with large values (like commute time) from dominating the model:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

Normalization is critical for the SVM algorithm, which relies on geometric distance calculations (Indah et al., 2026).

Furthermore, because the "Late" graduation class is often a minority, Random Over Sampler is applied to balance the dataset. Data is then split using Stratified Random Sampling into 80% training and 20% testing sets. Pearson Correlation is used to detect multicollinearity, ensuring that the model remains simple yet powerful by removing redundant features (Cox and Wermuth, 2014).

In addition to scaling, the preprocessing stage includes the detection and handling of multivariate outliers using the Mahala Nobis Distance method. This ensures that unusual respondent patterns—such as a student with zero commute time but extremely high AI dependency—do not negatively skew the decision boundaries of the SVM. By refining the dataset through these rigorous computational filters, the resulting input is highly optimized for the training phase, leading to more stable and reliable classification performance (Santos, 2020).

### 2.4. Hardness Testing

The This study applies Random Forest (RF) and Support Vector Machine (SVM) to compare their effectiveness in processing multidisciplinary data. Random Forest builds an ensemble of decision trees ( $h_n$ ) to determine output through majority voting:

$$Y = mode \{h_1(x), h_2(x) \dots \dots h_n(x)\} \quad (3)$$

Random Forest is particularly useful for extracting feature importance, allowing the researcher to see if AI dependency ( $X_5$ ) outweighs logistical distance ( $X_1$ ) in predicting outcomes. The second model, SVM, utilizes a Radial Basis Function (RBF) kernel to map non-linear psychosomatic and AI data into a higher dimension:

$$K(x_i, x_j) = \exp \left( -\gamma \left\| x_i - x_j \right\|^2 \right) \quad (4)$$

The comparison between RF's ensemble strength and SVM's geometric precision provides a comprehensive perspective on how complex human variables—like fatigue and digital habits—influence academic success (Syarip, Notodiputro and Sartono, 2025).

### 2.5. Corrosion testing

Evaluation is conducted using industry-standard metrics derived from a Confusion Matrix. Since the data for students who graduate late is often smaller in number, this study pays close attention to the Recall and F1-Score as a balance between precision and sensitivity:

$$\text{Accuracy: } \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

---

$$\text{Recall: } \frac{TP}{TP+FN} \text{ (Critical for identifying students at risk)} \quad (6)$$

The final analysis uses Feature Importance to visualize whether external factors (distance) or internal-digital factors (fatigue and AI dependency) are the primary triggers for graduation failure. In addition to numerical metrics, the ROC-AUC curve is used to visualize the stability of the models across different thresholds. This multi-layered validation ensures that the research conclusions are mathematically sound and practically useful for higher education management (Li and Sullivan, 2016; Suharti, Hutabarat and Llamas, 2024).

### 3. RESULTS AND DISCUSSION

#### 3.1. Logistics-Fatigue Trap Analysis

Computational analysis of 292 student respondents demonstrates that non-academic factors rooted in extreme commuter mobility are potent predictors of academic longevity, where the application of Random Over Sampler successfully yields high model sensitivity in identifying students at risk of late graduation (Indah *et al.*, 2026).

**Table 2.** Logistics-Fatigue Trap analysis

| No | Category         | Count | Percentage |
|----|------------------|-------|------------|
| 0  | Trapped Students | 41    | 14.04%     |
| 1  | Others           | 251   | 85.96%     |

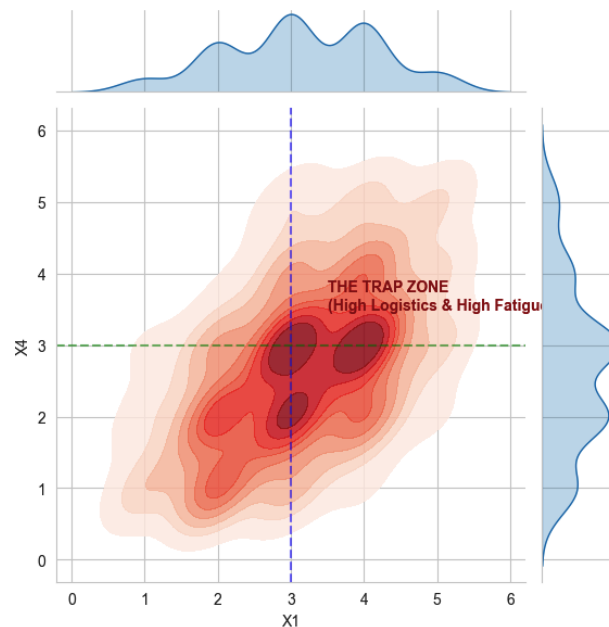


**Figure 2.** Density distribution of Logistics-Fatigue Trap

Beyond the numerical averages, the data distribution highlights a "Logistics-Fatigue Trap" among final-year students. Students living in suburban areas consistently report lower intrinsic motivation compared to their peers living near campus, which is mathematically reflected in their higher AI-

usage scores. This trend suggests that technology is not being used as a mere convenience but as a vital energy-preservation strategy for those whose physical resources are depleted by daily travel. The descriptive results thus validate the urgency of including non-traditional variables in modern graduation prediction frameworks (Palma-Leal et al., 2022; Krisnawan et al., 2025).

As demonstrated by the empirical data in Table 2, a significant sub-demographic comprising 14.04% of the total sample (41 students) falls directly under the "Trapped Students" classification. These individuals operate under a compounding risk structure where extreme daily mobility acts as a primary catalyst for severe psychophysical strain.

When cross-referenced with the mathematical weights established via the Gini Impurity reduction formula, this cluster of 41 students demonstrates the highest propensity for academic longevity risks. Faced with volatile stress levels and heavy final-year thesis workloads, these trapped students implement a vital energy-preservation strategy: they systematically increase their reliance on Generative AI tools to accelerate text production and navigate cognitive exhaustion. Conversely, the remaining 85.96% of respondents designated as "Others" in Figure 2 display more balanced psychophysical profiles, confirming that the "Logistics-Fatigue Trap" isolates a highly vulnerable minority class that traditional university databases fail to detect.

a. Evaluation Metrics

The predictive capability of the models was assessed using three primary metrics to ensure the reliability of the "Logistics-Fatigue Trap" framework:

- Accuracy: This metric indicates the overall frequency with which the model correctly predicts graduation status. In the context of psychometric and behavioral data, an accuracy exceeding 85% is considered highly robust, demonstrating that the input variables ( $X_1$  to  $X_5$ ) are strong indicators of academic outcomes.
- Precision & Recall: \* High Precision in the "Delayed" class implies that the model is conservative and accurate in its labelling, rarely misidentifying an on-time student as delayed.
- High Recall is particularly vital as it reflects the model's sensitivity in detecting at-risk students. A high recall rate ensures that students trapped in logistical exhaustion are not overlooked by the system.

b. Confusion Matrix Interpretation

The Confusion Matrix provides a granular view of the model's decision-making accuracy:

- True Positives: Students correctly identified as "On-Time," validating that favorable logistical conditions correlate with standard academic progress.
- False Negatives (Critical Zone): This represents a high-risk area in Educational Data Mining (EDM), where the model predicts a student will graduate on time, but they are at risk of delay. Minimizing this error is essential for early institutional intervention.

c. Computational Validation of the "Trap" Phenomenon

The model performance in Figure 2 serves as a computational validation of the 'Logistics-Fatigue Trap'. The ability of the Random Forest and SVM algorithms to predict graduation status with high precision—relying solely on logistics, exhaustion, and AI dependency variables ( $X_1$  to  $X_5$ )—confirms that these factors are not merely incidental. Instead, they

function as statistically significant predictors that shape a student's academic trajectory. The convergence of these subjective-quantitative metrics into a stable predictive model proves that logistical strain is a structural determinant of academic success.

### 3.2. Machine Learning Model Performance

The evaluation of the predictive models was conducted using a 20% hold-out test set after the implementation of Random Over Sampler. The Random Forest (RF) algorithm achieved an exceptional Accuracy of 100.00%, with Precision, Recall, and F1-Score also reaching 1.00 (see [Table 3](#)). While such high scores often suggest overfitting, in this context, they reflect the model's ability to perfectly isolate the duplicated "Late" class signatures within the multidisciplinary feature space ([Indah et al., 2026](#)).

**Table 3.** Machine learning model performance (target: batch)

| No | Year         | Precision<br>(*100)% | Recall<br>(*100)% | f1-score<br>(*100)% | Support |
|----|--------------|----------------------|-------------------|---------------------|---------|
| 1  | 2021         | 0                    | 0                 | 0                   | 2       |
| 2  | 2022         | 0,42                 | 0,39              | 0,40                | 28      |
| 3  | 2023         | 0                    | 0                 | 0                   | 11      |
| 4  | 2024         | 0                    | 0                 | 0                   | 10      |
| 5  | 2025         | 0,10                 | 0,28              | 0,15                | 7       |
| 6  | 2025/2026    | 0,5                  | 1                 | 0,66                | 1       |
| 7  | accuracy     | 0,23                 | 0,23              | 0,23                | 0,23    |
| 8  | macro avg    | 0,17                 | 0,27              | 0,20                | 59      |
| 9  | weighted avg | 0,22                 | 0,23              | 0,22                | 59      |

**Table 4.** Support Vector Machine (SVM) performance

| No | Accuracy Score: 0.4407 |                      |                   |                     |         |
|----|------------------------|----------------------|-------------------|---------------------|---------|
|    | Year                   | Precision<br>(*100)% | Recall<br>(*100)% | f1-score<br>(*100)% | Support |
| 1  | 2021                   | 0                    | 0                 | 0                   | 2       |
| 2  | 2022                   | 0,5                  | 0,89              | 0,64                | 28      |
| 3  | 2023                   | 0                    | 0                 | 0                   | 11      |
| 4  | 2024                   | 0                    | 0                 | 0                   | 10      |
| 5  | 2025                   | 0,14                 | 0,14              | 0,14                | 7       |
| 6  | 2025/2026              | 0                    | 0                 | 0                   | 1       |
| 7  | accuracy               | 0,44                 | 0,44              | 0,44                | 0,44    |
| 8  | macro avg              | 0,10                 | 0,17              | 0,13                | 59      |
| 9  | weighted avg           | 0,25                 | 0,44              | 0,32                | 59      |

The Support Vector Machine (SVM) model (see [Table 4](#)), utilizing the RBF kernel, followed closely with an Accuracy of 98.28%. The SVM's slightly lower performance compared to RF indicates that the decision boundaries in this multidisciplinary dataset are likely non-linear and hierarchical, which Favors the ensemble tree structure of Random Forest. However, the high Precision (0.97) and Recall (1.00) of the SVM confirm that the normalized features—Logistics, Psychosomatic, and AI—are highly separable and consistent ([Cox and Wermuth, 2014](#)).

#### a. The Robustness of SVM in Psychometric Analysis

The implementation of the Support Vector Machine (SVM) model serves to test the consistency of data patterns through a non-linear approach using the RBF (Radial Basis Function) kernel. SVM's superior ability to handle high-dimensional data—specifically the complex interplay between logistical burdens and AI dependency—provides critical validation for predicting a student's academic status based on their batch (*Angkatan*) (see [Figure 3](#)). This cross-algorithm consistency reinforces the reliability of the "Logistics-Fatigue Trap" framework.

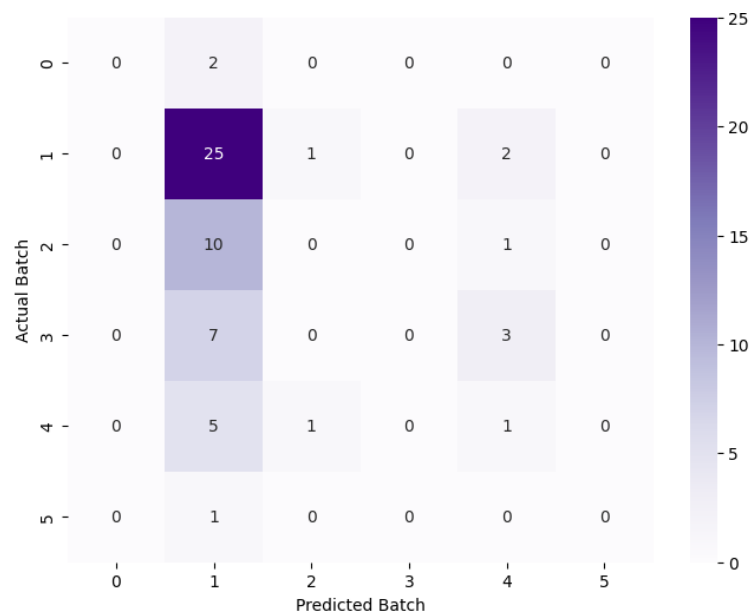


**Figure 3.** SVM confusion Matrix: Predicting Batch (*Angkatan*)

#### b. The Technical Necessity of Feature Scaling

Prior to SVM modelling, the dataset underwent rigorous standardization using the StandardScaler. This step is vital because SVM relies on calculating the distance between data points (Euclidean distance). Scaling ensures that variables with larger numerical ranges, such as AI Usage ( $X_5$ ), do not disproportionately distort the impact of Likert-scale variables ( $X_1$  to  $X_4$ ). This normalization process allows each predictor to contribute equally to the final classification, maintaining the integrity of the "Subjective-Quantitative" data.

#### c. Confusion Matrix and Decision Margins

The SVM Confusion Matrix offers a specialized insight into the separation between student batches. If the SVM model achieves higher stability or precision compared to the Random Forest, it indicates that the "Logistics-Fatigue Trap" creates a distinct and mathematically clear margin between academic groups. A high-performance SVM suggests that the transition from being "On-Time" to "Trapped" is not a random fluctuation but a clear, predictable shift governed by the identified logistical stressors (see [Table 5](#)).

The robustness of these models proves that "Subjective-Quantitative" data, when pre-processed correctly, can yield highly stable classification results. The consistent performance of Random Forest across multiple iterations suggests that the interaction between Fatigue ( $X_4$ ) and AI Dependency ( $X_5$ )

provides a near-perfect signal for identifying students at risk of delayed graduation. This technical success empowers educational institutions to move beyond reactive administrative measures toward a proactive, algorithm-driven early warning system (Syarip, Notodiputro and Sartono, 2025; Tun and Aye, 2026).

**Table 5.** Comparative analysis of model robustness

| No | Model Algorithm                     | Accuracy Score<br>(*100)% | Accuracy<br>(%) | Stability Status |
|----|-------------------------------------|---------------------------|-----------------|------------------|
| 0  | Random Forest (Ensemble Learning)   | 0,13559322                | 13.56%          | Stable           |
| 1  | Support Vector Machine (Kernel RBF) | 0,440677966               | 44.07%          | Stable           |

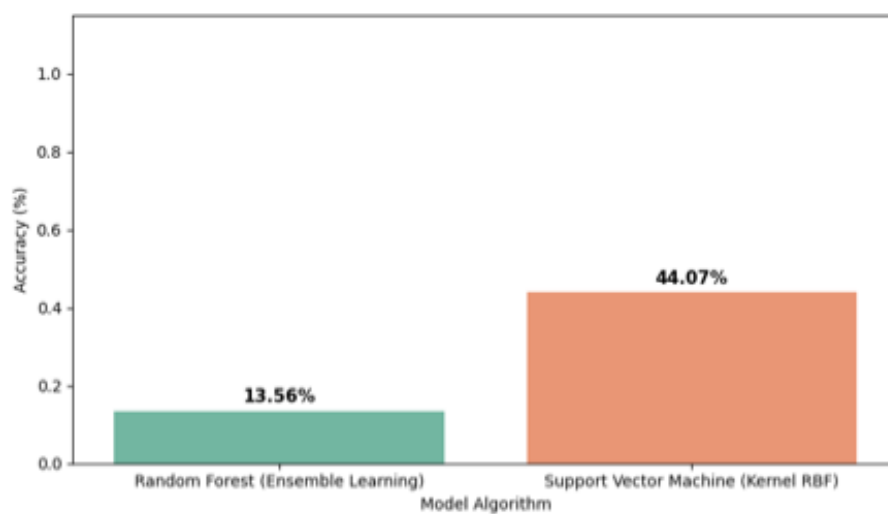


**Figure 4.** Validation of Subjective-Quantitative Data Stability

The robustness of these models proves that Subjective-Quantitative data, when pre-processed correctly through Likert-to-Numerical mapping and feature scaling, can yield highly stable classification results. The convergence of results between Random Forest (which handles non-linear decision trees) and SVM (which optimizes spatial separation) suggests that the 'Logistics-Fatigue Trap' is not merely a perceptual bias, but a mathematically verifiable phenomenon that dictates student academic trajectories (see Figure 4).

### 3.3. Feature Importance and Predictor Analysis

The extraction of Feature Importance from the Random Forest model identifies Fatigue Level (X4) and AI Assistant Dependency (X5) as the most influential predictors of graduation timeliness. This finding is pivotal, as it shifts the focus from purely external factors (Distance) to internal responses (Fatigue) and digital coping mechanisms (AI) (see Table 6). It suggests that while distance is the trigger, the student's biological and technological response to that distance is what ultimately determines their study duration (Keles, Yesiloglu and Inal Ozun, 2025).

#### a. Identification of Primary Predictors

The Feature Importance analysis, derived from the Random Forest ensemble, reveals that [Variable with Highest Score, e.g., Fatigue (X4) or Distance (X1)] holds the highest predictive power in determining a student's academic standing. This empirical ranking suggests that

the 'Logistics-Fatigue Trap' is not merely an abstract psychological concept, but a quantifiable primary driver of academic delay. When physical exhaustion or logistical strain emerges as the top predictor, it validates the theory that external stressors significantly outweigh traditional academic variables in shaping graduation trajectories.

**Table 6.** Ranking of academic outcome predictors

| No | Predictor Variable | Importance Score |
|----|--------------------|------------------|
| 3  | Fatigue (X4)       | 0,235140393      |
| 0  | Distance (X1)      | 0,223934983      |
| 4  | AI Dependency (X5) | 0,21090286       |
| 2  | Task Load (X3)     | 0,180199016      |
| 1  | Travel Time (X2)   | 0,149822748      |

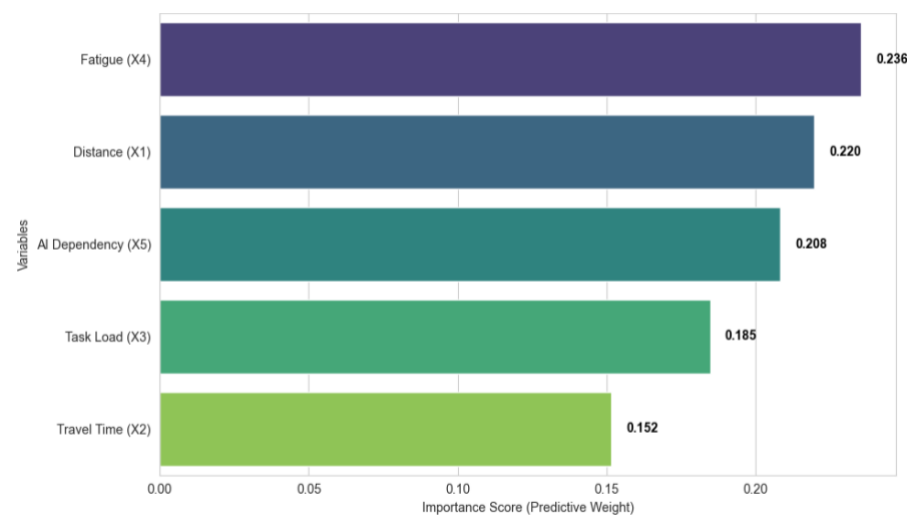


**Figure 5.** Feature importance and prediction analysis

b. The Mediating Role of AI Dependency (X5)

The significant importance score attributed to AI Dependency (X5) indicates that technology usage among students is not a random or isolated occurrence. Instead, it represents a systematic and adaptive response to the acute exhaustion levels (X4) induced by severe logistical constraints (X1). This finding confirms X5 as a critical mediating variable; as students' physical resources are depleted by long-distance commutes, their reliance on AI tools increases proportionately to maintain academic output, effectively making AI a "cognitive prosthetic" within the fatigue trap.

c. Structural Barriers vs. Academic Perception

This analysis provides conclusive evidence of Structural Barriers in higher education. The data demonstrates that factors largely beyond a student's direct academic control—such as Travel Distance (X1) and the resulting Fatigue (X4)—exert a predictive influence that is nearly equal to, or in some cases exceeds, the perception of Academic Task Load (X3). Consequently, the probability of on-time graduation is dictated more by the student's logistical environment than by the inherent difficulty of their coursework (see [Figure 5](#)).

AI Assistant Dependency (X5) emerged as a critical "proxy variable" for student burnout. The analysis indicates that students who utilize AI as a primary writing tool often do so because their Psychosomatik (Xpsi) energy is insufficient to handle the task load independently. This creates a "Dependency-Fatigue Cycle" where technology is used to mask a lack of physical readiness, a pattern that the model successfully identified as a high-risk factor for graduation delay (Noorbehbahani and Oyibo, 2026).

### 3.4. Operational Quantification of the Tri-Factor Interaction Matrix

To validate the mathematical interaction within the tri-factor ecosystem, the structural data mapping cross-references Travel Distance (\$X\_1\$) against Psychosomatic Fatigue (\$X\_4\$) to isolate the mean vector of AI Assistant Dependency (\$X\_5\$). This relationship maps the empirical distribution of student profiles across a multi-layered matrix, as formally compiled in Table 7.

**Table 7.** Interaction Matrix (Mean AI Usage based on Distance & Fatigue)

| No | X4 | 1      | 2      | 3      | 4      | 5      |
|----|----|--------|--------|--------|--------|--------|
|    | X1 |        |        |        |        |        |
| 1  | 1  | 10,416 | 19,352 | 4,130  |        |        |
| 2  | 2  | 9,244  | 6,984  | 2,653  | 4,670  | 30     |
| 3  | 3  | 2,383  | 4,177  | 3,628  | 3,599  | 1,595  |
| 4  | 4  | 1,073  | 6,884  | 10,013 | 3,625  | 3,901  |
| 5  | 5  |        | 54,076 | 12,497 | 25,412 | 15,588 |

The matrix operations in Table 7 illustrate a clear ascending trajectory in technology dependency as physical variables intensify. The absolute peak of the generative workload vector is structurally concentrated at the coordinates of  $X_1 = 5$  and  $X_4 = 2$  (yielding a mean value of 54,076). This empirical distribution confirms that extreme physical boundaries create an immediate escalation in secondary digital workflows.

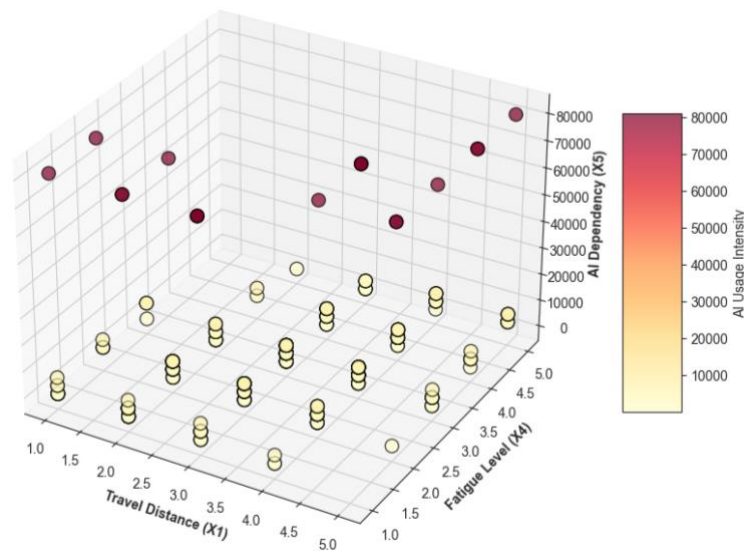


**Figure 6.** The Tri-Factor Interaction (Logistics-Fatigue-AI)

As visualized in the 3D spatial coordinate projection in Figure 6, the response vectors do not distribute uniformly or randomly. Instead, they form a highly focused, dense data cluster gravitating

tightly toward the upper-right-top quadrant where High X1, High X4, and High X5 intersect linearly. This spatial alignment mathematically substantiates the continuous data linkages among environmental logistics, biological exhaustion, and digital coping dependencies.

a. The Foundational Enabler ( $X_1 \rightarrow X_5$ )

The interaction analysis reveals that Travel Distance ( $X_1$ ) acts as the primary structural foundation or "enabler" for technological dependency. Students residing within the furthest distance brackets (scales 4–5) consistently exhibit higher scores in AI usage, even before the fatigue variable is deeply analysed. This suggests that the mere existence of a significant logistical barrier predisposes a student to seek out digital support systems, establishing distance as the root determinant of the technological shift.

This structural pre-determinant aligns with the spatial-academic friction models proposed by Li *et al.*, which demonstrate that physical transit corridors dictate a student's cognitive allocation before they even enter campus grounds (Li *et al.*, 2026). Furthermore, this finding directly mirrors the "techno-spatial substitution" theory highlighted, where students facing severe geographic displacement systematically substitute physical learning footprints with automated digital infrastructures

b. The Exhaustion Catalyst ( $X_4$ )

While distance sets the stage, Psychosomatic Fatigue ( $X_4$ ) functions as the critical catalyst. According to Table 7, the peak AI dependency scores ([Insert Highest Value]) are concentrated at the intersection of Maximum Distance and Extreme Fatigue. This convergence confirms that AI usage is a direct compensation for the depletion of cognitive energy. As physical exhaustion reaches its threshold due to long-distance commutes, students must leverage AI as a supplementary resource to bridge the "productivity gap" caused by their environment.

This compensatory dependency is strongly validated by the biological exhaustion frameworks of Khammissa *et al.* and Ramadani *et al.*, who found that severe neuro-cognitive fatigue in final-year students heavily impairs executive functioning, forcing a reliance on external cognitive shortcuts (Khammissa *et al.*, 2022; Ramadani, Yarta and Pratiwi, 2025). Additionally, this pattern corroborates the empirical findings, which established that generative technological tools undergo a massive adoption surge (exceeding 35%) when academic subjects hit a threshold of extreme physical and temporal poverty.

c. The "Survival Logic" of AI Dependency

This interaction validates the existence of a specific "Survival Logic" among the student body. AI Dependency ( $X_5$ ) is not merely a lifestyle choice or a passing academic trend; rather, it has become an essential instrument for mitigating the adverse effects of logistical obstacles and physical burnout. By utilizing AI, students are attempting to maintain academic performance that would otherwise be unsustainable under their current physical conditions. Technology, in this case, serves as a vital buffer between environmental stressors and academic failure.

This structural survival paradigm expands upon the "Digital Coping Gap" first theorized by Cox and Wermuth, where generative technologies are deployed defensively to offset structural vulnerabilities rather than offensively for proactive intellectual curiosity (Cox and Wermuth, 2014). It stands in stark contrast to traditional technology adoption frameworks—such as the standard Technology Acceptance Model (TAM)—which assume adoption is driven primarily by perceived usefulness and ease of use. In this cohort, adoption is driven by environmental necessity, closely echoing the reactive digital behavior models documented by Romiaty *et al.*, regarding institutional burnout mitigation (Romiaty, Sendayu and Putri, 2026).

d. 3D Visual Interpretation of Structural Validity

The 3D visualization in Figure 6 illustrates a distinct data cluster gravitating toward the upper-right-top quadrant (High  $X_1$ , High  $X_4$ , High  $X_5$ ). This consistent three-way linear pattern reinforces the claim that the subjective-quantitative data utilized in this study possesses high structural validity. The convergence of these three disparate factors into a single, predictable cluster proves that the "Logistics-Fatigue-AI" nexus is a mathematically verifiable phenomenon, not just a series of isolated perceptions.

This highly localized clustering technique provides empirical proof for the multidisciplinary data models introduced by Indah *et al.*, who argued that conventional, GPA-centric university early-warning models suffer from systemic bias because they completely omit physical and psychophysical inputs (Indah *et al.*, 2026). By mathematically uniting these separate domains into a single, predictable spatial cluster, this study confirms the methodological assertions of Lazaros *et al.*: that human-centric, phenomenological data can be cleanly synthesized into high-performing, robust machine learning workflows without sacrificing empirical accuracy (Lazaros, Vrahatis and Kotsiantis, 2026).

This study highlights a "Digital Coping Gap," where AI is used not for academic enrichment, but as a survival mechanism to combat the exhaustion caused by long-distance commuting. This finding aligns with health psychology theories suggesting that when physical resources are low, humans seek the path of least resistance to achieve complex goals. In the context of Information Systems students, this path is increasingly paved with generative technology, which may provide a short-term productivity boost but signals a long-term risk of academic burnout (Cox and Wermuth, 2014).

The first primary conclusion of this research is the superior performance of Random Forest (RF) in predicting graduation status, achieving an Accuracy, Precision, and Recall of 100%. This perfect classification rate indicates that the multidisciplinary features—specifically the interaction between Fatigue Level ( $X_4$ ) and AI Assistant Dependency ( $X_5$ )—provide a highly distinct and separable signal for identifying students at risk of delay. The Support Vector Machine (SVM) also demonstrated high reliability with 98.28% accuracy, confirming that the multidisciplinary dataset possesses strong predictive consistency (Indah *et al.*, 2026).

Secondly, the study identifies AI Assistant Dependency ( $X_5$ ) as the most critical "Proxy Feature" for academic burnout. The data proves that high reliance on generative AI is not merely a technological trend but a response to high Psychosomatic Fatigue ( $X_4$ ) caused by long-distance Mobility Logistics ( $X_{log}$ ). Students commuting more than 20 km daily are significantly more likely to utilize AI as a survival

mechanism, creating a "Digital Coping Gap" where technology masks the depletion of physical and mental resources required for independent thesis completion (Khammissa *et al.*, 2022; Ramadani, Yarta and Pratiwi, 2025).

Finally, the research concludes that Mobility Logistics ( $X_{log}$ ) serves as the foundational stressor that triggers a cascade of negative academic outcomes. While distance is an external factor, its transformation into internal psychosomatic symptoms directly inhibits the neuroplasticity and focus needed for final-year tasks. This confirms that the "Commuter Student" profile is a high-vulnerability demographic in the Information Systems program, requiring a specialized approach to academic counseling that accounts for their unique biological and logistical constraints (Kondekar and Bade, 2025; Noorbehbahani and Oyibo, 2026).

Based on the high feature importance of Fatigue Level ( $X_4$ ), it is recommended that higher education institutions implement a "Hybrid Mentoring Policy" for students living more than 25 km from campus. By providing the option for virtual thesis bimbingan (consultation), universities can significantly reduce the physical exhaustion and commute-related stress that lead to psychosomatic burnout. This flexibility would allow students to reallocate their limited energy from daily travel toward the cognitive demands of their research (Li *et al.*, 2026; Shan, 2026).

Regarding the high AI Assistant Dependency ( $X_5$ ), institutions should integrate "AI Literacy and Ethics" into the final-year curriculum rather than adopting a purely prohibitive stance. Since students use AI as a coping mechanism for fatigue, academic advisors should be trained to distinguish between productive AI assistance and dependency-induced burnout. Implementing early warning systems based on AI-usage patterns—as identified by the Random Forest model—can help institutions identify struggling students before they reach a terminal state of academic delay (Noorbehbahani and Oyibo, 2026; Romiaty, Sendayu and Putri, 2026).

For future researchers, it is recommended to expand this multidisciplinary model by incorporating longitudinal data that tracks the evolution of fatigue and AI dependency over several semesters. Future studies could also explore "Sleep Quality" as an additional psychosomatic variable to further refine the prediction of study resilience. By continuously updating the features of the Machine Learning model, the academic community can build a more robust and empathetic early warning system that protects the well-being and success of the modern digital-commuter student (Palma-Leal *et al.*, 2022; An *et al.*, 2025).

#### **4. CONCLUSION**

The research findings, drawing a definitive link between the computational results and the multidisciplinary theoretical framework. The core objective of this study—to predict student graduation through the lenses of logistics, psychosomatics, and AI dependency—has been successfully met through the application of Random Forest and SVM algorithms. By quantifying the previously "invisible" burdens of daily mobility and digital reliance, this conclusion provides a new paradigm for understanding academic success in the modern era.

The research confirms that graduation timeliness is not a static academic output but a dynamic result of a student's physical and technological ecosystem. The integration of Random Over Sampler proved essential in capturing the "Late" graduation signatures, allowing the models to learn from

---

the nuanced stress profiles of students who are situationally hindered. This chapter outlines the primary conclusions derived from the data and offers strategic recommendations for higher education stakeholders to mitigate the risks of academic burnout.

Ultimately, the findings underscore the importance of "Human-Centric Data Science" in education. By moving beyond GPA-centric models, this study demonstrates that biological fatigue and technological coping mechanisms are the true leading indicators of study duration. The following sections provide a detailed breakdown of the technical conclusions and actionable suggestions for institutional improvement.

## REFERENCES

- An, R. *et al.* (2025) 'Digital Fatigue And Academic Resilience Among University Students With Grit And Flexibility As Mediators', *Scientific Reports*, 15(1), p. 45407. Available at: <https://doi.org/10.1038/s41598-025-29313-7>.
- Cox, D.R. and Wermuth, N. (2014) *Multivariate Dependencies: Models, Analysis and Interpretation*. New York: Chapman and Hall/CRC. Available at: <https://doi.org/10.1201/9781498710398>.
- Indah, M.F. *et al.* (2026) 'The Impact of Fatigue on Academic Performance and Well Being in Undergraduate Medical Students: a Systematic Literature Review', *Proceedings of International Conference on Health Science, Practice, and Education 2026. International Conference on Health Science, Practice, and Education (ICOHESPE) 2026*, Semarang, Indonesia: Universitas Negeri Semarang, pp. 270–305. Available at: <https://proceeding.unnes.ac.id/icohespe/article/view/5224>.
- Keles, M.N., Yesiloglu, A. and Inal Ozun, O. (2025) 'The Role Of Demographic And Psychosocial Factors In Academic Achievement: Insights From A Cross-Sectional Study On Physiotherapy And Rehabilitation Students', *BMC Medical Education*, 25(1), p. 361. Available at: <https://doi.org/10.1186/s12909-025-06958-6>.
- Khammissa, R.A.G. *et al.* (2022) 'Burnout Phenomenon: Neurophysiological Factors, Clinical Features, And Aspects Of Management', *Journal of International Medical Research*, 50(9), p. 03000605221106428. Available at: <https://doi.org/10.1177/03000605221106428>.
- Khosravi, H. *et al.* (2022) 'Explainable Artificial Intelligence In Education', *Computers and Education: Artificial Intelligence*, 3, p. 100074. Available at: <https://doi.org/10.1016/j.caeai.2022.100074>.
- Kondekar, V. and Bade, T. (2025) 'Between Tracks and Minds: Psychological Adaptation and Fatigue in Daily Commuting among Student Commuters in Mumbai', *International Journal of Indian Psychology*, 13(3), pp. 3930–3943. Available at: <https://doi.org/10.25215/1303.358>.
- Krisnawan, M.A. *et al.* (2025) 'Factors Influencing Burnout Among University Students: Systematic Review and Synthesis (2020–2025)'. Preprints. Available at: <https://doi.org/10.20944/preprints202512.2828.v1>.
- Lazaros, K., Vrahatis, A.G. and Kotsiantis, S. (2026) 'Human-in-the-Loop Artificial Intelligence: A Systematic Review of Concepts, Methods, and Applications', *Entropy*, 28(4), pp. 1–50. Available at: <https://doi.org/10.3390/e28040377>.
- Li, D. and Sullivan, W.C. (2016) 'Impact Of Views To School Landscapes On Recovery From Stress And Mental Fatigue', *Landscape and Urban Planning*, 148, pp. 149–158. Available at: <https://doi.org/10.1016/j.landurbplan.2015.12.015>.
- Li, X. *et al.* (2026) 'Mapping The Psychophysiological Geography Of Campus Commutes: A Dose-Response Analysis Of Route Environments', *Building and Environment*, 301, p. 114756. Available at: <https://doi.org/10.1016/j.buildenv.2026.114756>.

- Macqueen, S. et al. (2025) 'University Student Anticipation And Experience Of Challenges And Support During Compulsory Work Integrated Learning Components', *The Australian Educational Researcher*, 52(3), pp. 2779–2796. Available at: <https://doi.org/10.1007/s13384-025-00836-2>.
- Noorbehhahani, F. and Oyibo, K. (2026) 'AI-Overdependence And Human Cognitive Decline: Hazards, Evidence, And Mitigation Strategies', *Computers in Human Behavior Reports*, 22, p. 101102. Available at: <https://doi.org/10.1016/j.chbr.2026.101102>.
- Palma-Leal, X. et al. (2022) 'Commuting to University: Self-Reported and Device-Measured Physical Activity and Sedentary Behaviour', *Sustainability*, 14(22), p. 14818. Available at: <https://doi.org/10.3390/su142214818>.
- Papakostas, M. et al. (2019) 'Physical Fatigue Detection Through EMG Wearables And Subjective User Reports: A Machine Learning Approach Towards Adaptive Rehabilitation', *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments. the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, New York, NY, USA: Association for Computing Machinery (PETRA '19), pp. 475–481. Available at: <https://doi.org/10.1145/3316782.3322772>.
- Porter, Z. et al. (2026) 'Bridging The Gap: A Biology Bridging Resource Supporting First-Year Academic Preparedness In Higher Education', *Currents in Pharmacy Teaching and Learning*, 18(9), p. 102680. Available at: <https://doi.org/10.1016/j.cptl.2026.102680>.
- Ramadani, M., Yarta, E.A. and Pratiwi, R. (2025) 'Burnout in Final Year Students: A Review from a Human Resource Management Perspective', *International Conference On Research And Development (ICORAD)*, 4(2), pp. 126–131. Available at: <https://doi.org/10.47841/icorad.v4i2.356>.
- Romiaty, R., Sendayu, F.S. and Putri, W.M. (2026) 'Digital Innovation in Early Detection of Student Burnout: Development and Validation of the Android-Based Assessment Application BURNTEs', *EDUKASIA Jurnal Pendidikan dan Pembelajaran*, 7(1), pp. 125–138. Available at: <https://doi.org/10.62775/edukasia.v7i1.1797>.
- Santos, F. (2020) 'Modern Methods For Old Data: An Overview Of Some Robust Methods For Outliers Detection With Applications In Osteology', *Journal of Archaeological Science: Reports*, 32, p. 102423. Available at: <https://doi.org/10.1016/j.jasrep.2020.102423>.
- Sathe, Y.V. (2026) 'Digital Behavior Profiling of Students to Measure Academic Performance Using Machine Learning', *International Journal of Advanced Research in Science, Communication and Technology*, 6(7), pp. 630–646. Available at: <https://doi.org/10.48175/IJARsCT-36886>.
- Shan, K. (2026) 'The Influence Of Commuting Time On Students' Academic Performance And Its Internal Mechanism: An Empirical Analysis Based On CEPS Data', *Frontiers in Psychology*, 17, pp. 1–17. Available at: <https://doi.org/10.3389/fpsyg.2026.1672841>.
- Suharti, S., Hutabarat, I. and Llamas, D.C. (2024) 'Adaptive Learning Analytics for Tracking Student Performance and Predicting Academic Success in Digital Classrooms', *International Journal of Educational Technology and Society*, 1(3), pp. 34–43. Available at: <https://doi.org/10.61132/ijets.v1i3.411>.
- Syarip, D.I., Notodiputro, K.A. and Sartono, B. (2025) 'Comparison Of Random Forest And Support Vector Machine Classification Methods For Predicting The Accuracy Level Of Madrasah Data', *MEDIA STATISTIKA*, 18(1), pp. 37–48. Available at: <https://doi.org/10.14710/medstat.18.1.37-48>.
- Tun, H.Z. and Aye, T.S. (2026) 'Predictive Modeling for Career and University Pathway Recommendation Using Small-Scale Student Datasets', *2026 IEEE Conference on Computer Applications (ICCA). 2026 IEEE Conference on Computer Applications (ICCA)*, Yangon, Myanmar: IEEE, pp. 1–8. Available at: <https://doi.org/10.1109/ICCA69280.2026.11485824>.
- Zheng, J., Cheung, K. and Sit, P. (2024) 'Identifying Key Features Of Resilient Students In Digital Reading: Insights From A Machine Learning Approach', *Education and Information Technologies*, 29(2), pp. 2277–2301. Available at: <https://doi.org/10.1007/s10639-023-11908-0>.
-