

When Machine Learning Meets SWOT Analysis: XGBoost Prediction for Student On-Time Graduation

Intan Siti Amaliyah¹, Samidi²

¹ Master's Program in Management Science, Faculty of Economics and Business, Universitas Padjadjaran

² Master's Program in Computer Science, Universitas Budi Luhur

Article Info

Article history:

Received September, 2026

Revised September, 2026

Accepted September, 2026

Keywords:

Feature Importance;
Higher Education Strategic
Management;
On-Time Graduation Prediction;
Swot Analysis;
XGBoost

ABSTRACT

On-time graduation is a key performance indicator for higher education institutions, including Universitas Padjadjaran, yet on-time graduation rates vary substantially across faculties (25.99%–87.80%). This study aims to build a predictive model of student on-time graduation and translate its results into managerial strategy recommendations. Using the CRISP-DM framework, six classification algorithms (Decision Tree, Random Forest, Naive Bayes, Logistic Regression, Gradient Boosted Tree, and XGBoost) were compared on academic and socioeconomic data from 25,872 students. The default (non-tuned) XGBoost model achieved the best performance (Accuracy 75.14%; AUC 0.832; F1-Score 76.82%), outperforming Logistic Regression and Gradient Boosted Tree, which both recorded 73.33% Accuracy. Feature importance analysis identified Faculty (35.76%) and Cumulative GPA (29.34%) as the two most dominant factors, jointly accounting for 65.10% of the total feature importance score. These dominant factors were then integrated into a SWOT analysis, yielding an IFAS score of 3.10 and an EFAS score of 2.95, placing Universitas Padjadjaran in Quadrant I (Growth) of the Internal-External Matrix. The TOWS matrix recommends integrating the XGBoost model into the university's integrated academic information system (SIAT) as an early warning system, prioritized for the five faculties with the lowest on-time graduation rates. This study demonstrates that integrating predictive analytics with strategic management can generate more targeted, evidence-based academic policy recommendations.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Name: Intan Siti Amaliyah

Institution: Master's Program in Management Science, Faculty of Economics and Business, Universitas Padjadjaran

Email: Intan.amaliyah@unpad.ac.id

1. INTRODUCTION

Timely student graduation is one of the key performance indicators of higher education institutions in Indonesia, as reflected in the Educational Efficiency Rate (*Angka Efisiensi Edukasi* [AEE]), which

forms part of the Key Performance Indicators (KPIs) for higher education institutions [1]. At Universitas Padjadjaran, of the 25,872 students analyzed, only 52.25% graduated on time, with substantial variation in timely graduation rates across faculties, ranging from 25.99% in the Faculty of Law to 87.80%

in the Faculty of Pharmacy. This disparity indicates that delayed graduation is not merely an issue attributable to individual students but also reflects structural differences across academic units, requiring data-driven and differentiated interventions. Efforts to predict student success and study delays have been extensively examined through educational data mining and learning analytics approaches [2], [3]. These studies generally stop at identifying the best-performing predictive models and the factors contributing to predictions, without explicitly translating such findings into applicable strategic recommendations for institutional-level decision-makers. Conversely, studies of strategic management in higher education employing SWOT/TOWS analysis [4], [5] are generally qualitative and descriptive and are not based on empirically measurable factor weights derived from institutional data [6]. This gap between the two approaches provides the starting point for the present study: how predictive analytics results, particularly feature importance scores derived from machine learning models, can be systematically used as a basis for weighting factors in SWOT/TOWS analysis to formulate more targeted managerial strategies.

Based on this rationale, this study aims to: (1) compare the performance of six classification algorithms, namely Decision Tree, Random Forest, Naive Bayes, Logistic Regression, Gradient Boosted Tree, and XGBoost, in predicting students' timely graduation; (2) identify the academic, structural, and socioeconomic factors that contribute most substantially to these predictions through feature importance analysis; and (3) integrate these dominant factors into SWOT analysis and the TOWS Matrix to formulate managerial strategic recommendations for Universitas Padjadjaran. The primary contribution of this study lies in extending the CRISP-DM framework to the stage of strategic deployment by using feature importance scores as the basis for quantitative weighting in the Internal Factor Analysis Summary (IFAS) Matrix. Consequently, the strategy formulation process no longer relies entirely

on subjective qualitative assessments but is grounded in empirical evidence generated through predictive modeling.

2. LITERATURE REVIEW

2.1 *Predicting Academic Success Using Machine Learning*

Various classification algorithms have been widely used to predict student academic success and study delays, including Decision Tree, Random Forest [7], Naive Bayes [8], Logistic Regression, Gradient Boosted Tree [9], and XGBoost [10]. Systematic literature reviews indicate that academic factors, such as grades and grade point average (GPA), and structural factors, such as study program and faculty affiliation, consistently represent the strongest predictors of academic success [2], [3], [11]. In contrast, socioeconomic factors tend to make weaker and less consistently linear contributions.

Recent studies in the Indonesian higher education context have further confirmed the relevance of this approach. [12] applied machine learning to predict timely graduation in Indonesian higher education institutions, while [13] compared the performance of various timely graduation prediction models using resampling techniques and hyperparameter tuning. [14], [15] respectively explored deep learning and educational data mining approaches to predict student graduation timing. A systematic review by [16] further confirmed that the performance of predictive models for academic success depends substantially on the quality of feature engineering and algorithm selection. These findings support the sequential approach adopted in this study, which involves model comparison followed by feature importance extraction.

2.2 *The CRISP-DM Framework*

The Cross-Industry Standard Process for Data Mining (CRISP-DM) is a widely recognized framework for data mining projects consisting of six stages:

business understanding, data understanding, data preparation, modeling, evaluation, and deployment [17]. This study extends the deployment stage beyond merely interpreting the contribution of predictive factors by translating the analytical findings into managerial strategic recommendations.

2.3 *SWOT Analysis, IFAS-EFAS Matrices, and the TOWS Matrix*

SWOT analysis is a matching tool that connects an organization's internal strengths and weaknesses with external opportunities and threats [4]. The IFAS and EFAS matrices are used to systematically weight and evaluate internal and external factors, while the TOWS Matrix [18] is used to formulate strategic alternatives based on combinations of these factors, namely SO, WO, ST, and WT strategies. Conventionally, the weighting of factors in these matrices is qualitative and based on expert judgment. This study proposes the use of feature importance scores derived from predictive modeling as a basis for weighting internal factors, thereby providing a more measurable operationalization of the internal analytical dimension. The integration of data-driven analytical findings into SWOT analysis has also increasingly been explored in educational contexts [19], further supporting the relevance of the combined predictive analytics and strategic management approach adopted in this study.

3. RESEARCH METHODS

This study employed a quantitative approach with a non-experimental design that combines predictive modeling and descriptive strategic analysis. The research object consisted of undergraduate students at Universitas Padjadjaran, with a population/sample of 25,872 students who had been recorded as graduates and classified into two graduation status groups: Timely Graduation and Delayed Graduation. The variables analyzed included academic data,

comprising cumulative grade point average (GPA), semester GPA, and total completed credit hours; structural data, comprising faculty and study program; demographic data, comprising city or regency of origin; socioeconomic data, comprising parental occupation; and student activity data, comprising the number of student activities and achievements.

The research procedures followed the CRISP-DM framework. During the data understanding and data preparation stages, the cohort year attribute was excluded from the modeling process because it posed a risk of data leakage, and all categorical attributes were transformed using the Nominal to Numerical technique through dummy coding. Feature selection was conducted using the Select by Weights method based on Information Gain scores, with the top 15 attributes retained. During the modeling stage, the six classification algorithms were trained and evaluated using their default configurations without hyperparameter tuning. This decision was based on preliminary testing, which indicated that the default configurations produced performance comparable to or better than the tuned models for this dataset. Model performance was evaluated using Accuracy, Precision, Recall, F1-Score, and the Area Under the Receiver Operating Characteristic Curve (AUC).

Before the modeling stage, the dataset was divided into training and testing datasets using a 70:30 ratio. Seventy percent of the data was used to develop and train the models, while the remaining 30% was used to evaluate the models' ability to predict previously unseen observations. This hold-out scheme was selected to provide an objective evaluation of model performance and reduce the risk of testing bias. All performance metrics reported in the Results and Discussion section were therefore based on performance on the testing dataset (30%), rather than on the training dataset. During the subsequent evaluation stage involving interpretation and deployment, feature importance scores were extracted from the best-performing model using the AttributeWeights operator in RapidMiner

Studio. Because the Nominal to Numerical transformation generated individual scores for each category, such as “Faculty = Medicine” and “Faculty = Law,” these scores were subsequently aggregated to the level of their parent attributes, such as “Faculty,” by summation. This aggregation enabled the findings to be interpreted meaningfully within the context of institutional strategy.

The dominant factors identified through the aggregation process were subsequently mapped into the SWOT components. Actionable internal factors, including cumulative GPA and semester GPA, completed credit hours, and the number of student activities and achievements, were categorized as Strengths or Weaknesses. Faculty and Study Program, which represent structural and heterogeneous contexts rather than factors with a unidirectional effect, were used as segmentation variables to determine priority targets for implementation rather than being treated as separate Strength or Weakness factors. External factors, comprising Opportunities and Threats, were derived from a review of higher education policy documents and relevant literature. The weights in the IFAS Matrix were calculated from renormalized feature importance scores, excluding Faculty, Study Program, City, and Father's Occupation, whereas the weights in the EFAS Matrix were based on qualitative estimates referring to [4], [5]. Ratings on a scale of 1 to 4 were determined based on actual performance conditions reflected in the

historical data. The total IFAS and EFAS scores were subsequently plotted in the Internal-External Matrix to determine the strategic quadrant and were then incorporated into the TOWS Matrix to formulate SO, WO, ST, and WT strategic alternatives.

4. RESULTS AND DISCUSSION

4.1 Data Characteristics and Patterns of Timely Graduation

Of the 25,872 students analyzed, 52.25% were classified as graduating on time. The Timely Graduation group consistently demonstrated higher academic performance from the first semester onward, with differences in semester GPA ranging from 0.17 to 0.22 points in every semester compared with the Delayed Graduation group. Timely graduation rates also varied substantially across faculties, ranging from 25.99% in the Faculty of Law to 87.80% in the Faculty of Pharmacy. Student engagement factors also showed a notable pattern: students with three or more recorded activities or achievements had a substantially higher timely graduation rate of 66.23%, compared with 50.50% among students with no recorded activities or achievements.

4.2 Predictive Model Selection Through Performance Comparison

Table 1. Performance Comparison of Six Classification Algorithms

Algorithm	Accuracy	Precision	Recall (Timely Graduation)	Recall (Delayed Graduation)	F1-Score	AUC
Decision Tree	59.39%	56.50%	96.74%	18.51%	71.34%	0.577
Random Forest	62.47%	58.57%	96.23%	25.53%	72.82%	0.799
Naive Bayes	69.18%	68.03%	77.39%	60.20%	72.40%	0.766
Logistic Regression	73.33%	70.33%	84.66%	60.93%	76.84%	0.818
Gradient Boosted Tree	73.33%	70.63%	83.80%	61.87%	76.65%	0.818
XGBoost	75.14%	74.91%	78.84%	71.10%	76.82%	0.832

Source: Model testing results obtained using RapidMiner Studio and processed by the researchers.

As presented in Table 1, XGBoost achieved the best overall performance among the six algorithms tested, with an Accuracy of 75.14%, a Precision of 74.91%, and an AUC of 0.832, the highest AUC value among all models. Although Logistic Regression achieved a slightly higher F1-Score of 76.84%, XGBoost demonstrated a more balanced ability to detect both graduation classes. Its Recall for the Delayed Graduation class reached 71.10%, substantially higher than those of Logistic Regression and Gradient Boosted Tree, which achieved Recall

values of approximately 61% and 62%, respectively, for this class. This capability is particularly important for an early warning system that aims to identify students at risk of delayed graduation as accurately as possible. Based on its overall performance advantage and balanced classification capability, XGBoost was selected as the primary model for subsequent analyses, including feature importance extraction and managerial strategy formulation.

4.3 Dominant Factors: Feature Importance Analysis

Table 2. Fourteen Attributes with the Highest Feature Importance Scores in the XGBoost Model After Aggregation

Rank	Attribute	Feature Importance Score	Normalized Score
1	Faculty, aggregated across all categories	5,737.66	35.76%
2	Cumulative GPA	4,708.01	29.34%
3	Study Program, aggregated across all categories	1,507.98	9.40%
4	Semester 7 GPA	719.50	4.48%
5	Completed Credit Hours	663.01	4.13%
6	Semester 6 GPA	464.51	2.89%
7	Semester 1 GPA	432.09	2.69%
8	Semester 5 GPA	340.59	2.12%
9	Semester 2 GPA	328.80	2.05%
10	Semester 3 GPA	328.28	2.05%
11	City/Regency of Origin, aggregated across all categories	294.50	1.84%
12	Semester 4 GPA	255.56	1.59%
13	Number of Student Activities/Achievements	143.38	0.89%
14	Father's Occupation, aggregated across all categories	122.73	0.76%

Source: Processed by the researchers based on XGBoost AttributeWeights results in RapidMiner Studio. The model was retrained without the Cohort Year attribute.

Feature importance extraction from the best-performing XGBoost model identified Faculty (35.76%) and Cumulative GPA (29.34%) as the two most dominant factors, jointly accounting for 65.10% of the total feature importance score, as presented in Table 2. These factors were followed by Study Program (9.40%), Semester 7 GPA (4.48%), and Completed Credit Hours (4.13%). The interpretation of these findings indicates that Faculty and Study Program primarily represent structural contexts rather than factors with a unidirectional effect. This interpretation

is supported by the substantial variation in timely graduation rates across faculties, ranging from 25.99% to 87.80%. These variables were therefore treated as segmentation factors for determining institutional priorities rather than as direct intervention variables.

In contrast, Cumulative GPA and semester GPA measures demonstrated a clear academic pattern. Higher academic performance was associated with a greater likelihood of timely graduation. Semester 7 GPA showed the largest difference between the two groups, at 0.21 points, suggesting

that academic performance near the end of the study period represents a particularly important indicator of graduation outcomes.

Conversely, socioeconomic and demographic factors, although retained during the feature selection stage based on Information Gain scores, demonstrated substantially weaker contributions and did not exhibit clear patterns. Variation in timely graduation rates across students' cities or regencies of origin was relatively narrow, ranging from 47% to 56%, with no clear geographical pattern. Similarly, variation across categories of fathers' occupations was limited, ranging from 49.63% to 53.98%. The contributions of these two factors to the total feature importance score were only 1.84% for City/Regency of Origin and 0.76% for Father's Occupation, substantially lower than those of academic and structural factors. These findings suggest that data-driven intervention strategies should prioritize academic factors and structural contexts that demonstrate clearer predictive patterns rather than demographic variables that are empirically less

discriminative within the context of this study.

4.4 Integration into SWOT Analysis and IFAS-EFAS Weighting

The dominant factors were subsequently mapped into the SWOT framework. Cumulative GPA and academic performance during the later semesters, represented by Semester GPA for Semesters 5 to 7, were categorized as internal strengths. In contrast, achievement gaps during the earlier semesters, represented by Semester GPA for Semesters 1 to 4, together with low participation in student activities and achievements, were categorized as internal weaknesses.

On the external side, the Educational Efficiency Rate policy [1], the availability of educational analytics technologies, and the Merdeka Belajar-Kampus Merdeka policy were identified as opportunities. Accreditation requirements, limited capacity among academic advisors, and the increasing tendency of students to undertake part-time employment were identified as threats.

Table 3. Internal Factor Analysis Summary (IFAS) Matrix

Internal Factors	Weight	Rating	Score
S1: Cumulative GPA	0.56	3	1.68
S2: Semester 7 GPA	0.09	3	0.27
S3: Semester GPA Momentum in Semesters 5 and 6	0.09	3	0.27
W1: Academic Achievement Gap in Semesters 1 to 4	0.16	4	0.64
W2: Completed Credit Hours, indicating weak progress monitoring	0.08	2	0.16
W3: Low participation in student activities and achievements	0.02	4	0.08
Total IFAS	1.00		3.10

Source: Processed by the researchers. Weights were renormalized from the feature importance scores presented in Table 2 with reference to [4], [5].

Table 4. External Factor Analysis Summary (EFAS) Matrix

External Factors	Weight	Rating	Score
O1: KPI 1 Policy and the Educational Efficiency Rate	0.22	4	0.88
O2: Availability of educational analytics technologies	0.18	3	0.54
O3: MBKM Policy, including recognition of off-campus credits	0.15	3	0.45
T1: Accreditation requirements related to Educational Efficiency Rate performance	0.18	3	0.54
T2: Limited capacity of academic advisors	0.15	2	0.30
T3: Increasing tendency of students to work part-time due to economic pressures	0.12	2	0.24

External Factors	Weight	Rating	Score
Total EFAS	1.00		2.95

Source: Processed by the researchers based on policy document reviews and relevant literature. The weights represent qualitative estimates based on [4], [5].

The combination of an IFAS score of 3.10 and an EFAS score of 2.95 places Universitas Padjadjaran in Quadrant I of the Internal-External Matrix, corresponding to a Growth and Build position. According to [4], this position indicates that intensive strategies represent the most appropriate strategic orientation. The institution is therefore positioned to leverage its existing internal strengths to capture external opportunities rather than primarily adopting defensive or turnaround strategies.

4.5 Strategy Formulation Through the TOWS Matrix

The results of the SWOT analysis were subsequently integrated into the TOWS Matrix to formulate strategic alternatives. The Growth and Build position indicates that Universitas Padjadjaran has relatively strong internal capabilities and a favorable external environment for improving timely graduation performance. The proposed strategies therefore emphasize the use of academic strengths and data-driven capabilities to address persistent graduation challenges.

Table 5. Strategic Alternatives Formulated Through the TOWS Matrix

Strategy Type	Strategic Alternatives
SO Strategies	SO1: Develop an integrated early warning system for timely graduation using predictive analytics to identify students with a high risk of delayed graduation based on academic performance indicators. SO2: Utilize educational analytics technologies to support data-driven academic monitoring and intervention at the faculty and study program levels.
WO Strategies	WO1: Strengthen early-semester academic intervention programs to address achievement gaps in Semesters 1 to 4 through systematic monitoring and targeted academic support. WO2: Expand student participation in academic and extracurricular activities while integrating these activities into structured student development programs.
ST Strategies	ST1: Utilize strong academic performance indicators and predictive analytics to support compliance with accreditation requirements related to timely graduation and educational efficiency. ST2: Develop differentiated intervention strategies across faculties and study programs to optimize limited academic advising capacity.
WT Strategies	WT1: Establish a structured monitoring mechanism for students demonstrating early signs of academic difficulties and delayed study progress. WT2: Strengthen coordination among academic advisors, study programs, and faculties to provide targeted support for students facing academic and external socioeconomic challenges.

Source: Developed by the researchers based on the integration of predictive analytics results, the SWOT analysis, and the TOWS Matrix.

Among the proposed alternatives, the development of an integrated early warning system represents the most directly data-driven strategy because it builds upon the predictive capability demonstrated by the XGBoost model. Academic

performance indicators, particularly Cumulative GPA, semester GPA, and completed credit hours, can be continuously monitored to identify students who require earlier intervention. However, the results also demonstrate that a single institutional

intervention is unlikely to be equally effective across all academic units. Faculty and Study Program emerged as highly important structural predictors, indicating substantial heterogeneity in graduation patterns. Accordingly, the implementation of intervention strategies should be differentiated according to the characteristics and graduation patterns of individual faculties and study programs.

The findings further suggest that intervention should begin early in the student lifecycle. Academic achievement gaps during Semesters 1 to 4 were identified as an important internal weakness, indicating that waiting until the final stage of study to intervene may be less effective. Early academic monitoring, targeted mentoring, and structured support mechanisms may therefore provide greater opportunities to prevent delayed graduation before students enter the final stages of their studies. Overall, the integration of predictive analytics into SWOT and TOWS analysis demonstrates how machine learning can extend beyond prediction and support strategic decision-making. Rather than treating predictive models solely as technical tools for classification, the findings show that feature importance can provide an empirical basis for identifying institutional priorities and translating analytical evidence into actionable managerial strategies.

5. CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This study demonstrates that the XGBoost model, using its default configuration, achieved the best performance among the six algorithms tested in predicting timely graduation among students at Universitas Padjadjaran, with an Accuracy of 75.14%, an AUC of 0.832, and an F1-Score of 76.82%. Faculty and Cumulative GPA were identified as the two most

dominant factors in the prediction, jointly accounting for 65.10% of the total feature importance score. The integration of these dominant factors into the SWOT analysis resulted in a WO strategy recommendation involving the development of an early warning system based on a predictive model integrated with the academic information system and prioritized for faculties with the lowest timely graduation rates. These findings demonstrate that integrating predictive analytics with strategic management can generate more targeted and evidence-based policy recommendations than conventional descriptive approaches.

5.2 Limitations and Future Research

This study has several limitations. First, the research design is cross-sectional and correlational; therefore, the feature importance scores reflect predictive contributions rather than causal relationships. Second, the findings are derived from data obtained from a single university, and their generalizability to other institutions requires further examination. Third, the weighting of external factors in the EFAS Matrix remains based on qualitative estimates. Future research is recommended to: (1) conduct external validation of the model using student cohorts or institutions beyond the present study context; (2) develop dynamic predictive models that are updated each semester rather than relying on static models; and (3) explore quantitative approaches for weighting external factors, such as structured expert judgment using methods including the Analytic Hierarchy Process (AHP) or Delphi technique, to strengthen the objectivity of the EFAS Matrix.

5.3 Managerial Implications

From a managerial perspective, university- and faculty-level decision-makers are recommended to integrate the results of this predictive analysis into data-driven academic decision-making. Priority should be given to lower-tier

faculties that demonstrate the lowest timely graduation rates while also showing the strongest association between student engagement and graduation outcomes. The development of an XGBoost-based early warning system integrated with the Integrated Academic Information System (SIAT) is

recommended. Such a system should include an individual student-level risk score for delayed graduation generated at the end of each semester, a notification mechanism for academic advisors, and structured follow-up procedures for students identified as having a high risk of delayed graduation.

AUTHOR CONTRIBUTIONS

Intan Siti Amaliyah: Conceptualization, Methodology, Data Curation, Formal Analysis, Investigation, Writing – Original Draft.

Samidi: Supervision, Validation, Writing – Review and Editing.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no financial or personal conflicts of interest that could have influenced the objectivity of this research.

FUNDING STATEMENT

This research received no specific grant or funding from any public, commercial, or not-for-profit funding agency.

DATA AVAILABILITY STATEMENT

The student academic data used in this study are confidential and cannot be publicly disclosed due to the institutional privacy policies of the university from which the data were obtained. The data may be made available by the corresponding author upon reasonable request and subject to permission from the relevant institution.

ETHICS STATEMENT AND DATA USE PERMISSION

This study used secondary data consisting of anonymized academic and student records obtained with permission from the relevant institution. The dataset contained no personally identifiable information, such as student names or student identification numbers. As the study did not involve direct interaction with human participants and used data that had been fully anonymized before analysis, formal ethical committee review was not required under the institutional policies applicable at the time the research was conducted.

GENERATIVE ARTIFICIAL INTELLIGENCE USE STATEMENT

During the preparation of this manuscript, the authors used a generative artificial intelligence tool (Claude, Anthropic) to assist in drafting the initial narrative and refining the language based on research findings generated and analyzed by the authors. Following the use of this tool, the authors reviewed, edited, and validated the entire manuscript and remain fully responsible for the content of the published article.

REFERENCES

- [1] Kementerian Pendidikan dan Kebudayaan Republik Indonesia, *Keputusan Menteri Pendidikan dan Kebudayaan Nomor*

- 3/M/2021 tentang Indikator Kinerja Utama Perguruan Tinggi Negeri dan Lembaga Layanan Pendidikan Tinggi di Kementerian Pendidikan dan Kebudayaan. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2021.
- [2] E. Alyahyan and D. Düşteğör, "Predicting academic success in higher education: literature review and best practices," *Int. J. Educ. Technol. High. Educ.*, vol. 17, no. 1, p. 3, 2020, doi: <https://doi.org/10.1186/s41239-020-0177-7>.
 - [3] A. Namoun and A. Alshantiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Appl. Sci.*, vol. 11, no. 1, p. 237, 2020, doi: <https://doi.org/10.3390/app11010237>.
 - [4] F. R. David and F. R. David, *Strategic management: A competitive advantage approach*. Pearson, 2017.
 - [5] T. L. Wheelen, J. D. Hunger, A. N. Hoffman, and C. E. Bamford, *Strategic Management and Business Policy: Globalization, Innovation and Sustainability, 15th ed.* Pearson, 2018.
 - [6] R. Komaladewi, P. Usmanij, A. R. Amani, and F. E. Saputra, "Exploring social media marketing: A key driver of social commerce in Indonesia," *J. Int. Coun. Small Bus.*, pp. 1–18, 2025, doi: <https://doi.org/10.1080/26437015.2025.2558206>.
 - [7] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: <https://doi.org/10.1023/A:1010933404324>.
 - [8] H. Zhang, "The Optimality of Naive Bayes," in *Proceedings of the 17th International Florida Artificial Intelligence Research Society Conference (FLAIRS 2004)*, AAAI Press, 2004, pp. 562–567.
 - [9] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Ann. Stat.*, pp. 1189–1232, 2001, doi: <https://doi.org/10.1214/aos/1013203451>.
 - [10] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
 - [11] B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student'performance prediction using machine learning techniques," *Educ. Sci.*, vol. 11, no. 9, p. 552, 2021, doi: <https://doi.org/10.3390/educsci11090552>.
 - [12] M. A. S. Pawitra, H.-C. Hung, and H. Jati, "A Machine Learning Approach to Predicting On-Time Graduation in Indonesian Higher Education," *Elinvo (Electronics, Informatics, Vocat. Educ.)*, vol. 9, no. 2, pp. 294–308, 2024, doi: <https://doi.org/10.21831/elinvo.v9i2.77052>.
 - [13] R. Bakri, S. Alam, N. P. Astuti, and M. I. Bakhtiar, "Optimizing machine learning models for graduation on time prediction: A comparative study with resampling and hyperparameter tuning," *J. Online Inform.*, vol. 10, no. 2, pp. 270–285, 2025, doi: <https://doi.org/10.15575/join.v10i2.1590>.
 - [14] N. V. Darenoh, F. A. Bachtiar, and R. S. Perdana, "Prediction of On-time Student Graduation with Deep Learning Method," *J. ICT Res. Appl.*, vol. 18, no. 1, p. 1, 2024, doi: <https://doi.org/10.5614/itbj.ict.res.appl.2023.18.1.1>.
 - [15] A. Santoso, H. Retnawati, Kartianom, E. Apino, I. Rafi, and M. N. Rosyada, "Predicting time to graduation of Open University students: An educational data mining study," *Open Educ. Stud.*, vol. 6, no. 1, p. 20220220, 2024, doi: <https://doi.org/10.1515/edu-2022-0220>.
 - [16] L. R. Pelima, Y. Sukmana, and Y. Rosmansyah, "Predicting university student graduation using academic performance and machine learning: a systematic literature review," *IEEE Access*, vol. 12, pp. 23451–23465, 2024, doi: <https://doi.org/10.1109/ACCESS.2024.3361479>.
 - [17] P. Chapman *et al.*, "CRISP-DM 1.0: Step-by-step data mining guide," *SPSS inc*, vol. 9, no. 13, pp. 1–73, 2000.
 - [18] H. Weihrich, "The TOWS matrix – A tool for situational analysis," *Long Range Plann.*, vol. 15, no. 2, pp. 54–66, 1982, doi: [https://doi.org/10.1016/0024-6301\(82\)90120-0](https://doi.org/10.1016/0024-6301(82)90120-0).
 - [19] R. Nuraeni, M. Hardini, J. Parker, and M. G. Ilham, "Swot analysis of ai-based learning recommendation systems for student engagement," *Int. Trans. Artif. Intell.*, vol. 4, no. 1, pp. 1–12, 2025, doi: <https://doi.org/10.33050/italic.v4i1.894>.