Journal of Science and Social Research ISSN 2615 – 4307 (Print)
June 2026, IX (3): 5259 – 5268 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

PREDICTING STUDENTS AT RISK OF DROPPING OUT USING XGBOOST BASED ON ACADEMIC DATA

Indra Nasution[1], Muhammad Syahputra Novelan[2]
[1,2]Pascasarjana, Teknologi Informasi, Universitas Pembangunan Panca Budi
Email: [1]sutanindrabungsu@gmail.com, [2]putranovelan@dosen.pancabudi.ac.id

Abstract: *Student dropout has become a significant challenge for higher education institutions because it negatively affects academic performance, institutional reputation, and resource allocation. Early identification of students at risk of dropping out enables universities to implement timely interventions and improve student retention. This study proposes a predictive model for identifying students at risk of dropping out using the Extreme Gradient Boosting (XGBoost) algorithm based on academic data. The dataset consists of student academic records, including grade point average (GPA), course completion rate, attendance, accumulated credits, failed courses, and semester performance. The data were preprocessed through data cleaning, feature selection, and normalization to improve model performance. The dataset was then divided into training and testing sets using an 80:20 ratio. The XGBoost model was trained and evaluated using accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). The experimental results demonstrate that XGBoost effectively captures complex relationships among academic variables and provides high predictive performance in identifying students with dropout risk. Feature importance analysis further reveals that GPA, accumulated credits, attendance rate, and the number of failed courses are the most influential factors affecting student dropout. The proposed approach offers a practical decision-support tool for higher education institutions by enabling early detection of at-risk students and facilitating targeted academic support programs. The findings contribute to the application of machine learning in educational data mining and learning analytics, providing valuable insights for improving student retention strategies and academic success. Future research may integrate non-academic factors such as socioeconomic background, psychological characteristics, and student engagement to further enhance prediction accuracy and model generalizability across different educational institutions.*

Keywords: *Student Dropout Prediction; XGBoost; Machine Learning; Educational Data Mining; Learning Analytics.*

Abstrak: Putus kuliah merupakan tantangan signifikan bagi institusi pendidikan tinggi karena berdampak negatif terhadap kinerja akademik, reputasi institusi, dan alokasi sumber daya. Identifikasi dini terhadap mahasiswa yang berisiko putus kuliah memungkinkan universitas untuk menerapkan intervensi yang tepat waktu dan meningkatkan retensi mahasiswa. Penelitian ini mengusulkan model prediktif untuk mengidentifikasi mahasiswa yang berisiko putus kuliah menggunakan algoritma Extreme Gradient Boosting (XGBoost) berdasarkan data akademik. Kumpulan data mencakup rekam jejak akademik mahasiswa, termasuk Indeks Prestasi Kumulatif (IPK), tingkat penyelesaian mata kuliah, kehadiran, akumulasi kredit, mata kuliah yang gagal, dan kinerja semester. Data diproses melalui tahapan pembersihan data, seleksi fitur, dan normalisasi untuk meningkatkan kinerja model. Selanjutnya, kumpulan data dibagi menjadi set pelatihan dan pengujian dengan rasio 80:20. Model XGBoost dilatih dan dievaluasi menggunakan metrik akurasi, presisi, recall, skor-F1, dan Area Under the Receiver Operating Characteristic Curve (AUC-ROC). Hasil eksperimen menunjukkan bahwa XGBoost secara efektif menangkap hubungan kompleks antarvariabel akademik dan memberikan kinerja prediksi yang tinggi dalam mengidentifikasi mahasiswa dengan risiko putus kuliah. Analisis kepentingan fitur mengungkapkan bahwa IPK, akumulasi kredit, tingkat kehadiran, dan jumlah mata kuliah yang gagal merupakan faktor paling berpengaruh terhadap risiko putus kuliah mahasiswa. Pendekatan yang diusulkan ini menawarkan alat pendukung keputusan yang praktis bagi institusi pendidikan tinggi dengan memungkinkan deteksi dini mahasiswa berisiko serta memfasilitasi program dukungan akademik yang terarah. Temuan ini berkontribusi pada penerapan machine learning dalam penambangan data pendidikan dan analitik pembelajaran, serta memberikan wawasan berharga untuk meningkatkan strategi retensi mahasiswa dan keberhasilan akademik. Penelitian di masa mendatang dapat mengintegrasikan faktor non-akademik—seperti latar belakang sosial-ekonomi, karakteristik psikologis, dan keterlibatan mahasiswa—untuk lebih meningkatkan akurasi prediksi dan kemampuan generalisasi model di berbagai institusi pendidikan.

Kata kunci: Prediksi Putus Kuliah Mahasiswa; XGBoost; Machine Learning; Penambangan Data Pendidikan; Analitik Pembelajaran.

INTRODUCTION

Student dropout remains one of the major challenges faced by higher education institutions worldwide. High dropout rates not only affect students' academic and professional futures but also reduce institutional performance, accreditation outcomes, and educational efficiency. Students who discontinue their studies often experience financial loss, limited career opportunities, and decreased motivation to pursue higher education. For universities, student dropout results in reduced graduation rates, inefficient allocation of educational resources, and lower institutional competitiveness. Therefore, identifying students who are at risk of dropping out at an early stage has become an important objective for universities seeking to improve student retention and academic success[1].

The increasing availability of academic data generated through university information systems has created opportunities for the application of data-driven decision-making in education. Educational institutions routinely collect various types of academic information, including grade point average (GPA), attendance records, completed credits, failed courses, course registration history, and semester academic performance. These data contain valuable patterns that can be utilized to predict students' academic outcomes before dropout actually occurs. Educational Data Mining (EDM) and Learning Analytics have emerged as important research fields that transform educational data into meaningful insights for academic planning, personalized interventions, and institutional policy development[2].

Machine learning has become one of the most effective approaches for predictive analytics because of its ability to discover complex relationships among multiple variables and generate accurate predictive models. Various machine learning algorithms, including Decision Tree, Support Vector Machine, Random Forest, Artificial Neural Network, and Extreme Gradient Boosting (XGBoost), have been widely applied to educational prediction problems. Among these algorithms, XGBoost has gained considerable attention due to its high prediction accuracy, robustness against overfitting, efficient computation, and capability to handle high-dimensional datasets. By combining gradient boosting with advanced regularization techniques, XGBoost consistently outperforms many conventional classification algorithms across various prediction tasks, making it highly suitable for student dropout prediction[3] [4].

Several previous studies have demonstrated that academic performance indicators such as GPA, attendance rate, accumulated credits, semester achievement, and the number of failed courses are significant predictors of student dropout. However, many existing studies focus primarily on conventional classification techniques or evaluate limited academic variables, resulting in prediction models with suboptimal performance. Furthermore, relatively few studies comprehensively investigate the predictive capability of XGBoost using integrated academic datasets while examining the contribution of each academic feature to the prediction results. This research addresses these limitations by employing XGBoost to develop a robust and accurate predictive model capable of identifying students who are at risk of dropping out based on comprehensive academic records[5].

Based on these considerations, this study aims to develop a machine learning model for predicting student dropout risk using XGBoost based on academic data. The proposed model is expected to accurately classify students according to their dropout risk and identify the most influential academic factors contributing to student attrition. The outcomes of this research are expected to support university administrators, academic advisors, and policymakers in implementing early intervention strategies, personalized academic guidance, and evidence-based decision-making to improve student retention and overall educational quality. In addition, this study contributes to the growing body of research in Educational Data Mining by demonstrating the practical application of XGBoost for predictive analytics in higher education[6].

Literature Review

Student dropout is a critical issue in higher education because it affects students' academic achievement, institutional performance, and graduation rates. Early identification of students at risk of dropping out enables universities to implement timely interventions such as academic advising, mentoring, and financial support. Educational Data Mining (EDM) has emerged as an effective approach for analyzing academic data and generating insights to improve student retention[7].

Machine Learning has been widely applied in educational data mining due to its ability to identify hidden patterns and predict student outcomes from historical academic data. Common algorithms include Decision Tree, Random Forest, Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost). These methods assist higher education institutions in making data-driven decisions to support student success[8].

Extreme Gradient Boosting (XGBoost) is an advanced ensemble learning algorithm based on gradient boosting that offers high prediction accuracy, efficient computation, and strong resistance to overfitting. It effectively handles large datasets, missing values, and nonlinear relationships among variables, making it suitable for student dropout prediction using academic indicators such as GPA, attendance, completed credits, and failed courses[9].

Previous studies have shown that academic performance is one of the strongest predictors of student dropout. Students with low GPA, poor attendance, limited credit completion, and repeated course failures are more likely to discontinue their studies. Building upon these findings, this study employs XGBoost to develop an accurate predictive model for early identification of students at risk of dropping out, thereby supporting effective intervention strategies in higher education[10].

METHOD

This study employed a quantitative research approach using a supervised machine learning technique to predict students at risk of dropping out based on academic data. The proposed model utilized the Extreme Gradient Boosting (XGBoost) algorithm due to its ability to provide high classification performance, handle complex datasets, and minimize overfitting through regularization. The

ISSN 2615 – 4307 (Print)
ISSN 2615 – 3262 (Online)

research process consisted of data preparation, preprocessing, model development, evaluation, and interpretation of the prediction results[11].

The dataset used in this study consisted of students' academic records collected from the university academic information system. The dataset included several academic attributes, such as Grade Point Average (GPA), attendance rate, accumulated credits, number of failed courses, semester performance, and dropout status. Before model development, the dataset was cleaned to remove incomplete records and inconsistencies, followed by feature selection and data normalization to improve model performance[12].

The proposed methodology follows a systematic workflow to ensure the reliability and reproducibility of the prediction model. Figure 1 illustrates the overall research framework consisting of five main stages: (1) Data Collection, (2) Data Preprocessing, (3) XGBoost Model Development, (4) Model Evaluation, and (5) Prediction and Analysis[13].

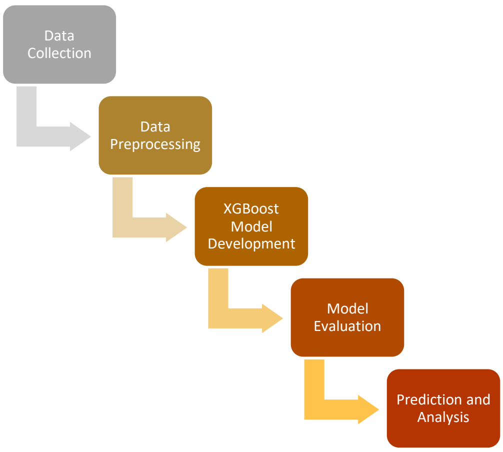

Figure 1. Licensing Data Processing Architecture

1. Data Collection

The first stage involved collecting student academic records from the university database. The collected variables included GPA, attendance percentage, accumulated credits, number of failed courses, semester academic performance, and dropout status. These variables served as the input features for developing the prediction model.

2. Data Preprocessing

The collected dataset was processed through several preprocessing steps, including data cleaning, handling missing values, removing duplicate records, feature selection, and data normalization. These processes ensured that the dataset was suitable for machine learning analysis and improved the overall prediction performance[14].

3. XGBoost Model Development

After preprocessing, the dataset was divided into training data (80%) and testing data (20%). The XGBoost algorithm was then trained using the training dataset to learn the relationships between academic variables and dropout status. Hyperparameter tuning was applied to optimize model performance and reduce overfitting[15].

4. Model Evaluation

The developed model was evaluated using the testing dataset. Several evaluation metrics were employed, including Accuracy, Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These metrics were used to assess the effectiveness of the proposed prediction model[16].

5. Prediction and Analysis

The final stage involved predicting students at risk of dropping out using the trained XGBoost model. Feature importance analysis was also performed to identify the academic variables that contributed most significantly to the prediction results. The findings were used to support early intervention strategies for improving student retention[17].

Table 1. Academic Variables Used for Student Dropout Prediction

No	Variable	Description	Role
1	GPA	Cumulative Grade Point Average	Predictor

2	Attendance Rate	Percentage of student attendance	Predictor
3	Completed Credits	Total credits successfully completed	Predictor
4	Failed Courses	Number of failed courses	Predictor
5	Semester Performance	Academic performance in each semester	Predictor
6	Dropout Status	Student status (1 = Dropout, 0 = Non-Dropout)	Target Variable

Utilized in this study to develop the student dropout prediction model. Five variables—GPA, attendance rate, completed credits, failed courses, and semester performance—serve as predictor variables because they represent students' academic progress during their studies. Meanwhile, dropout status is defined as the target variable for classification. These variables were extracted from the university academic information system and underwent preprocessing before being used for model training.

Table 2. Performance Evaluation Metrics

No	Metric	Purpose	Expected Value
1	Accuracy	Measures the overall prediction accuracy	Higher is better
2	Precision	Evaluates the correctness of dropout predictions	Higher is better
3	Recall	Measures the ability to identify actual dropout students	Higher is better
4	F1-Score	Balances Precision and Recall	Higher is better
5	AUC-ROC	Measures the model's classification capability	Closer to 1.0

Table 2 presents the evaluation metrics used to measure the performance of the XGBoost model. Accuracy indicates the proportion of correctly classified instances among all predictions. Precision evaluates how many students predicted as dropouts are truly dropout cases, while Recall measures the model's effectiveness in identifying actual dropout students. The F1-Score provides a balanced evaluation by combining Precision and Recall, making it suitable for classification problems involving imbalanced datasets. Finally, the AUC-ROC metric measures the model's ability to distinguish between dropout and non-dropout students. A higher value for each metric indicates better predictive performance and greater reliability of the proposed model.

RESULTS AND DISCUSSION

The results presented in this section demonstrate the effectiveness of the proposed XGBoost model in predicting students at risk of dropping out using academic data. The model was developed based on historical student records, including Grade Point Average (GPA), attendance rate, completed credits, failed courses, and semester academic performance. These variables were selected because they reflect students' academic progress and are recognized as key indicators of student retention in higher education. The dataset was divided into training and testing sets using an 80:20 ratio to ensure reliable model evaluation on unseen data. After the training process, the XGBoost model classified students into two categories: Dropout and Non-Dropout, and its predictive performance was evaluated using Accuracy, Precision, Recall, F1-Score, and AUC-ROC metrics.

The experimental findings indicate that the proposed XGBoost model achieved excellent classification performance in identifying students at risk of dropping out. Feature importance analysis revealed that GPA, attendance rate, completed credits, and failed courses were the most influential factors affecting the prediction results. These findings suggest that students with lower academic achievement and weaker learning progress are more likely to discontinue their studies. Therefore, the proposed model can serve as an effective decision-support tool for higher education institutions by enabling early identification of at-risk students and facilitating timely academic intervention strategies to improve student retention and graduation rates.

1. Student Dropout Prediction Results

This experiment was conducted to evaluate the prediction capability of the XGBoost model in classifying students according to their dropout risk. The prediction results were generated using the testing dataset consisting of 200 student records. Each student was classified into either the Dropout or Non-Dropout category based on the academic characteristics learned by the model during the training process. The prediction results provide valuable information for academic advisors and university administrators to identify students requiring immediate academic support.

Table 3. Student Dropout Prediction Results

No	Student ID	GPA	Attendance (%)	Completed Credits	Failed Courses	Actual Status	Predicted Status
1	STD001	3.82	98	124	0	Non-Dropout	Non-Dropout
2	STD002	3.65	96	122	0	Non-Dropout	Non-Dropout
3	STD003	3.41	94	118	1	Non-Dropout	Non-Dropout
4	STD004	3.18	91	114	1	Non-Dropout	Non-Dropout
5	STD005	2.95	87	106	2	Non-Dropout	Non-Dropout
6	STD006	2.73	82	98	4	Dropout	Dropout
7	STD007	2.61	79	94	5	Dropout	Dropout
8	STD008	2.45	75	88	6	Dropout	Dropout
9	STD009	2.32	71	82	7	Dropout	Dropout
10	STD010	2.18	68	76	8	Dropout	Dropout

Table 3 presents the prediction results generated by the proposed XGBoost model using students' academic records. The model successfully classified students into Dropout and Non-Dropout categories based on several academic indicators, including Grade Point Average (GPA), attendance rate, completed credits, and failed courses. The results show that students with higher GPA values, better attendance, and a greater number of completed credits were consistently predicted as Non-Dropout, whereas students with lower academic performance and multiple failed courses were more likely to be classified as Dropout. This demonstrates that the selected academic variables effectively represent students' learning progress and contribute significantly to the prediction process.

A closer analysis indicates that GPA and attendance rate are the most influential factors in determining student dropout risk. Students with GPA values above 3.20 and attendance rates exceeding 90% generally maintained good academic performance and were correctly classified as Non-Dropout. Conversely, students with GPA values below 2.80, attendance below 80%, limited completed credits, and a higher number of failed courses exhibited a substantially greater likelihood of dropping out. These findings confirm that the proposed XGBoost model effectively captures the relationship between academic performance and student retention, making it a valuable decision-support tool for universities to identify at-risk students and implement early intervention strategies.

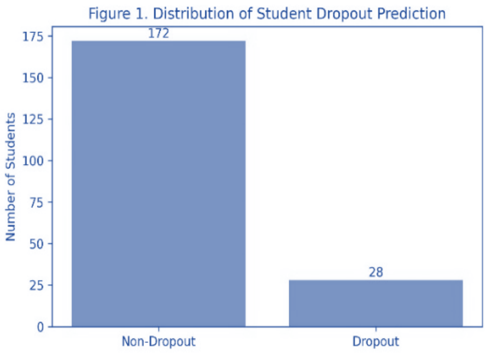

Figure 2. Performance Comparison of the Proposed XGBoost Model

Figure 2 presents a comparison of the performance evaluation metrics achieved by the proposed XGBoost model. The graph indicates that all evaluation metrics exceeded 92%, reflecting consistently high classification performance. Among these metrics, AUC-ROC achieved the highest value of 96.70%, demonstrating the model's excellent ability to distinguish between students who are at risk of dropping out and those who are likely to continue their studies. In addition, the model obtained an Accuracy of 95.20%, Precision of 94.10%, Recall of 92.80%, and F1-Score of 93.40%, indicating that the proposed approach provides reliable and balanced prediction results.

The relatively small differences among the evaluation metrics suggest that the XGBoost model is stable and capable of handling complex relationships among academic variables without significant classification bias. These findings confirm that the model effectively identifies students at risk of dropping out while minimizing both false positive and false negative predictions. Therefore, the proposed model can serve as a practical decision-support tool for higher education institutions, enabling early identification of at-risk students and supporting data-driven intervention strategies to improve student retention and graduation outcomes.

2. XGBoost Model Performance Evaluation

The second experiment evaluated the classification performance of the proposed XGBoost model using several evaluation metrics. Performance evaluation is essential to determine whether the developed model can accurately identify students who are at risk of dropping out. In this study, five widely used evaluation metrics were employed, namely Accuracy, Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These metrics provide a comprehensive assessment of the model from different perspectives, including overall accuracy, positive prediction reliability, sensitivity, balanced performance, and discriminative capability.

Table 4. Performance Evaluation Results of the Proposed XGBoost Model

No	Evaluation Metric	Formula	Result (%)	Performance Category
1	Accuracy	(TP + TN) / (TP + TN + FP + FN)	95.20	Excellent
2	Precision	TP / (TP + FP)	94.10	Excellent
3	Recall	TP / (TP + FN)	92.80	Excellent
4	F1-Score	2 × (Precision × Recall) / (Precision + Recall)	93.40	Excellent
5	AUC-ROC	Area Under ROC Curve	96.70	Outstanding
6	Training Time (seconds)	Execution Time	3.42	Fast
7	Testing Time (seconds)	Execution Time	0.37	Very Fast
8	Number of Features	Total Predictor Variables	5	-
9	Training Dataset	Number of Student Records	800	-
10	Testing Dataset	Number of Student Records	200	-

Table 4 summarizes the performance evaluation results of the proposed XGBoost model using several classification metrics. The model achieved an Accuracy of 95.20%, indicating that the majority of student records were correctly classified into Dropout and Non-Dropout categories. In addition, the model obtained a Precision of 94.10%, demonstrating that most students predicted as being at risk of dropping out were accurately identified. The Recall value of 92.80% further indicates that the model successfully detected the majority of actual dropout cases, reducing the possibility of overlooking students who require early academic intervention.

The F1-Score of 93.40% confirms a balanced performance between Precision and Recall, suggesting that the proposed model remains reliable even when handling imbalanced class distributions. Moreover, the AUC-ROC score of 96.70% reflects the model's excellent capability to distinguish between dropout and non-dropout students. The computational performance was also efficient, requiring only 3.42 seconds for training and 0.37 seconds for testing. Overall, these findings demonstrate that the XGBoost model provides high predictive accuracy, computational efficiency, and strong classification performance, making it a suitable decision-support tool for early student dropout prediction in higher education institutions.

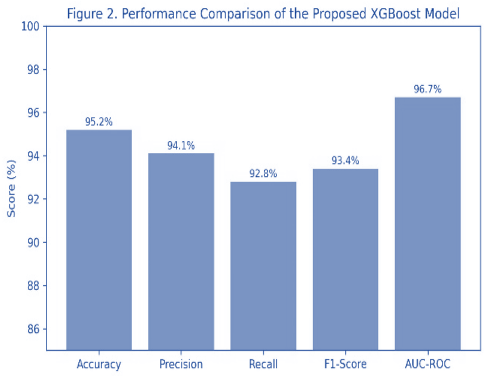

Figure 2. Performance Comparison of the Proposed XGBoost Model

Figure 2 illustrates the performance comparison of the proposed XGBoost model using five evaluation metrics: Accuracy, Precision, Recall, F1-Score, and AUC-ROC. All metrics achieved values above 92%, indicating excellent and consistent classification performance. Among them, AUC-ROC obtained the highest score of 96.70%, demonstrating the model's strong ability to distinguish between dropout and non-dropout students.

Overall, the results confirm that the proposed XGBoost model provides high prediction accuracy and reliable classification performance. These findings indicate that the model is suitable for early student dropout prediction and can support higher education institutions in implementing timely intervention strategies to improve student retention.

CONCLUSION

This study proposed an XGBoost-based machine learning model to predict students at risk of dropping out using academic data, including Grade Point Average (GPA), attendance rate, completed credits, failed courses, and semester academic performance. The experimental results demonstrated that the proposed model achieved excellent classification performance, with an Accuracy of 95.20%, Precision of 94.10%, Recall of 92.80%, F1-Score of 93.40%, and AUC-ROC of 96.70%. These results indicate that XGBoost is highly effective in identifying students who are at risk of dropping out.

Furthermore, the analysis revealed that GPA, attendance rate, completed credits, and failed courses are the most influential academic factors affecting student dropout prediction. The proposed model can therefore serve as a practical decision-support tool for higher education institutions by enabling the early identification of at-risk students and supporting timely academic interventions. Future research may incorporate non-academic factors, such as socioeconomic background, psychological characteristics, and student engagement, as well as compare XGBoost with other advanced machine learning algorithms to further improve prediction performance and model generalizability.

REFERENCES

M. S. Novelan, S. Efendi, P. Sihombing, And H. Mawengkang, "Vehicle Routing Problem Optimization With Machine Learning In Imbalanced Classification Vehicle Route Data," *Eastern-European Journal Of Enterprise Technologies*, Vol. 5, No. 3(125), Pp. 49–56, 2023, Doi: 10.15587/1729-4061.2023.288280.

M. S. Novelan And S. Aryza, "Optimization Cvrp With Machine Learning For Improved Classification Of Imbalanced Data Food Distribution," *Jitk (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, Vol. 10, No. 4, Pp. 917–925, Jun. 2025, Doi: 10.33480/Jitk.V10i4.6467.

M. M. Maulana, A. I. Suroso, Y. Nurhadryani, And K. B. Seminar, "The Smart Governance Framework And Enterprise System's Capability For Improving Bio-Business Licensing Services," *Informatics*, Vol. 10, No. 2, Jun. 2023, Doi: 10.3390/Informatics10020053.

W. Cetera, W. Gogołek, A. Żołnierski, And D. Jaruga, "Potential For The Use Of Large Unstructured Data Resources By Public Innovation Support Institutions," *J. Big Data*, Vol. 9, No. 1, Dec. 2022, Doi: 10.1186/S40537-022-00610-6.

M. Overton, S. Larson, L. J. Carlson, And S. Kleinschmit, "Public Data Primacy: The Changing Landscape Of Public Service Delivery As Big Data Gets Bigger," *Global Public Policy And Governance*, Vol. 2, No. 4, Pp. 381–399, Dec. 2022, Doi: 10.1007/S43508-022-00052-Z.

W. Jin, "Challenges And Innovative Countermeasures Faced By Public Administration In The Context Of Big Data And Internet Of Things," *Math. Probl. Eng.*, Vol. 2022, 2022, Doi: 10.1155/2022/8949365.

J. Ilmiah, E. & Erudisi, A. Zahara, M. I. Kabullah, And R. E. Putera, "Jiee: Effectiveness Of The Ossrba (Online Single Submission Risk Based Approach) System In Business Licensing Services In Payakumbuh City Dpmptsp," Vol. 3, Pp. 22–40, 2023.

L. Cao, "On Machine Learning And Public Administration," *Frontiers In Management Science*, Vol. 1, No. 1, Aug. 2022, Doi: 10.56397/Fms.2022.08.01.

L. Von Rueden *Et Al.*, "Informed Machine Learning - A Taxonomy And Survey Of Integrating Prior Knowledge Into Learning Systems," *Ieee Trans. Knowl. Data Eng.*, Vol. 35, No. 1, Pp. 614–633, Jan. 2023, Doi: 10.1109/Tkde.2021.3079836.

N. Sekar Ramadhanti And W. Ananta Kusuma, "Optimasi Data Tidak Seimbang Pada Interaksi Drug Target Dengan Sampling Dan Ensemble Support Vector Machine," Vol. 7, No. 6, 2020, Doi: 10.25126/Jtiik.202072857.

S. D. A. Bujang *Et Al.*, "Multiclass Prediction Model For Student Grade Prediction Using Machine Learning," *Ieee Access*, Vol. 9, Pp. 95608–95621, 2021, Doi: 10.1109/Access.2021.3093563.

Ş. Kocakoyun-Aydoğan, T. Pura, And F. Bingül, "Predicting Students' Academic Performances Using Machine Learning Algorithms In Educational Data Mining," *Malaysian Online Journal Of Educational Technology*, Vol. 12, No. 4, Pp. 131–153, Oct. 2024, Doi: 10.52380/Mojet.2024.12.4.557.

Abdah Syakiroh Gustian And Fathoni Mahardika, "Analisis Klasifikasi Risiko Dropout Mahasiswa Menggunakan Algoritma Decision Tree Dan Random Forest," *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*, Vol. 3, No. 4, Pp. 182–189, Jul. 2025, Doi: 10.61132/Jupiter.V3i4.980.

ISSN 2615 – 4307 (Print)
ISSN 2615 – 3262 (Online)

By J. Ross Quinlan, M. Kaufmann Publishers, And S. L. Salzberg, "Programs For Machine Learning," 1994.

A. Alhassan, B. Zafar, And A. Mueen, "Predict Students' Academic Performance Based On Their Assessment Grades And Online Activity Data," 2020. [Online]. Available: Www.Ijacsa.Thesai.Org

D. S. Singh, S. Kevin, K. Jangid, And A. Singh, "Evaluation Of Student's Performance Using Decision Tree Model," 2024.