JEMSI: Jurnal Ekonomi Manajemen Sistem Informasi
E-ISSN: 2686-5238
P-ISSN: 2686-4916
https://dinastirev.org/JEMSI dinasti.info@gmail.com +62 811 7404 455
DOI: https://doi.org/10.38035/jemsi.v7i2
https://creativecommons.org/licenses/by/4.0/
Received: 17 November 2025 | **Revised:** 30 November 2025 | **Accepted:** 19 Desember 2025

Perancangan *Business Intelligence* untuk Prediksi Risiko Drop Out Mahasiswa berdasarkan Analisis Data Historis

Daffa Imani Saputra[1], Darius Andana Haris[2]
[1]Universitas Tarumanagara, Jakarta, Indonesia, daffa.825220114@stu.untar.ac.id
[2]Universitas Tarumanagara, Jakarta, Indonesia, darius@fti.untar.ac.id

Corresponding Author: daffa.825220114@stu.untar.ac.id[1]

Abstract *This research aims to design a Business Intelligence (BI) system capable of predicting student dropout risks based on historical academic data. The system was developed using a data-driven approach through the stages of extract, transform, and load (ETL), followed by data processing and analysis using several classification algorithms, including Decision Tree, Random Forest, and Logistic Regression. The best model was evaluated using performance metrics such as accuracy, precision, and recall. The implementation utilized the Streamlit framework to build an interactive dashboard that visualizes analytical results across various interfaces, including the Analytical Dashboard, Individual Prediction, and Student Analysis menus. The results indicate that the designed BI system can display predictive information in real time and assist academic stakeholders in conducting early interventions for high-risk students.*

Keywords: *Business Intelligence, Dropout Prediction, Historical Data, Machine Learning, Analytical Dashboard.*

Abstrak: Penelitian ini bertujuan untuk merancang sistem Business Intelligence (BI) yang mampu memprediksi risiko drop out mahasiswa berdasarkan data akademik historis. Sistem dikembangkan dengan pendekatan data-driven melalui tahapan extract, transform, dan load (ETL), kemudian dilakukan pemrosesan dan analisis menggunakan beberapa algoritma klasifikasi seperti Decision Tree, Random Forest, dan Logistic Regression. Model terbaik dievaluasi menggunakan metrik akurasi, presisi, dan recall. Implementasi dilakukan menggunakan framework Streamlit untuk menyajikan dashboard interaktif yang memvisualisasikan hasil analisis risiko dalam berbagai tampilan seperti Dashboard Analitik, Prediksi Individu, dan Analisis Mahasiswa. Hasil penelitian menunjukkan bahwa sistem BI yang dirancang mampu menampilkan informasi prediksi risiko secara real-time dan membantu pihak kampus dalam melakukan intervensi dini terhadap mahasiswa berisiko tinggi.

Kata Kunci: Business Intelligence, Prediksi Drop Out, Data Historis, Machine Learning, Dashboard Analitik.

PENDAHULUAN

Fenomena mahasiswa *drop out* merupakan tantangan serius yang dihadapi perguruan tinggi di Indonesia. Kegagalan mahasiswa menyelesaikan studi tepat waktu menimbulkan dampak signifikan, baik bagi individu maupun institusi pendidikan, seperti kerugian finansial, menurunnya akreditasi, serta berkurangnya efisiensi sistem akademik. Untuk menghadapi hal tersebut, pendekatan prediktif berbasis data menjadi solusi strategis melalui penerapan *Business Intelligence* (BI), yang berperan penting dalam mendeteksi risiko akademik secara dini dan membantu pengambilan keputusan berbasis data. Dengan menggabungkan data historis mahasiswa seperti Indeks Prestasi Kumulatif (IPK), tingkat kehadiran, dan jumlah SKS yang ditempuh, BI memungkinkan pihak perguruan tinggi melakukan intervensi preventif sebelum mahasiswa mencapai kondisi berisiko tinggi terhadap *drop out*.

Data historis akademik terbukti memiliki nilai prediktif yang tinggi terhadap potensi mahasiswa *drop out*. Variabel seperti IPK, SKS lulus, dan tingkat kehadiran dapat digunakan untuk membangun model klasifikasi risiko akademik yang akurat. Namun, sebagian besar Sistem Informasi Akademik (SIA) di perguruan tinggi masih berfungsi sebatas alat administratif tanpa fitur analisis prediktif yang memadai. Padahal, dengan mengintegrasikan BI ke dalam SIA, data pasif dapat diubah menjadi informasi strategis yang memberikan *insight* komprehensif terhadap performa mahasiswa. Integrasi ini tidak hanya meningkatkan fungsi analitik kampus, tetapi juga mendukung tata kelola akademik berbasis bukti melalui visualisasi data pada *dashboard* interaktif.

Dalam penelitian ini, penerapan BI difokuskan pada analisis data akademik mahasiswa yang mencakup IPK, jumlah SKS yang telah ditempuh, tingkat kehadiran, dan status cuti studi. Data tersebut digunakan untuk membangun model klasifikasi risiko *drop out* dengan menerapkan algoritma *machine learning* seperti *Decision Tree*, *Support Vector Machine* (SVM), *Naïve Bayes*, *Logistic Regression*, *XGBoost*, *Gradient Boosting*, dan *Random Forest*. Hasil prediksi kemudian dievaluasi menggunakan metrik seperti akurasi, *precision*, *recall*, F1-score, dan ROC AUC untuk menentukan model dengan performa terbaik. Evaluasi tersebut bertujuan untuk memastikan sistem dapat memberikan hasil prediksi yang akurat dan relevan bagi kebutuhan akademik.

Penerapan BI dalam penelitian ini dibatasi pada penggunaan data sekunder dari satu institusi pendidikan tinggi, dengan fokus utama pada pengembangan *dashboard* interaktif yang menampilkan hasil analisis risiko akademik mahasiswa dalam tiga kategori, yaitu rendah, sedang, dan tinggi. Visualisasi ini berfungsi sebagai alat bantu analisis bagi pihak kampus dalam memantau dan mengevaluasi kondisi akademik mahasiswa secara real-time. Dengan demikian, sistem yang dikembangkan tidak hanya berperan sebagai alat analitik, tetapi juga sebagai *decision support system* yang mendukung strategi pencegahan *drop out* secara proaktif dan berbasis data.

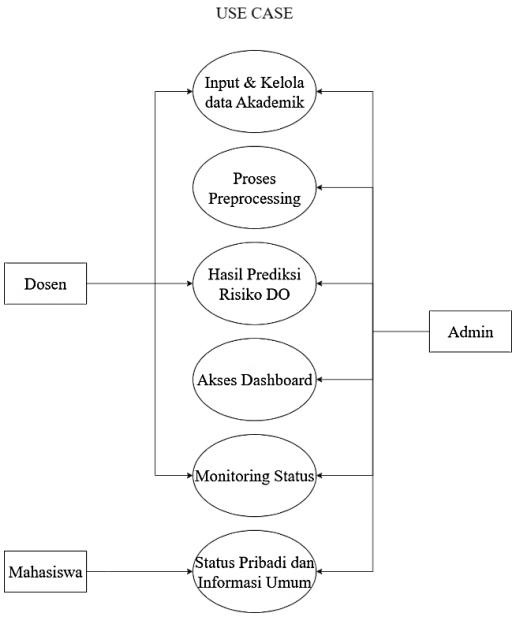
Tujuan dari penelitian ini adalah merancang dan mengimplementasikan sistem *Business Intelligence* yang terintegrasi dengan SIA untuk memprediksi risiko *drop out* mahasiswa berdasarkan data akademik historis. Manfaat penelitian ini meliputi peningkatan kemampuan institusi dalam melakukan deteksi dini terhadap mahasiswa berisiko, memperkuat peran SIA sebagai sistem analitik, serta menyediakan referensi empiris bagi pengembangan teknologi pendidikan berbasis data. Dengan demikian, penelitian ini diharapkan dapat berkontribusi terhadap transformasi digital pendidikan tinggi menuju tata kelola yang lebih adaptif, efisien, dan berorientasi pada keberhasilan akademik mahasiswa.

METODE

Program aplikasi yang dikembangkan dalam penelitian ini merupakan sistem *Business Intelligence* (BI) berbasis *machine learning* yang dirancang untuk memprediksi risiko *drop out* mahasiswa menggunakan data akademik historis dari Sistem Informasi Akademik (SIA). Aplikasi ini dibangun menggunakan bahasa pemrograman Python dengan framework Streamlit

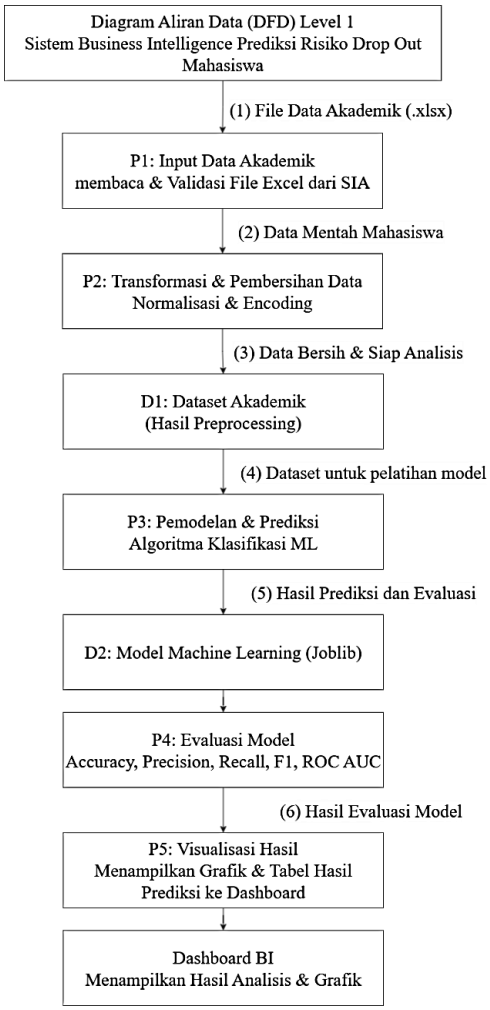
sebagai antarmuka utama dan didukung oleh library seperti Scikit-Learn, Plotly, Matplotlib, Openpyxl, dan Joblib. Tujuan utama dari sistem ini adalah menyediakan *dashboard* analitik interaktif yang mampu menampilkan hasil klasifikasi risiko mahasiswa secara akurat, cepat, dan mudah dipahami oleh pihak kampus.

Tahapan pengembangan dilakukan melalui proses yang sistematis, meliputi analisis kebutuhan sistem, perancangan arsitektur ETL (*Extract, Transform, Load*), penerapan algoritma klasifikasi seperti Naïve Bayes, Logistic Regression, SVM, Decision Tree, XGBoost, Gradient Boosting, dan Random Forest, serta pembuatan *dashboard* BI interaktif. Evaluasi sistem dilakukan menggunakan metrik seperti akurasi, presisi, *recall*, F1-score, dan ROC AUC untuk menilai performa model prediksi. Selain itu, sistem diuji melalui *unit testing* untuk memastikan setiap modul berjalan sesuai fungsi, mulai dari input data, klasifikasi, hingga visualisasi hasil.



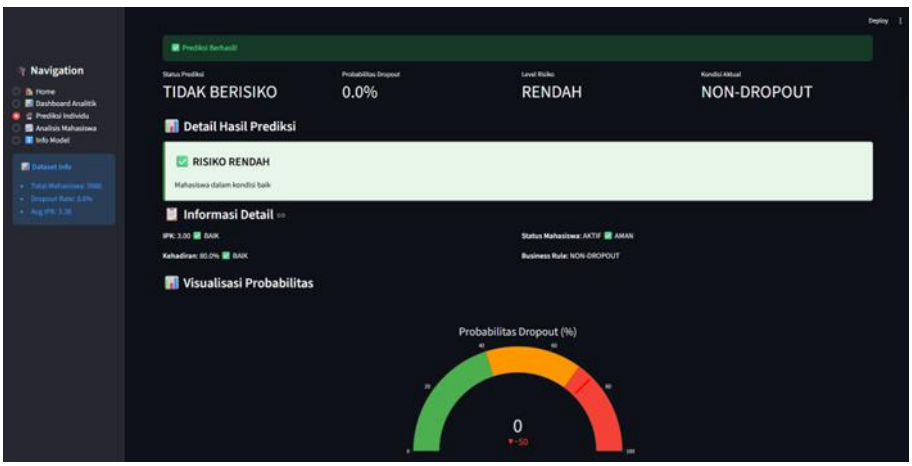
Gambar 1. *Use Case* Diagram Sistem BI Prediksi Risiko *Drop Out* Mahasiswa

Sistem ini dirancang dengan tiga jenis pengguna, yaitu administrator akademik, dosen pembimbing, dan mahasiswa, yang masing-masing memiliki hak akses berbeda. Admin bertugas mengunggah dan mengelola data, dosen dapat memantau mahasiswa bimbingannya berdasarkan hasil analisis risiko, sedangkan mahasiswa hanya dapat melihat status akademik pribadinya. Proses sistem dijelaskan melalui *Data Flow Diagram* (DFD) dan *Flowchart* ETL yang menggambarkan aliran data mulai dari unggahan file Excel, transformasi data, pemodelan, hingga visualisasi hasil analisis di dashboard Streamlit.



Gambar 2. Diagram Aliran Data (DFD) Sistem BI Prediksi Risiko Drop Out Mahasiswa

Antarmuka pengguna dirancang agar intuitif dan mudah digunakan, mencakup form *upload* data, pemilihan algoritma klasifikasi, tombol proses analisis, serta tampilan hasil berupa tabel, grafik interaktif, dan indikator performa model. Dashboard yang dihasilkan mampu menampilkan *confusion matrix*, ROC curve, serta distribusi kategori risiko mahasiswa, sehingga membantu pihak kampus dalam pengambilan keputusan berbasis data. Evaluasi sistem juga dilakukan berdasarkan *Key Performance Indicator* (KPI) untuk mengukur kecepatan sistem, kualitas data, tingkat adopsi pengguna, dan kepuasan terhadap kemudahan penggunaan aplikasi.



Gambar 3. Dashboard Prediksi dan Evaluasi Model

HASIL DAN PEMBAHASAN

Pada poin ini menjelaskan proses implementasi aplikasi *Business Intelligence* (BI) untuk prediksi risiko *dropout* mahasiswa, mencakup spesifikasi sistem, rancangan implementasi, prosedur penggunaan, serta hasil pengujian oleh pengguna (*User Acceptance Testing*). Implementasi dilakukan agar sistem dapat berjalan secara stabil, efisien, dan sesuai kebutuhan pengguna di lingkungan akademik. Dari sisi perangkat keras, aplikasi memerlukan prosesor minimal Intel Core i3, RAM 4 GB, dan penyimpanan SSD 128 GB agar proses pelatihan model dan eksekusi dashboard berjalan lancar. Sedangkan untuk perangkat lunak, digunakan Python 3.x beserta library seperti Streamlit, Scikit-Learn, Plotly, Matplotlib, dan Openpyxl. Personel yang terlibat dalam pengelolaan sistem terdiri dari administrator akademik sebagai pengelola data dan model, dosen pembimbing sebagai pengguna utama dalam pemantauan risiko mahasiswa, serta mahasiswa sebagai pengguna akhir yang dapat melihat hasil analisis individual mereka.

```
✅ Data gabungan berhasil disimpan di data\clean_dataset.xlsx
   No       NIM                    Nama     Status  SKS   IPK  Kehadiran Angkatan Prodi Semester
0   1  825160011  REFINDO ISHAK HAMONANGAN      LULUS  144  2.97       0.97     2020    SI   GANJIL
1   2  825160019            EDEN - HIZKIA  NON AKTIF   35  1.83       0.00     2020    SI   GANJIL
2   3  825160020          MUHAMMAD ALWALID      LULUS  146  3.36       0.86     2020    SI   GANJIL
3   4  825160022                  FALANDY       LULUS  146  3.15       0.87     2020    SI   GANJIL
4   5  825160031          WILLIAM HERMAWAN      LULUS  147  3.14       0.84     2020    SI   GANJIL
```

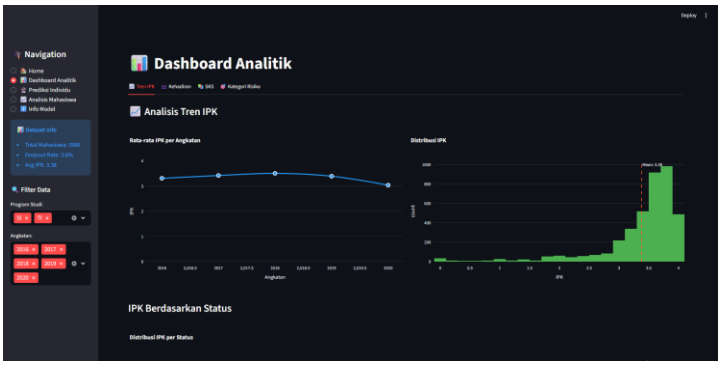
Gambar 4. Hasil Preprocessing Data

Implementasi sistem dilakukan dengan pendekatan *file-based*, di mana seluruh data dan model disimpan dalam format Excel (.xlsx) serta file model seperti pkl dan json, tanpa menggunakan basis data server seperti MySQL. Pendekatan ini dipilih agar aplikasi lebih ringan dan mudah dijalankan di berbagai perangkat. Struktur program dibangun secara modular dengan memisahkan fungsi utama seperti pemrosesan data (*utils*), pengaturan tampilan (*config*), dan halaman *dashboard* (*pages*). Framework Streamlit digunakan sebagai penggerak utama antarmuka sekaligus server lokal yang menampilkan visualisasi interaktif tanpa memerlukan backend tambahan. Seluruh komponen visual, termasuk grafik, tabel, dan indikator statistik, ditampilkan menggunakan kombinasi Streamlit dan Plotly.



Gambar 5. Struktur Folder dan Navigasi Aplikasi

Prosedur operasional aplikasi dilakukan secara sederhana. Pengguna cukup menjalankan perintah "*streamlit run app.py*" melalui terminal, kemudian sistem otomatis terbuka di *web browser*. Terdapat lima halaman utama dalam aplikasi, yaitu: (1) *Home*, menampilkan ringkasan data dan statistik mahasiswa; (2) *Dashboard Analitik*, menyajikan analisis mendalam terhadap IPK, kehadiran, SKS, serta kategori risiko dropout; (3) *Prediksi Individu*, menampilkan hasil prediksi risiko mahasiswa berdasarkan input data tertentu; (4) *Analisis Mahasiswa*, menampilkan seluruh hasil prediksi dan memungkinkan filter berdasarkan program studi dan angkatan; serta (5) *Informasi Model*, menjelaskan metode machine learning, performa model, dan logika bisnis yang digunakan dalam proses prediksi.



Gambar 6. Tampilan Menu Dashboard Analitik

Pemeliharaan sistem dilakukan secara berkala untuk menjaga kinerja aplikasi dan memastikan model prediksi tetap akurat seiring pembaruan data akademik mahasiswa. Prosedur yang dilakukan meliputi pemeriksaan fungsi aplikasi, pembaruan library Python, *backup* file penting seperti dataset dan model, serta *retraining* model ketika performa mulai menurun. Administrator akademik bertanggung jawab dalam memperbarui dataset setiap semester dan memastikan integrasi antar komponen sistem tetap berjalan baik. Dengan pemeliharaan rutin ini, aplikasi *Business Intelligence* untuk prediksi risiko *dropout* mahasiswa dapat terus berfungsi optimal, memberikan hasil analisis yang andal, serta mendukung pengambilan keputusan akademik berbasis data di lingkungan perguruan tinggi.

KESIMPULAN

Sistem *Business Intelligence* (BI) yang dikembangkan berhasil mengintegrasikan data akademik seperti IPK, SKS, dan kehadiran untuk memprediksi risiko *drop out* mahasiswa secara efisien melalui proses ETL dan penerapan algoritma seperti *Decision Tree*, *Random Forest*, serta *Logistic Regression*. Dashboard berbasis Streamlit juga mampu menyajikan hasil analisis secara interaktif dan mudah dipahami oleh pengguna.

Secara keseluruhan, sistem ini efektif mendukung upaya pencegahan mahasiswa berisiko tinggi dengan menyediakan informasi risiko secara *real-time*. Ke depan, pengembangan dapat difokuskan pada penambahan faktor non-akademik, integrasi otomatis dengan database kampus, serta peningkatan tampilan antarmuka agar sistem lebih responsif dan akurat.

REFERENSI

D. M. Córdova-Esparza *et al.*, “Predicting and Preventing School Dropout with Business Intelligence: Insights from a Systematic Review,” Apr. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/info16040326.

L. Aulck, N. Velagapudi, J. Blumenstock, and J. West, “Predicting Student Dropout in Higher Education,” Mar. 2016, [Online]. Available: http://arxiv.org/abs/1606.06364

A. Edenito, “Analisis Sentimen Masyarakat Pada Media Sosial Twitter Tehadap Saham Syariah Menggunakan Metode Naive Bayes Classifier Dan Support Vector Machine (Svm),” pp. 1–23, 2016.

V. P. Nugraha, “Prediksi Risiko Cuti Akademik Mahasiswa Berdasarkan Performa Awal dengan Decision Tree Predicting Student Academic Leave Risk Based on Early Performance Using Decision Tree Algorit ... Prediksi Risiko Cuti Akademik Mahasiswa Berdasarkan Performa Awal deng,” no. July, pp. 0–12, 2025, doi: 10.34010/jati.vxix.xxxxx.

Haditama, M. R. (2023). Analisis dan Pembuatan Dashboard Prediksi Kelulusan Mahasiswa Menggunakan Metode Random Forest, Naïve Bayes dan Support Vector Machine. Skripsi. Universitas Islam Negeri Syarif Hidayatullah Jakarta.

Kurniawati, "KLASIFIKASI DATA MAHASISWA LAMPAU MENGGUNAKAN METODE DECISION TREE DAN SUPPORT VECTOR MACHINE THESIS Oleh: KURNIAWATI NIM. 200605220016 PROGRAM STUDI MAGISTER INFORMATIKA FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG 2," 2024.

T. A. Marzuqi, E. Kristiani, and Marcel, "Prediksi Mahasiswa Drop-Out Di Universitas XYZ," J. Teknol. Inf. dan Ilmu Komput., vol. 11, no. 6, pp. 1345–1350, Dec. 2024, doi: 10.25126/jtiik.2024118689.

Z. S. Bischof, F. E. Bustamante, and N. Feamster, "Characterizing and Improving the Reliability of Broadband Internet Access," Sep. 2017, [Online]. Available: http://arxiv.org/abs/1709.09349

Y. Yang, Z. Wen, J. Cao, J. Shen, H. Yin, and X. Zhou, "EPARS: Early Prediction of At-risk Students with Online and Offline Learning Behaviors," Jun. 2020, [Online]. Available: http://arxiv.org/abs/2006.03857

A. Menolli, F. Horita, J. J. L. Dias, and R. Coelho, "BI-based methodology for analyzing higher education: A case study of dropout phenomenon in information systems courses," in ACM International Conference Proceeding Series, Association for Computing Machinery, Nov. 2020. doi: 10.1145/3411564.3411636.

S. Hafidzul Haq et al., "Peran Manajemen Sekuriti Dalam Meningkatkan Kesadaran Keamanan Data Mahasiswa Pada Sistem Informasi Akademik Ubhara Jaya," Orbit J. Ilmu Multidisiplin Nusant., vol. 1, no. 2, pp. 35–50, 2025, doi: 10.63217/orbit.v1i2.77.

A. S. Puspaningrum, "Pengukuran Kesesuaian Fungsional Dengan Pendekatan Berorientasi Tujuan Pada Sistem Informasi Akademik (SIA) Berdasarkan Model Kualitas ISO/IEC 25010," Tesis, no. Institut Teknologi Sepuluh Nopember, 2017.

P. A. Sunarya, M. Asri, N. Azizah, and C. P. Lim, "Evaluation of Educational Information Systems for Data and Decision Management," J. MENTARI Manajemen, Pendidik. dan Teknol. Inf., vol. 3, no. 2, pp. 118–126, 2025, doi: 10.33050/mentari.v3i2.738.

A. P. Anugraeni, W. Purnomo, and N. Y. Setiawan, "Pembangunan Dashboard Business Intelligence Visualisasi Data Akademik Pada MTsN 3 Malang," J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 1, no. 1, pp. 1–10, 2017, [Online]. Available: http://j-ptiik.ub.ac.id

H. F. Ramadhan, A. Fauzi, C. N. Rupelu, D. P. Aprillia, N. D. Anjani, and Halimatusadiah, "Pengaruh Business Intelligence Terhadap Perusahaan Dalam Pengambilan Keputusan : Business Intelligence , Arsitektur Bi Dan Data Warehouse (Kajian Studi Business Intelligence)," JEMSI (Jurnal Ekon. Manaj. Sist. Informasi), vol. 3, no. 6, pp. 639–644, 2022, doi: 10.31933/jemsi.v3i6.

R. Nabilatulrahmah et al., "Analisis Risiko Akademik Siswa Dengan Metode K-Means: Studi Kasus Smpn 10 Palembang," JATI (Jurnal Mhs. Tek. Inform., vol. 9, no. 4, pp. 6766–6773, 2025, doi: 10.36040/jati.v9i4.14181.

H. Abijono, P. Santoso, and N. L. Anggreini, "Algoritma Supervised Learning Dan Unsupervised Learning Dalam Pengolahan Data," J. Teknol. Terap. G-Tech, vol. 4, no. 2, pp. 315–318, 2021, doi: 10.33379/gtech.v4i2.635.

D. M. A. M. Sanjaya, A. A. I. I. Paramitha, and N. W. Utami, "Penerapan Data Mining untuk Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Algoritma C4.5: Studi Kasus STMIK Primakara," J. Ilm. Ilmu Terap. Univ. Jambi, vol. 6, no. 1, pp. 84–97, 2022, doi: 10.22437/jiituj.v6i1.19600.

D. A. Sa'adah, "Tren Integrasi Kecerdasan Buatan dalam Sistem Informasi Manajemen Pendidikan: Peluang dan Tantangan," vol. 1, no. 1, pp. 1–7, 2025.

A. W. Abdullah and A. Muhid, "Social Support, Academic Satisfaction, and Student Drop Out Tendency/ Dukungan Sosial, Academic Satisfaction, dan Kecenderungan Drop Out

pada Mahasiswa," Psikoislamika J. Psikol. dan Psikol. Islam, vol. 18, no. 1, pp. 174–187, 2021, doi: 10.18860/psikoislamika.v18i1.11546.

S. Linawati, S. Nurdiani, K. Handayani, and L. Latifah, "Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Random Forest Dan C4.5," J. Khatulistiwa Inform., vol. 8, no. 1, pp. 47–52, 2020, doi: 10.31294/jki.v8i1.7827.

S. D. Puspa, J. Riyono, C. E. Pujiastuti, and F. Puspitasari, "Pelatihan Pembuatan Dashboard dan Visualisasi Data dengan Looker Studio untuk Penguatan Literasi Data dan Digital," J. Pengabdi. Masy. dan Apl. Teknol., vol. 3, no. 2, pp. 78–85, 2024, doi: 10.31284/j.adipati.2024.v3i2.6272.

A. A. Jabar, R. F. Wijaya, and S. Wahyuni, "Dashboard Visualisasi Data Kecerdasan Bisnis Untuk Mendukung Pengambilan Keputusan Bisnis Pada Pt Bmpt," Djtechno J. Teknol. Inf., vol. 6, no. 1, pp. 75–89, 2025, doi: 10.46576/djtechno.v6i1.6038.

D. Khoirunnisa, "Sentimen Analisis terhadap penggunaan ChatGPT menggunakan metode Support Vector Machine (Studi kasus : Opini penggunaan ChatGPT pada Twitter)," 2024.