

Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma Support Vector Machine Dan Deep Learning

Data Mining Application To Generate Graduation Patterns Of Informatics Engineering Study Program Students With Support Vector Machine And Deep Learning Algorithms

Christanty Febriani Harimu¹, Vivi Peggie Rantung², Gladly C. Rorimpandey³

^{1,2,3} Teknik Informatika, Fakultas Teknik, Universitas Negeri Manado

Article Info	ABSTRAK
Article history: Received: Oct 9, 2024 Revised: Nov 16, 2024 Accepted: Nov 28, 2024	Mahasiswa adalah salah satu aspek terpenting ketika mengevaluasi keberhasilan pendirian program studi di Universitas. Kelulusan mahasiswa penting karena ini mempengaruhi evaluasi negara dan mempengaruhi sertifikasi kurikulum. Penelitian ini bertujuan untuk menerapkan teknik data mining dengan algoritma Support Vector Machine dan Deep Learning untuk menghasilkan prediksi pola kelulusan mahasiswa di prodi Teknik Informatika. Metode penelitian ini yaitu menggunakan algoritma Support Vector Machine dan Deep Learning yaitu melibatkan tahap perencanaan, pengumpulan data, pemrosesan data atau implementasi data dan pengujian. Perencanaan yaitu menentukan faktor-faktor yang mempengaruhi kelulusan mahasiswa. Selanjutnya dilakukan pengumpulan data dari yang berasal dari sistem informasi akademik universitas. Pemrosesan data melibatkan evaluasi model yang dilakukan untuk memastikan akurasi dan efektivitas model dalam memprediksi pola kelulusan, menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score kemudian diimplementasikan ke dalam aplikasi data mining yang dikembangkan. Pengujian menggunakan blacbox testing untuk memastikan tidak ada bug dan aplikasi berjalan sesuai yang diharapkan. Hasil dari penelitian ini yaitu mampu mengidentifikasi pola dan tren yang tidak terlihat secara langsung, memberikan prediksi yang lebih akurat mengenai kemungkinan kelulusan mahasiswa.
Kata kunci Data Mining, Deep Learning, Kelulusan Mahasiswa, Support Vector Machine	
Keywords Data Mining, Deep Learning, Student Graduation, Support Vector Machine	ABSTRACT <i>Students are one of the most important aspects when evaluating the success of establishing a study program at a University. Student graduation is important because it affects state evaluation and affects curriculum certification. This research aims to apply data mining techniques with Support Vector Machine and Deep Learning algorithms to produce predictions of student graduation patterns in the Informatics Engineering study program. The method of this research is to use the Support Vector Machine and Deep Learning algorithms, which involve the planning, data collection, data processing or data implementation and testing stages.</i>

Planning is to determine the factors that affect student graduation. Furthermore, data collection is carried out from the university's academic information system. Data processing involves model evaluation to ensure the accuracy and effectiveness of the model in predicting graduation patterns, using evaluation metrics such as accuracy, precision, recall, and F1-score and then implemented into the developed data mining application. Testing uses blacbox testing to ensure there are no bugs and the application runs as expected. The results of this research are able to identify patterns and trends that are not directly visible, providing more accurate predictions regarding the possibility of student graduation.

Corresponding Author:

Vivi Peggie Rantung,
Program Studi Teknik Informatika,
Universitas Negeri Manado,
Jl. Kampus Unima, Tonsaru, Kec. Tondano Selatan, Kab. Minahasa, Sulawesi Utara, 95618.
Email: vivirantung@unima.ac.id

PENDAHULUAN

Perkembangan teknologi saat ini semakin berkembang pesat, khususnya di bidang teknologi informasi. Implementasi teknologi informasi dan komunikasi (TIK) dalam organisasi merupakan hal yang wajib dilakukan di era sekarang ini, untuk memenangkan persaingan (Rantung et al., 2019). Bukan hanya dikalangan masyarakat, teknologi informasi juga diperlukan dalam pendidikan agar dapat memperlancar aktivitas dalam kegiatan pendidikan. Pendidikan adalah proses untuk melahirkan generasi berkualitas tinggi berikutnya (Loho et al., 2022).

Mahasiswa adalah salah satu aspek terpenting ketika mengevaluasi keberhasilan pendirian program studi di Universitas. Kelulusan mahasiswa penting karena ini mempengaruhi evaluasi negara dan mempengaruhi sertifikasi kurikulum. Masa studi seorang mahasiswa sangat penting bagi mahasiswa itu sendiri dan tentunya bagi program studi, guna menentukan langkah-langkah yang diperlukan agar berhasil menyelesaikan jalur pembelajaran. Dalam hal ini Prodi Teknik Informatika memiliki kesulitan untuk mengetahui pola kelulusan mahasiswa untuk itu dibutuhkan sistem yang bisa memprediksi pola kelulusan mahasiswa. Memprediksi kelulusan mahasiswa adalah suatu kegiatan untuk menentukan suatu kondisi dimasa depan berdasarkan data yang ada (Bangun et al., 2022).

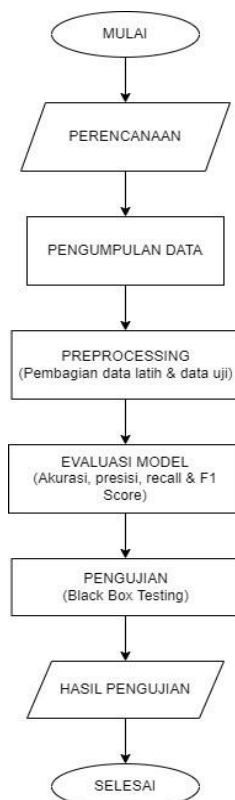
Setiap Universitas maupun Program studi memiliki *database* yang menyimpan data akademik dan biografi mahasiswa. Salah satu metode yang dapat digunakan untuk menganalisa *database* tersebut adalah metode *data mining* (Setiyani, 2020). *Data mining* adalah proses eksplorasi dan analisis data yang digunakan untuk menemukan pola, tren, dan hubungan yang bermanfaat dari kumpulan data yang besar.

Penelitian ini mengembangkan aplikasi *data mining* yang bertujuan untuk menghasilkan pola kelulusan mahasiswa pada program studi Teknik Informatika dengan menggunakan algoritma *Support Vector Machine* dan *Deep Learning*. Aplikasi ini dirancang untuk menganalisis data akademik mahasiswa, seperti nilai, jumlah kredit yang diambil, dan berbagai faktor lainnya yang berpotensi mempengaruhi kelulusan.

METODE PENELITIAN

Algoritma *Support Vector Machine* (SVM) adalah algoritma pembelajaran mesin yang digunakan untuk tugas klasifikasi dan regresi. Tujuannya adalah untuk menemukan *hyperplane* yang memisahkan dua kelas data dengan margin maksimum. *Hyperplane* ini dipilih sedemikian rupa sehingga jarak antara *hyperplane* dan titik terdekat dari setiap kelas (disebut vektor dukungan) adalah maksimum. SVM mencoba untuk memaksimalkan margin ini untuk meningkatkan kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya. SVM dapat bekerja dengan baik dalam dimensi rendah maupun tinggi, dan sering digunakan dalam berbagai bidang seperti klasifikasi gambar, pengenalan teks dan bioinformatika.

Deep Learning merupakan sebuah teknologi yang digadang-gadang menjadi salah satu teknologi terpenting setelah internet dan sering digambarkan sebagai teknologi yang akan menggantikan posisi manusia di lapangan pekerjaan (Nugroho et al. 2020). Dalam konteks deep learning, *Multi-Layer Perceptron* (MLP) adalah jenis jaringan saraf tiruan yang terdiri dari beberapa lapisan *neuron*, termasuk lapisan *input*, satu atau lebih lapisan tersembunyi, dan lapisan *output* (Lubis and Kustanto 2020). Pada MLP, data *input* diterima oleh lapisan *input*, yang kemudian diproses melalui lapisan-lapisan tersembunyi sebelum menghasilkan *output*. Setiap *neuron* dalam lapisan tersembunyi melakukan kombinasi linier dari *input* yang diterima dan menerapkan fungsi aktivasi *non-linear* seperti *ReLU*, *sigmoid*, atau *tanh*. Proses ini, yang dikenal sebagai *forward propagation*, memungkinkan MLP memodelkan hubungan *non-linear* antara *input* dan *output*. Setelah *forward propagation*, *output* jaringan dibandingkan dengan nilai target untuk menghitung *error* atau *loss* menggunakan fungsi loss seperti *Mean Squared Error* (MSE) atau *Categorical Cross-Entropy*. Meskipun sering digantikan oleh arsitektur jaringan saraf yang lebih kompleks seperti *Convolutional Neural Networks* (CNNs) dan *Recurrent Neural Networks* (RNNs) dalam tugas-tugas tertentu, MLP tetap penting dalam pemahaman dasar *deep learning*. Berikut adalah penjelasan tentang tahapan-tahapan yang terdapat pada penelitian ini :



Gambar 1. Tahapan Penelitian

Perencanaan, pada fase ini adalah identifikasi tujuan dan kebutuhan proyek, yaitu menentukan pola kelulusan mahasiswa dan faktor-faktor yang mempengaruhi kelulusan (Nur 2022). Data yang telah dikumpulkan kemudian melalui proses pembersihan data untuk memastikan kualitas dan konsistensi, termasuk penanganan data yang hilang atau tidak valid.

Pengumpulan data yaitu pada tahapan ini, data yang dikumpulkan mencakup informasi akademik mahasiswa, seperti nilai mata kuliah, dan Indeks Prestasi Kumulatif (IPK) dari semester 1 sampai semester 8, catatan kehadiran, serta data demografis termasuk usia, jenis kelamin, dan latar belakang pendidikan. Sumber data utama berasal dari sistem informasi akademik universitas, yang menyediakan akses ke rekam jejak akademik mahasiswa selama periode studi mereka.

Tahap berikutnya adalah pemrosesan data, di mana data akan diubah menjadi format yang sesuai untuk analisis dengan algoritma SVM dan DL. Pelatihan model dilakukan dengan membagi data menjadi data latih dan data uji untuk menguji performa model (Sirait et al. 2021). Evaluasi model dilakukan untuk memastikan akurasi dan efektivitas model dalam memprediksi pola kelulusan, menggunakan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Tahap selanjutnya adalah implementasi model ke dalam aplikasi *data mining* yang dikembangkan. Proses ini melibatkan integrasi model dengan sistem yang ada, serta pengembangan antarmuka pengguna yang intuitif untuk memungkinkan pengguna mengakses dan memanfaatkan hasil prediksi.

Pengujian, metode yang digunakan dalam melakukan pengujian adalah *Black-Box Testing* (uji coba fungsionalitas sebuah aplikasi atau program yang sedang dikembangkan) dengan melakukan pengujian terhadap masukan dan keluaran yang dihasilkan sistem apakah sudah berjalan sesuai dengan fungsinya masing-masing. Dari hasil evaluasi tersebut akan ditarik kesimpulan berdasarkan hasil pengujian (*black box testing*) terhadap sistem yang dibuat (Munaiseche et al., 2019).

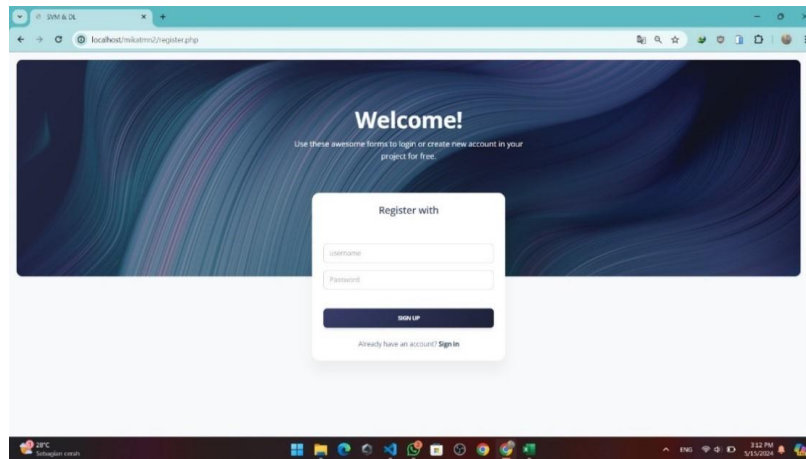
HASIL DAN PEMBAHASAN

Perencanaan

Langkah awal dalam penelitian ini adalah mengidentifikasi variabel-variabel penting yang mempengaruhi kelulusan mahasiswa, seperti nilai mata kuliah, dan Indeks Prestasi Kumulatif (IPK) dari semester 1 sampai semester 8. Data ini kemudian dikumpulkan dari database prodi teknik informatika, yang melibatkan lebih dari 1000 mahasiswa selama periode lima tahun terakhir. Selanjutnya, data yang telah dikumpulkan melalui proses pembersihan untuk mengatasi masalah data yang hilang dan inkonsistensi. Data tersebut kemudian diproses lebih lanjut melalui teknik pra-pemrosesan seperti normalisasi dan transformasi untuk memastikan keseragaman dan kesiapan data untuk dianalisis.

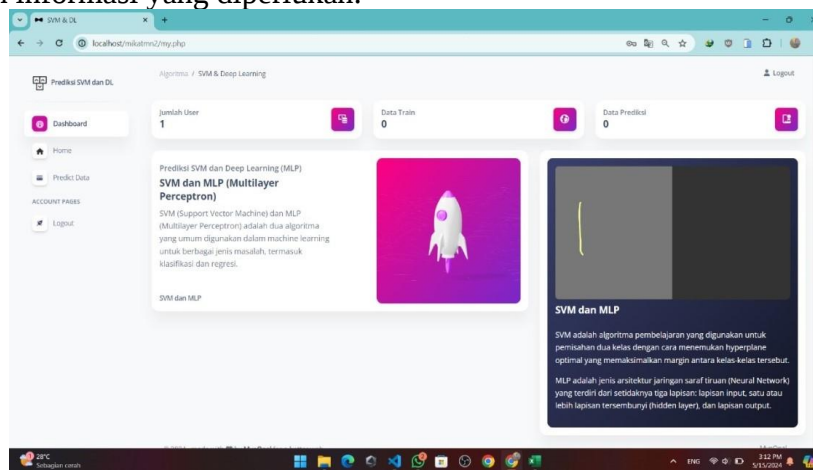
Pengumpulan Data

Pada tahap ini, data dikumpulkan dari berbagai sumber yang relevan, termasuk database akademik prodi teknik informatika yang berisi catatan nilai mata kuliah, dan Indeks Prestasi Kumulatif (IPK) dari semester 1 sampai semester 8. Data yang didapat dari database yaitu 1046 data mahasiswa dari angkatan 2016 sampai 2022. Proses ini melibatkan pengumpulan data historis dari mahasiswa yang telah lulus maupun yang masih aktif dalam kurun waktu lima tahun terakhir.



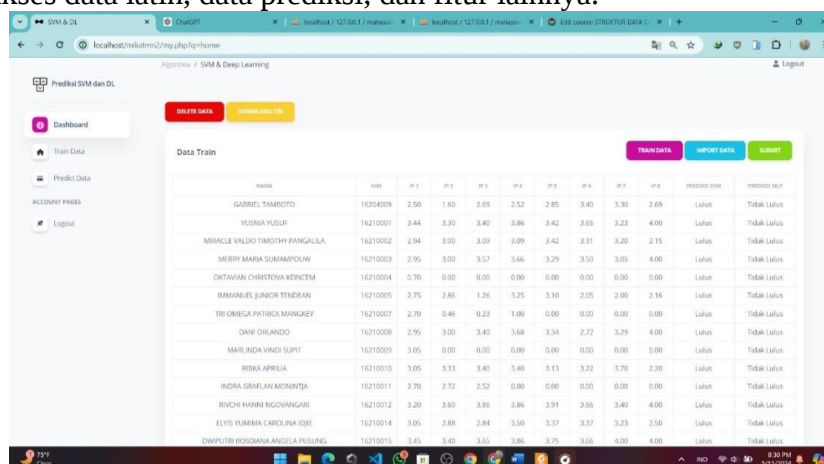
Gambar 5. Halaman Register

Menampilkan halaman registrasi di mana pengguna baru dapat membuat akun dengan memasukkan informasi yang diperlukan.



Gambar 6. Halaman Dashboard

Tampilan halaman dashboard aplikasi setelah pengguna berhasil login. Disini pengguna dapat mengakses data latih, data prediksi, dan fitur lainnya.



Gambar 7. Halaman Train Data

Halaman ini memungkinkan pengguna untuk mengupload data, mengkonfigurasi parameter pelatihan, dan memulai proses pelatihan model.

NAMA	ID	IPK.1	IPK.2	IPK.3	IPK.4	IPK.5	IPK.6	IPK.7	IPK.8	PREDIKSI DLM	PREDIKSI DL
GABRIEL TAMBOTO	16204009	2.50	1.60	2.69	2.52	2.85	3.40	3.30	2.69	Tidak Lulus	Tidak Lulus
YUSNA YUSUF	16210001	3.44	3.30	3.40	3.86	3.42	3.65	3.23	4.00	Lulus	Lulus
MIRACLE VALDO TIMOTHY PANGALILA	16210002	2.94	3.00	3.09	3.09	3.42	3.31	3.20	2.15	Lulus	Tidak Lulus
MERRY MARIA SUMAMPOLUW	16210003	2.95	3.00	3.57	3.66	3.29	3.50	3.05	4.00	Lulus	Lulus
OKTAVIAN CHRISTOVA KENCISEM	16210004	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	Tidak Lulus	Tidak Lulus
IRMANUEL JUNIOR TENDEAN	16210005	2.75	2.86	1.26	3.25	3.10	2.05	2.00	2.16	Tidak Lulus	Tidak Lulus
TBI OMS GA PATRICK MANGKEY	16210007	2.70	0.46	0.23	1.00	0.00	0.00	0.00	0.00	Tidak Lulus	Tidak Lulus
DANI ORLANDO	16210008	2.95	3.00	3.40	3.68	3.34	2.72	3.29	4.00	Lulus	Lulus
MARLINDA VINDI SURIT	16210009	3.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	Tidak Lulus	Tidak Lulus
REBEA APRILIA	16210010	3.09	3.33	3.40	3.40	3.13	3.22	3.79	2.20	Lulus	Tidak Lulus
INDRA GUNAWAN MONTONG	16210011	2.70	2.72	2.52	0.00	0.00	0.00	0.00	0.00	Tidak Lulus	Tidak Lulus
REYCH HANNI NGZANGARI	16210012	3.20	3.60	3.86	3.86	3.91	3.66	3.40	4.00	Lulus	Lulus
ELVIS YUMAMA CAROLINA EDE	16210014	3.05	2.88	2.84	3.50	3.37	3.37	3.23	2.50	Lulus	Tidak Lulus

Gambar 8. Halaman Prediksi Data

Halaman ini memberikan informasi tentang status prediksi, tingkat akurasi, dan rekomendasi untuk tindakan selanjutnya.

Pengujian

Metode *blackbox testing* ini dipilih karena memfokuskan pada fungsi sistem tanpa memperhatikan bagaimana sistem tersebut bekerja secara internal. Metode ini sangat cocok untuk digunakan pada tahap pengujian aplikasi ini karena membantu mengevaluasi apakah sistem berfungsi sesuai dengan yang diharapkan dan memastikan bahwa sistem bekerja dengan baik sebelum diluncurkan.

KESIMPULAN

Penelitian ini mengembangkan aplikasi data mining yang bertujuan untuk menghasilkan pola kelulusan mahasiswa pada program studi Teknik Informatika dengan menggunakan algoritma Support Vector Machine (SVM) dan Deep Learning (DL). Aplikasi ini dirancang untuk menganalisis data akademik mahasiswa, seperti nilai, jumlah kredit yang diambil, dan berbagai faktor lainnya yang berpotensi mempengaruhi kelulusan. Dengan memanfaatkan kemampuan SVM dan DL, aplikasi ini mampu mengidentifikasi pola dan tren yang tidak terlihat secara langsung, memberikan prediksi yang lebih akurat mengenai kemungkinan kelulusan mahasiswa. Hasil dari penelitian ini diharapkan dapat membantu pihak akademik dalam merancang strategi yang lebih efektif untuk meningkatkan tingkat kelulusan dan memberikan wawasan yang lebih mendalam tentang faktor-faktor kunci yang mempengaruhi kesuksesan akademik mahasiswa. Simpulan merupakan ringkasan dari hasil penelitian yang utama sesuai dengan tujuan penelitian yang telah ditetapkan. Apabila ada temuan baru yang berbeda dari hasil penelitian juga dituliskan pada bagian ini. Kesimpulan dituliskan dalam satu paragraf, bukan dalam bentuk poin-poin.

UCAPAN TERIMA KASIH

Saya ingin menyampaikan terima kasih atas semua dukungan yang luar biasa pada penelitian ini. Yang pertama saya ingin menyampaikan terima kasih kepada Koordinator dan Staff Program Studi Teknik Informatika yang telah memberikan izin untuk saya melakukan penelitian di Prodi dan memberikan akses untuk mengumpulkan data yang dibutuhkan untuk penelitian ini.

Yang kedua saya ingin mengucapkan terima kasih kepada Dosen Pembimbing Akademik atas arahan dan bimbingan yang diberikan dengan penuh kesabaran kepada saya dalam proses penelitian ini. Yang terakhir saya mengucapkan terima kasih kepada keluarga

bahkan teman-teman yang tetap setia memberikan dukungan serta doa sehingga penelitian ini berjalan dengan lancar.

DAFTAR PUSTAKA

- Akour, M., Sghaier, H. Al, & Qasem, O. Al. (2020). The effectiveness of using deep learning algorithms in predicting students achievements. *Indonesian Journal of Electrical Engineering and Computer Science*, 19(1), 388–394. <https://doi.org/10.11591/ijeecs.v19.i1.pp388-394>
- Bangun, O., Mawengkang, H., & Efendi, S. (2022). Metode Algoritma Support Vector Machine (SVM) Linier Dalam Memprediksi Kelulusan Mahasiswa. *Jurnal Media Informatika Budidarma*, 6(4), 2006. <https://doi.org/10.30865/mib.v6i4.4572>
- Farida, I., & Hendric, S. W. H. L. (2019). Prediksi Pola Kelulusan Mahasiswa Menggunakan Teknik Data Mining Classification Emerging Pattern. *Petir*, 12(1), 1–17. <https://doi.org/10.33322/petir.v12i1.414>
- Haryatmi, E., & Pramita Hervianti, S. (2021). Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 386–392. <https://doi.org/10.29207/resti.v5i2.3007>
- Hernández-Blanco, A., Herrera-Flores, B., Tomás, D., & Navarro-Colorado, B. (2019). A Systematic Review of Deep Learning Approaches to Educational Data Mining. *Complexity*, 2019. <https://doi.org/10.1155/2019/1306039>
- Informatics, J. O. F., & Technology, I. (n.d.). *Journal of informatics, business, education and innovation technology*. 51–59.
- Irfani, F. F. (2020). Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine. *JBMI (Jurnal Bisnis, Manajemen, Dan Informatika)*, 16(3), 258–266. <https://doi.org/10.26487/jbmi.v16i3.8607>
- Khasanah, N., Salim, A., Afni, N., Komarudin, R., & Maulana, Y. I. (2022). Prediksi Kelulusan Mahasiswa Dengan Metode Naive Bayes. *Technologia : Jurnal Ilmiah*, 13(3), 207. <https://doi.org/10.31602/tji.v13i3.7312>
- Loho, C., Rantung, V. P., & Rorimpandey, G. C. (2022). Data Mining Rekomendasi Sekolah Calon Siswa SMA di Kota Tomohon Menggunakan Metode K-Means Clustering. *JOINTER : Journal of Informatics Engineering*, 3(02), 1–9. <https://doi.org/10.53682/jointer.v3i02.34>
- Martanto, M., Ali, I., & Mulyawan, M. (2019). Prediksi Tingkat Kelulusan Mahasiswa Menggunakan Machine Learning dengan Teknik Deep Learning. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(2–2), 191–194. <https://doi.org/10.30591/jpit.v4i2-2.1877>
- Munaiseche, C. P. C., Rantung, V. P., Bawiling, N. S., & Manggopa, H. K. (2019). A knowledge based system for diagnosing heart diseases. *Journal of Physics: Conference Series*, 1402(2). <https://doi.org/10.1088/1742-6596/1402/2/022090>
- Ningrum, A. A., Syarif, I., Gunawan, A. I., Satriyanto, E., & Muchtar, R. (2021). Algoritma

- Deep Learning-LSTM untuk Memprediksi Umur Transformator. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(3), 539–548. <https://doi.org/10.25126/jtiik.2021834587>
- Rantung, V. P., Munaiseche, C. P. C., Rorimpandey, G. C., Sangkop, F. I., Pardanus, R. H. W., & Hoppenbrouwers, S. (2020). Web-based application design for agile stakeholder communication. *Journal of Physics: Conference Series*, 1469(1). <https://doi.org/10.1088/1742-6596/1469/1/012062>
- Rantung, V. P., Munaiseche, C. P. C., & Komansilan, T. (2019). Analysis of higher education executive information system in Indonesia. *Journal of Physics: Conference Series*, 1402(7). <https://doi.org/10.1088/1742-6596/1402/7/077048>
- Rorimpandey, G. C., Sangkop, F. I., Rantung, V. P., Zwart, J. P., Liando, O. E. S., & Mewengkang, A. (2018). Data Model Performance in Data Warehousing. *IOP Conference Series: Materials Science and Engineering*, 306(1). <https://doi.org/10.1088/1757-899X/306/1/012044>
- Rosnita, L., & Putra, R. K. (2023). *Decision Support System Prediction Graduation of*. 9(1), 58–64.
- Sangkop, F. I., Rorimpandey, G. C., Rantung, V. P., Munaiseche, C. P. C., Tombiling, G. W., Runtuwene, J. P. A., & Rompas, P. T. D. (2020). *Web based Home Automation System Prototype using Raspberry Pi. Eic* 2018, 292–297. <https://doi.org/10.5220/0009010002920297>
- Santa, K. (2023). Literasi Digital untuk mengurangi dampak konten negatif bagi ibu-ibu rumah tangga di Minahasa Utara. *Jurnal Locus Penelitian Dan Pengabdian*, 2(4), 312–318. <https://doi.org/10.58344/locus.v2i4.968>
- Setiyani, L. (2020). *ANALISIS PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN METODE DATA MINING NAÏVE BAYES: SYSTEMATIC REVIEW*. 13(1), 35–43. <https://doi.org/10.30998/faktorexacta.v13i1.5548>
- Supit, M. A. (2022). Pemodelan Proses Bisnis dengan Business Process Management Notation pada Fakultas Teknik Universitas Negeri Manado. *Eduetik: Jurnal Pendidikan Teknologi Informasi Dan Komunikasi*, 2(1), 52–64.
- Thaniket, R., Kusrini, & Luthf, E. T. (2019). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Algoritma Support Vector Machine. *JURNAL FATEKSA: Jurnal Teknologi Dan Rekayasa*, 13(2), 69–83.
- Undap, M., Rantung, V. P., & Rompas, P. T. D. (2021). Analisis Sentimen Situs Pembajak Artikel Penelitian Menggunakan Metode Lexicon-Based. *Jointer - Journal of Informatics Engineering*, 2(02), 39–46. <https://doi.org/10.53682/jointer.v2i02.44>
- Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25. <https://doi.org/10.36787/jti.v13i1.78>