

PENERAPAN ALGORITMA C4.5 UNTUK ANALISIS PREDIKSI KUALITAS TIDUR BERDASARKAN GAYA HIDUP

¹Andi Sanjaya, ²Naufal Daffa Erlangga, ³Mohammad Badrul*, ⁴Fajar Mustofa,
⁵Muhammad Rakha Eka, ⁶Muhammad Pashya Mahesa, ⁷Christian Nikolaus Leonardo

¹Program Studi Informatika, ^{2,4,5,6,7}Prodi Sistem Informasi,
³Prodi Teknologi Informasi, Fakultas Teknik dan Informatika,
Universitas Bina Sarana Informatika
Jl. Kramat Raya No.98, Senen, Jakarta Pusat, 10450, Indonesia

Correspondent Author: mohammad.mbl@bsi.ac.id

e-mail: andi.aij@bsi.ac.id, erlangga.nde@gmail.com, mohammad.mbl@bsi.ac.id,
mustofafajar2004@gmail.com, pashya.mahesa@gmail.com, nikko.181204@gmail.com

Abstrak

Insomnia adalah masalah tidur yang umum yang ditandai dengan kesulitan tidur sehingga mengakibatkan kualitas tidur yang buruk. Kondisi ini dapat menyebabkan berbagai dampak negatif seperti kelelahan, penurunan produktivitas, dan masalah kesehatan mental. Siklus tidur alami sendiri dapat terganggu oleh kebiasaan gaya hidup modern, misalnya tingkat stres yang tinggi di tempat kerja, kurangnya aktivitas fisik, dan durasi tidur yang tidak memadai. Untuk menganalisis dan mengklasifikasikan faktor-faktor gaya hidup tersebut, sebuah model prediktif akan dibangun menggunakan dataset "Sleep_Health_and_Lifestyle_Dataset" dari Kaggle, yang terdiri dari variabel seperti durasi tidur, kualitas tidur, tingkat aktivitas fisik, tingkat stres, kategori BMI, dan tekanan darah. Penelitian ini diimplementasikan melalui perangkat lunak RapidMiner dengan menggunakan algoritma data mining C4.5, yang memiliki keunggulan dalam menangani atribut numerik dan kategorikal.

Kata kunci: Kualitas tidur, insomnia, gaya hidup, algoritma C4.5, data mining, RapidMiner

Abstract

Insomnia is a common sleep disorder characterized by difficulty sleeping, which results in poor sleep quality. This condition can lead to various negative impacts such as fatigue, decreased productivity, and mental health problems. The natural sleep cycle can be disrupted by modern lifestyle habits, for example, high levels of stress at work, a lack of physical activity, and inadequate sleep duration. To analyze and classify these lifestyle factors, a predictive model will be built using the "Sleep Health and Lifestyle Dataset" from Kaggle, which consists of variables such as sleep duration, sleep quality, physical activity level, stress level, BMI category, and blood pressure. This research is implemented using the RapidMiner software with the C4.5 data mining algorithm, which has the advantage of handling both numerical and categorical attributes.

Keyword: Quality of sleep, insomnia, lifestyle, C4.5 algorithm, data mining, RapidMiner

1 Pendahuluan

Jika kita ingin hidup dengan baik, kita harus aktif tidur, karena orang modern sering mengabaikan pentingnya tidur untuk kesehatan kita. Orang yang memiliki kualitas tidur yang



DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

buruk tidak hanya merasa lelah, tetapi mereka juga kurang produktif, lebih rentan terhadap penyakit jangka panjang, dan gangguan kesehatan mental [1]. Gaya hidup seseorang telah memasukkan banyak kebiasaan yang, dengan semua kompleksitasnya, dapat mengganggu siklus tidur alami seseorang. Banyak elemen gaya hidup modern telah menjadi bagian dari rutinitas sehari-hari banyak orang. Ini termasuk konsumsi terlalu banyak kafein, tekanan kerja yang berlebihan, dan terpapar cahaya biru dari ponsel hingga larut malam [2]. Salah satu masalah tidur yang paling umum pada orang dewasa adalah insomnia. Sulit tidur dapat membuat tubuh merasa tidak nyaman dan menghabiskan lebih banyak waktu tidur, menyebabkan kualitas tidur yang buruk. Kualitas tidur adalah ketika seseorang tidur dengan baik tanpa masalah tidur. Penurunan aktivitas sehari-hari, rasa lelah, lemah, tanda vital yang tidak stabil, kondisi neurologis yang buruk, penyembuhan luka yang lambat, dan penurunan kekebalan tubuh adalah beberapa akibat dari kualitas tidur yang buruk [3].

2 Tinjauan Literatur

Penelitian terdahulu telah banyak membahas hubungan antara kualitas tidur dan parameter kesehatan, serta penerapan data mining di bidang kesehatan. Studi literatur menunjukkan bahwa algoritma C4.5 sering dibandingkan dengan metode lain seperti *Naïve Bayes* dalam konteks prediksi medis, di mana C4.5 menunjukkan kemampuan yang baik dalam membentuk struktur keputusan yang jelas. Selain itu, penelitian lain juga menyoroti penggunaan *Decision Trees* untuk mendeteksi pola tidur sehat berdasarkan kebiasaan harian, menegaskan bahwa variabel gaya hidup adalah prediktor yang valid.

Algoritma C4.5 bekerja dengan konsep *entropy* dan *information gain* untuk memecah atribut menjadi pohon keputusan. Dalam penerapannya untuk klasifikasi perilaku atau penyakit, seperti diabetes mellitus dan perilaku penggunaan perangkat seluler, metode ini terbukti efektif dalam memetakan atribut prediktor ke dalam kelas target. Penelitian ini mengembangkan studi sebelumnya dengan memfokuskan analisis pada dataset *Sleep Health and Lifestyle* yang mengintegrasikan variabel fisik (BMI, tekanan darah) dan psikologis (tingkat stres) secara bersamaan.

3 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data sekunder "*Sleep Health and Lifestyle Dataset*" dari repositori Kaggle yang terdiri dari 374 record. Atribut data mencakup *Gender*, *Usia*, *Durasi Tidur*, *Kualitas Tidur*, *Aktivitas Fisik*, *Tingkat Stres*, *Kategori BMI*, *Tekanan Darah*, *Detak Jantung*, dan *Langkah Harian*.



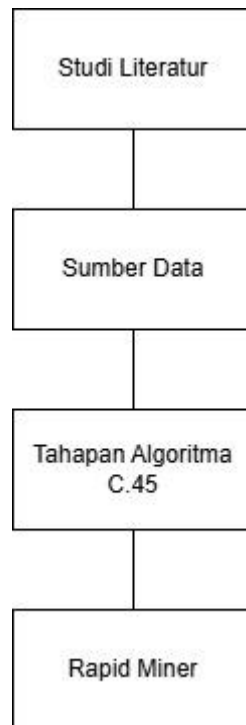
DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

Person ID	Gender	Age	Occupation	Deep Durability of Sleep	Cal Activity	Stress Level	MI Category	Blood Pressure	Heart Rate	Daily Step	sleep Disorder	
1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	None
2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	None
3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	None
4	Male	28	Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
5	Male	28	Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
6	Male	28	Software Engineer	5.9	4	30	8	Obese	140/90	85	3000	Insomnia
7	Male	29	Teacher	6.3	6	40	7	Obese	140/90	82	3500	Insomnia
8	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	None
9	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	None
10	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	None
11	Male	29	Doctor	6.1	6	30	8	Normal	120/80	70	8000	None
12	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	None
13	Male	29	Doctor	6.1	6	30	8	Normal	120/80	70	8000	None
14	Male	29	Doctor	6	6	30	8	Normal	120/80	70	8000	None
15	Male	29	Doctor	6	6	30	8	Normal	120/80	70	8000	None
16	Male	29	Doctor	6	6	30	8	Normal	120/80	70	8000	None
17	Female	29	Nurse	6.5	5	40	7	Normal Weight	132/87	80	4000	Sleep Apnea
18	Male	29	Doctor	6	6	30	8	Normal	120/80	70	8000	Sleep Apnea
19	Female	29	Nurse	6.5	5	40	7	Normal Weight	132/87	80	4000	Insomnia
20	Male	30	Doctor	7.6	7	75	6	Normal	120/80	70	8000	None

Tahapan penelitian diawali dengan pra-pemrosesan data (*data preprocessing*) untuk menangani missing values dan melakukan transformasi atribut. Atribut numerik seperti Stress Level dan Blood Pressure dikonversi menjadi data kategorikal (diskritisasi) untuk menyederhanakan proses percabangan pohon keputusan. Dataset kemudian dibagi (*splitting*) dengan rasio 70% sebagai data latih dan 30% sebagai data uji.



Gambar 1. Tahapan Alur Penelitian



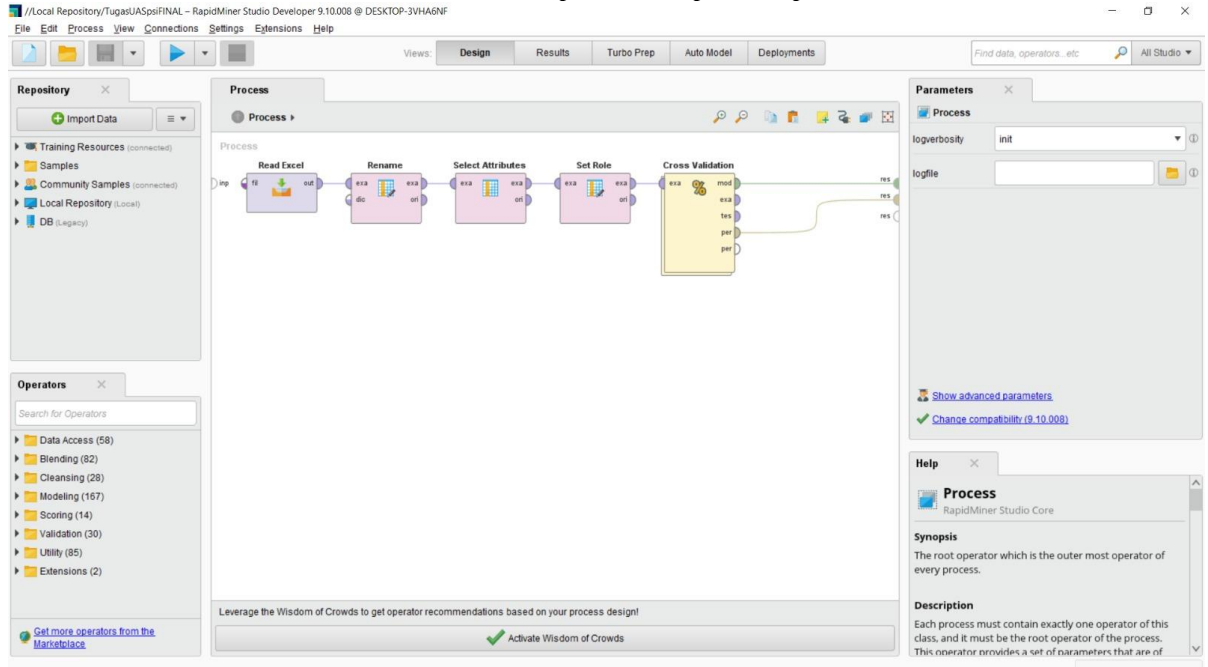
DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

Proses pembentukan model dilakukan menggunakan perangkat lunak RapidMiner Studio. Atribut Person ID dihapus pada tahap seleksi fitur karena tidak relevan, sementara atribut Sleep Disorder ditetapkan sebagai label target. Validasi model dilakukan menggunakan metode Cross Validation untuk memastikan objektivitas hasil, dan performa diukur menggunakan *Confusion Matrix* yang meliputi parameter *Accuracy*, *Precision*, dan *Recall*.

Gambar 2. Desain Proses Implementasi pada RapidMiner



4 Hasil dan Pembahasan (or Results and Analysis)

Berdasarkan eksperimen yang dilakukan menggunakan *RapidMiner*, model C4.5 berhasil terbentuk dengan tingkat akurasi sebesar 90.63% +/- 2.88%.

	True None	True Sleep Apnea	True Insomnia	Class Precision
Pred. None	209	8	7	93.30%
Pred. Sleep Apnea	6	67	7	83.75%
Pred. Insomnia	4	3	63	90.00%
Class Recall	95.43%	83.75%	81.82%	

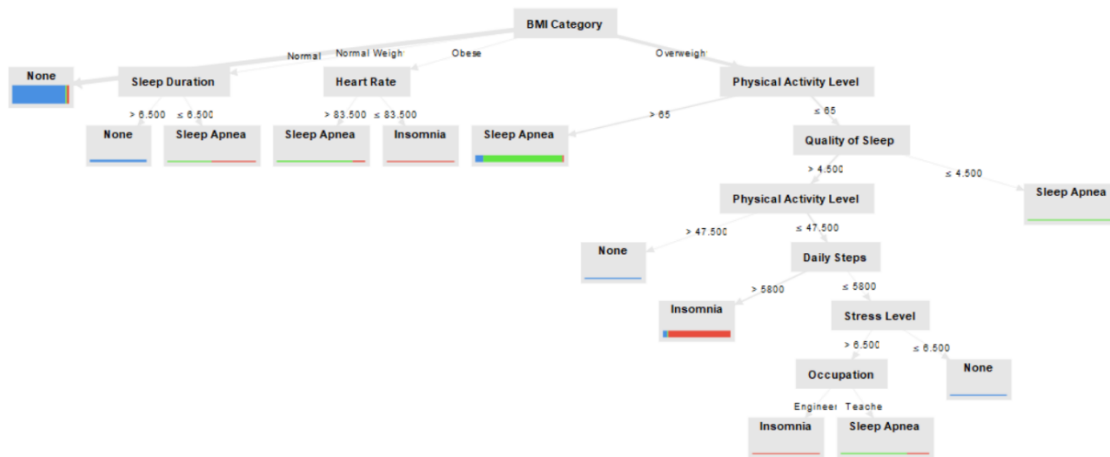
Dari *Confusion Matrix*, model mampu memprediksi kelas "None" (sehat) dengan sangat baik (209 prediksi benar), serta mengidentifikasi kasus *Sleep Apnea* dan "Insomnia" dengan presisi yang memadai, meskipun terdapat sedikit kesalahan klasifikasi antar-kelas gangguan tidur. Tingkat classification error yang rendah, yaitu 9.37%, menunjukkan model ini cukup andal.



DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>



Gambar 3. Struktur Pohon Keputusan Algoritma C4.5

Analisis terhadap struktur pohon keputusan yang dihasilkan mengungkap temuan menarik yang sedikit berbeda dari asumsi awal. Meskipun "Langkah Harian" memiliki Information Gain tinggi pada perhitungan awal, struktur pohon final menunjukkan bahwa *BMI Category* (Kategori Berat Badan) dan *Daily Steps* (Langkah Harian) saling berinteraksi kuat sebagai penentu utama.

Pola aturan (*rules*) yang diekstraksi dari model menunjukkan hubungan kausalitas yang logis:

1. Individu dengan *BMI Category* "Normal" memiliki kecenderungan terbesar untuk tidak memiliki gangguan tidur.
2. Pada individu dengan kategori *Overweight* atau *Obese*, risiko gangguan tidur meningkat signifikan jika dikombinasikan dengan aktivitas fisik rendah. Secara spesifik, jika seseorang *Overweight*, memiliki aktivitas fisik rendah (≤ 65), dan langkah harian rendah (≤ 5800), mereka sangat rentan terkena *Insomnia*.
3. Tingkat stres dan jenis pekerjaan juga menjadi faktor pembeda pada cabang daun terakhir; misalnya, profesi *Engineer* dengan stres tinggi cenderung mengalami *Insomnia*, sedangkan *Teacher* pada kondisi serupa lebih rentan terhadap *Sleep Apnea*.

Hasil ini mengonfirmasi bahwa menjaga berat badan ideal dan aktivitas fisik adalah kunci utama pencegahan gangguan tidur, melebihi faktor demografis semata.

5 Kesimpulan

Penelitian ini menyimpulkan bahwa algoritma C4.5 efektif dalam mengklasifikasikan kualitas tidur berdasarkan gaya hidup dengan akurasi mencapai 90.63%, menjadikannya model yang valid untuk



DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

deteksi dini. Transformasi atribut numerik menjadi kategorikal terbukti meningkatkan kualitas pohon keputusan yang terbentuk. Temuan substansial dari penelitian ini adalah identifikasi tingkat aktivitas fisik (langkah harian) dan kategori BMI sebagai faktor determinan paling dominan. Kombinasi antara gaya hidup sedenter (kurang gerak), tingkat stres tinggi, dan berat badan berlebih memiliki korelasi sangat kuat dengan diagnosis Insomnia dan Sleep Apnea. Disarankan bagi penelitian selanjutnya untuk menggunakan dataset dengan volume lebih besar dan membandingkan metode ini dengan algoritma lain seperti SVM atau KNN untuk validasi performa yang lebih komprehensif.

Referensi

- [1] K. Tidur, K. Fakhruddin, and S. Entas, "Perbandingan Algoritma C4.5 dan Naïve Bayes dalam Prediksi," *Jurnal Ilmu Kesehatan Umum (Vitamin)*, vol. 3, no. 4, 2025, doi: 10.61132/vitamin.v3i4.1773.
- [2] Imam Nawawi and Zaehol Fatah, "Penerapan Decision Trees dalam Mendeteksi Pola Tidur Sehat Berdasarkan Kebiasaan Gaya Hidup," *JURNAL ILMIAH SAINS TEKNOLOGI DAN INFORMASI*, vol. 2, no. 4, pp. 34–41, Oct. 2024, doi: 10.59024/jiti.v2i4.969.
- [3] E. Maulida, A. Fitri Hidayati, F. Y. Silalahi, and Y. D. Prasetyo, "Implementasi Algoritma K-Means Clustering dalam Penentuan Gangguan Tidur Seseorang berdasarkan Gaya Hidup," 2024.
- [4] S. D. Putra, D. Aryani, H. Syofyan, and V. Yasin, "Aplikasi Augmented Reality Geometri Sekolah Dasar Untuk Bangun Datar dan Ruang Menggunakan Metode Marker Based Tracking," *J. Media Inform. Budidarma*, vol. 7, no. 1, pp. 250–259, 2023, doi: <http://dx.doi.org/10.30865/mib.v7i1.5281>.
- [5] I. Ihramsyah, V. Yasin, and J. Johan, "Perancangan Aplikasi Sistem Informasi Penjualan Makanan Cepat Saji Berbasis Web Studi Kasus Kedai Cheese. Box," *J. Widya*, vol. 4, no. 1, pp. 117–139, 2023, doi: <https://doi.org/10.54593/awl.v4i1.170>.
- [7] N. Amalya and V. Yasin, "Perbandingan Algoritma Resilient Backpropagation Dan Conjugate Gradient Polak-Ribiere Dalam Memprediksi Penyakit Stroke," *J. Inov. Sist. Inf. Ilmu Komput.*, vol. 1, no. 1, pp. 12–19, 2023, [Online]. Available: <https://jisiilkom.org/index.php/journal/article/view/6>
- [8] H. Iswanto, V. Yasin, J. Johan, and R. Hartawan, "MONITORING APPLICATION FOR SUBMISSION OF PBJ TRANSACTION REPORTS ON PPAATK," *JISICOM (Journal Inf. Syst. Informatics Comput.*, vol. 7, no. 1, pp. 179–187, 2023, doi: <https://doi.org/10.52362/jisicom.v7i1.704>.
- [9] A. Jazaudhi'fi, A. V. Vitianingsih, Y. Kristyawan, A. L. Maukar, and V. Yasin, "Recommendation System to Determine Achievement Students Using Naïve Bayes and Simple Additive Weighting (SAW) Methods," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 15, no. 1, pp. 67–79, 2024, doi: <https://doi.org/10.31849/digitalzone.v15i1.19746>.
- [10] M. Z. Ardha, V. Yasin, and S. Solikhun, "The Application of the Fletcher-Reeves Algorithm to Predict Spinach Vegetable Production in Sumatra," *Int. J. Eng. Comput. Sci. Appl.*, vol. 2, no. 1, pp. 11–22, 2023, doi: <https://doi.org/10.30812/ijecsa.v2i1.2417>.
- [11] M. Usnaini, V. Yasin, and A. Z. Sianipar, "Perancangan sistem informasi inventarisasi aset berbasis web menggunakan metode waterfall," *J. Manajemen Inform. Jayakarta*, vol. 1, no. 1, pp. 36–55, 2021, doi: <https://doi.org/10.52362/jmijayakarta.v1i1.415>.
- [12] A. Mardian, T. Budiman, R. Haroen, and V. Yasin, "Perancangan Aplikasi Pemantauan Kinerja Karyawan Berbasis Android di PT. Salestrade Corp. Indonesia," *J. Manajemen Inform. Jayakarta*, vol. 1, no. 3, pp. 169–185, 2021, doi: <https://doi.org/10.52362/jmijayakarta.v1i3.481>.
- [13] V. Yasin, P. Peniarsih, A. Gozali, and I. Junaedi, "Application of expert system diagnosis of color blindness with ishihara method with microsoft vb 6.0," *Int. J. Informatics, Econ. Manag. Sci.*, vol. 1, no. 1, pp. 13–31, 2022, doi: <https://doi.org/10.52362/ijiems.v1i1.678>.
- [14] I. Junaedi, D. Abdillah, and V. Yasin, "Analisis Perancangan Dan Pembangunan Aplikasi



DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

- Business Intelligence Penerimaan Negara Bukan Pajak Kementerian Keuangan RI,” *J. Inf. Syst. Applied, Manag. Account. Res.*, vol. 4, no. 3, pp. 88–101, 2020, [Online]. Available: <https://journal.stmikjayakarta.ac.id/index.php/jisamar/article/view/249>
- [15] M. Awaludin and V. Yasin, “Application of Oriented Fast and Rotated BRIEF (ORB) and BruteForce Hamming in Library OpenCV for Classification,” *J. Inf. Syst. Applied, Manag. Accounting, Reserarch. Jakarta.*, vol. 4, no. 3, pp. 51–59, 2020, [Online]. Available: <https://journal.stmikjayakarta.ac.id/index.php/jisamar/article/view/247>
- [16] A. Rohmadi and V. Yasin, “Desain dan penerapan website tata kelola percetakan pada CV Apicdesign Kreasindo Jakarta dengan metode prototyping,” *J. Inf. Syst. Informatics Comput.*, vol. 4, no. 1, pp. 70–85, 2020, [Online]. Available: <https://journal.stmikjayakarta.ac.id/index.php/jisicom/article/view/210>
- [17] N. Mulyana, A. Sulistyanto, and V. Yasin, “PERANCANGAN SISTEM INFORMASI PENGELOLAAN ASET IT BERBASIS WEB PADA PT MANDIRI AXA GENERAL INSURANCE,” *J. Manajemen Inform. Jayakarta*, vol. 1, no. 3, pp. 243–257, 2021, doi: <https://doi.org/10.52362/jmijayakarta.v1i3.498>.
- [18] P. Setiani, I. Junaedi, A. Z. Sianipar, and V. Yasin, “Perancangan sistem informasi pelayanan penduduk berbasis website di rw 010 Kelurahan Keagungan Kecamatan Tamansari-Jakarta Barat,” *J. Manajemen Inform. Jayakarta*, vol. 1, no. 1, pp. 20–35, 2021, doi: <https://doi.org/10.52362/jmijayakarta.v1i1.414>.
- [19] I. Junaedi, N. Nuswantari, and V. Yasin, “Perancangan Dan Implementasi Algoritma C4. 5 Untuk Data Mining Analisis Tingkat Risiko Kematian Neonatum Pada Bayi,” *J. Inf. Syst. Informatics Comput.*, vol. 3, no. 1, pp. 29–44, 2019, [Online]. Available: <https://journal.stmikjayakarta.ac.id/index.php/jisicom/article/view/203>
- [20] M. Hendriawan, T. Budiman, V. Yasin, and A. S. Rini, “PENGEMBANGAN APLIKASI E-COMMERCE DI PT. PUTRA SUMBER ABADI MENGGUNAKAN FLUTTER,” *J. Inf. Syst. Informatics Comput.*, vol. 5, no. 1, pp. 69–88, 2021, doi: <https://doi.org/10.52362/jisicom.v5i1.371>.
- [21] H. Heriyanto, V. Yasin, and A. B. Yulianto, “Vipos application development design,” *J. Eng. Technol. Comput.*, vol. 1, no. 1, pp. 19–31, 2022, [Online]. Available: <https://journal.binainternusa.org/index.php/jetcom/article/view/3>
- [22] J. S. B. Sitepu, V. Yasin, and A. B. Yulianto, “Design information systems for the management and inventory of goods with web-based a priori algorithm methods,” *J. Eng. Technol. Comput.*, vol. 1, no. 1, pp. 1–9, 2022, [Online]. Available: <https://journal.binainternusa.org/index.php/jetcom/article/view/1>
- [23] D. Malini, V. Yasin, and A. Z. Sianipar, “Perancangan sistem informasi penjadwalan dan distribusi menu diet pasien di RS. Kramat 128 dengan framework laravel,” *J. Sains dan Teknol. Widyaloka*, vol. 1, no. 1, pp. 49–65, 2022, doi: <https://doi.org/10.54593/jstekwid.v1i1.48>.
- [24] H. W. Wibawa, H. M. Ali, D. N. R. Permada, and V. Yasin, “Analysis Of the Effect Of E-Accounting And E-Payroll on The Effectiveness Of Internal Control,” *Int. J. Informatics, Econ. Manag. Sci.*, vol. 1, no. 1, pp. 78–101, 2022, doi: <https://doi.org/10.52362/ijiems.v1i1.706>.



DOI: <https://doi.org/10.52362/jmijayakarta.v6i2.2240>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>