
PREDIKSI KEBERHASILAN SISWA MELALUI PENINGKATAN KUALITAS RENCANA PELAKSANAAN PEMBELAJARAN DENGAN MENGGUNAKAN ALGORITMA C4.5

Pancatma Hesthi Kawuri[#], Chairani^{*}

¹Fakultas Komputer Informatics & Business Institute Darmajaya

Jl. 2.A. Pagar Alam No. 93, Bandar Lampung - Indonesia 35142

Telp. (0721) 787214 Fax. (0721) 700261

E-mail: xxx1@yyy.zzz, xxx2@yyy.zzz

Abstract -This study aims to predict student success through improving the quality of lesson implementation plans using the C4.5 algorithm. Utilizing relevant datasets, this research implements the C4.5 algorithm to generate accurate predictions regarding student success. The research findings indicate an accuracy rate of 93.67%, with precision reaching 91.86%, and a recall/confusion matrix of 89.66%. Additionally, the Area Under Curve (AUC) achieved is 0.986, indicating high prediction quality. By employing this method, refining lesson implementation plans can enhance the likelihood of student success.

Keywords: Student success prediction, lesson implementation plan quality, C4.5 algorithm, accuracy, precision, recall, AUC.

Abstrak — Studi ini bertujuan untuk memprediksi keberhasilan siswa melalui peningkatan kualitas rencana pelaksanaan pembelajaran menggunakan algoritma C4.5. Dengan menggunakan dataset yang relevan, penelitian ini mengimplementasikan algoritma C4.5 untuk menghasilkan prediksi yang akurat terkait keberhasilan siswa. Hasil penelitian menunjukkan tingkat akurasi sebesar 93,67%, dengan precision mencapai 91,86%, dan recall/confusion matrix sebesar 89,66%. Selain itu, Area Under Curve (AUC) yang dihasilkan mencapai 0,986, menunjukkan kualitas prediksi yang tinggi. Dengan menggunakan metode ini, penyempurnaan rencana pelaksanaan pembelajaran dapat meningkatkan kemungkinan keberhasilan siswa.

Kata kunci: Prediksi keberhasilan siswa, kualitas rencana pelaksanaan pembelajaran, algoritma C4.5, akurasi, precision, recall, AUC.

I. PENDAHULUAN

Pendidikan adalah fondasi utama dalam mempersiapkan generasi mendatang untuk menghadapi tantangan dunia yang terus berkembang. Dalam era di mana teknologi informasi semakin mendominasi, penting bagi sistem pendidikan untuk terus berinovasi dalam menyajikan materi pembelajaran yang efektif dan relevan. Salah satu aspek penting dalam memajukan pendidikan adalah penggunaan teknologi informasi dalam proses pembelajaran.[1][2]

Dalam konteks ini, algoritma C4.5 menjadi sebuah alat yang potensial dan menarik dalam mendukung proses pembelajaran di sekolah menengah pertama. Algoritma C4.5, yang merupakan salah satu algoritma dalam data mining dan machine learning, memiliki kemampuan untuk menghasilkan model keputusan berbasis pohon yang dapat digunakan dalam berbagai domain, termasuk dalam analisis dan pengambilan keputusan pada data pembelajaran siswa.[4][5]

Rencana pelaksanaan pembelajaran siswa sekolah menengah pertama menggunakan algoritma C4.5 bertujuan untuk memanfaatkan kecanggihan teknologi ini dalam meningkatkan efektivitas pembelajaran di kelas. Melalui penerapan algoritma C4.5, diharapkan pembelajaran menjadi lebih adaptif, personal, dan mampu menyesuaikan diri dengan kebutuhan serta kemampuan individual siswa.[1]

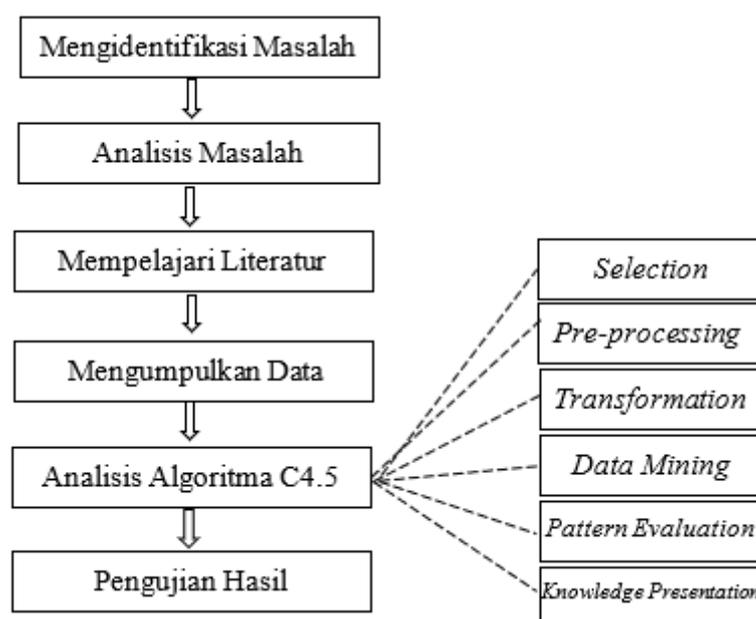
Penerapan algoritma C4.5 dalam pembelajaran siswa sekolah menengah pertama juga dapat memberikan berbagai manfaat, antara lain:

1. Memungkinkan identifikasi pola pembelajaran siswa secara lebih tepat dan mendalam.
2. Memfasilitasi penyusunan kurikulum dan materi pembelajaran yang lebih disesuaikan dengan kebutuhan siswa.
3. Mendorong penggunaan metode pembelajaran yang lebih interaktif dan berbasis masalah.
4. Meningkatkan keterlibatan siswa dalam proses pembelajaran dengan menyajikan materi yang lebih relevan dan menarik bagi mereka.

Dengan demikian, rencana pelaksanaan pembelajaran menggunakan algoritma C4.5 di sekolah menengah pertama merupakan langkah inovatif yang dapat membawa manfaat besar dalam meningkatkan kualitas pendidikan. Melalui pendekatan ini, diharapkan siswa dapat memperoleh pengalaman pembelajaran yang lebih bermakna dan relevan dengan perkembangan teknologi dan tuntutan zaman.[6]

II. METEDOLOGI PENELITIAN

Kerangka pikir yang digunakan dalam penelitian ini dapat dilihat pada gambar 1.



Tahapan penelitian pada gambar 3.1 di atas dapat diuraikan berikut ini:[7]

a. Mengidentifikasi Masalah

Sebelum melakukan penelitian, terlebih dahulu peneliti menentukan masalah. Masalah dalam penelitian ini adalah bagaimana meningkatkan kualitas rencana pelaksanaan pembelajaran dengan menggunakan metode *decision tree* dengan Algoritma C4.5 pada SMP Negeri 1 Jati Agung Kabupaten Lampung Selatan. [2][8]

b. Analisis Masalah

Setelah masalah diketahui, maka harus ditentukan ruang lingkup dari penelitian agar analisis masalah dapat lebih terarah.

c. Mempelajari Literatur

Literatur diambil dari buku dan jurnal yang berkaitan dengan kepuasan pengunjung menggunakan teknik data mining.

d. Mengumpulkan Data

Data yang dikumpulkan adalah data kuisisioner yang dibuat oleh SMP Negeri 1 Jati Agung Kabupaten Lampung Selatan

e. *Data Mining* menggunakan Algoritma C4.5

Merupakan algoritma yang digunakan untuk membentuk pohon keputusan (Decision Tree). Pohon keputusan merupakan metode klasifikasi dan prediksi yang terkenal. Pohon keputusan berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Banyak algoritma yang dapat dipakai dalam pembentukan pohon keputusan, mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule[4]

f. *Pattern Evaluation*

Setelah bahan mentah emas selesai diolah, maka perlu diuji juga apakah sudah sesuai dengan standar kualitasnya atau belum. Baru setelah itu bisa didistribusikan ke toko emas dan di-*display* kepada pelanggan. Setelah selesai melakukan proses data mining, pola-pola yang dihasilkan dari proses tersebut perlu untuk dievaluasi. Tujuan dari dilakukannya evaluasi adalah untuk menguji hipotesis awal. Setelah teruji data bisa dipresentasikan kepada pengguna[9].[10]

g. *Knowledge Presentation*

Pemrosesan data yang dilakukan secara komputerisasi dalam berbagai bidang, baik dalam penanganan transaksi bisnis maupun untuk kalangan pemerintah dan sosial telah menghasilkan data yang berukuran besar. Data yang dihasilkan tersebut merupakan tambang emas yang dapat diolah untuk mendapatkan pengetahuan. Data yang berskala besar tersebut

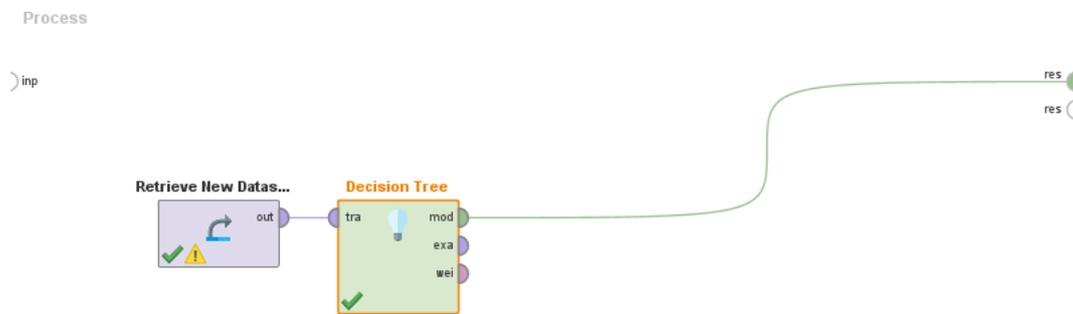
Untuk melakukan analisis ini, teknologi data mining digunakan sebagai langkah kunci dalam proses Knowledge Discovery in Database (KDD). KDD merupakan suatu rangkaian terstruktur

yang bertujuan untuk menemukan pola-pola yang sah, inovatif, berguna, dan dapat dipahami dari kumpulan data yang besar dan kompleks (Maimon dan Lior, 2010). Proses KDD mencakup langkah-langkah seperti pembersihan data, integrasi data, pemilihan data, transformasi data, data mining, evaluasi pola, dan penyajian pengetahuan.[11]

III. HASIL DAN PEMBAHASAN

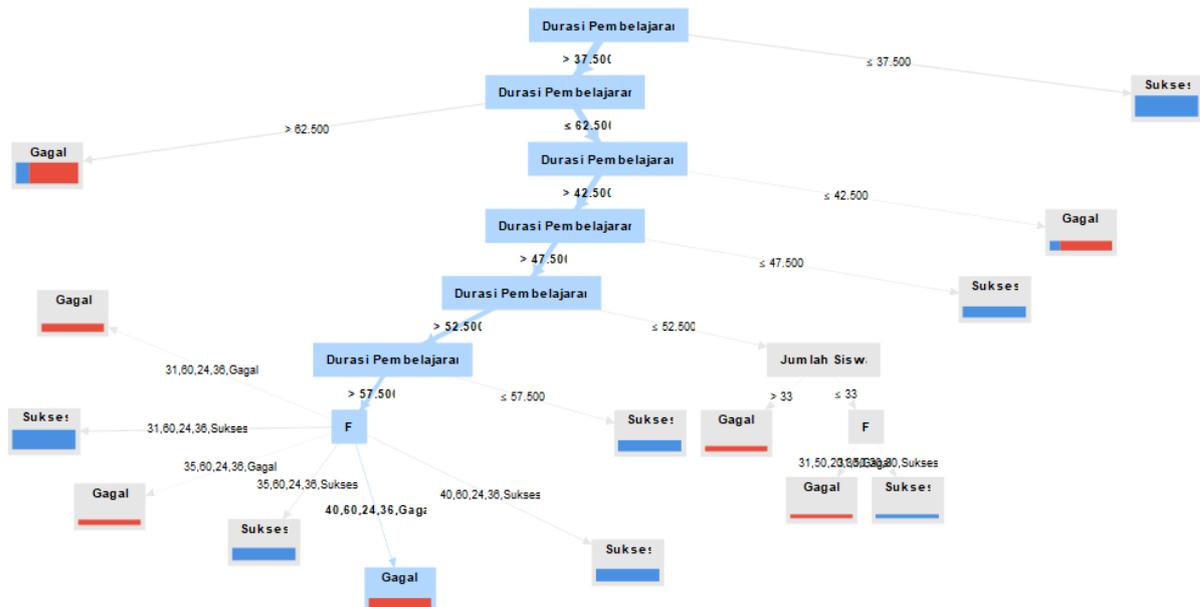
1. Penelitian ini memanfaatkan metode Decision Tree C.45.

Dalam penggunaan algoritma decision tree, data yang diambil dan dimasukkan ke dalam perangkat lunak RapidMiner dimaksudkan untuk mengungkap pola-pola yang terdapat dalam struktur pohon keputusan. Hasil analisis dari pohon keputusan dapat divisualisasikan dalam bentuk gambar yang terlampir.



Gambar 4.2. Tampilan Proses Decision Tree C4.5

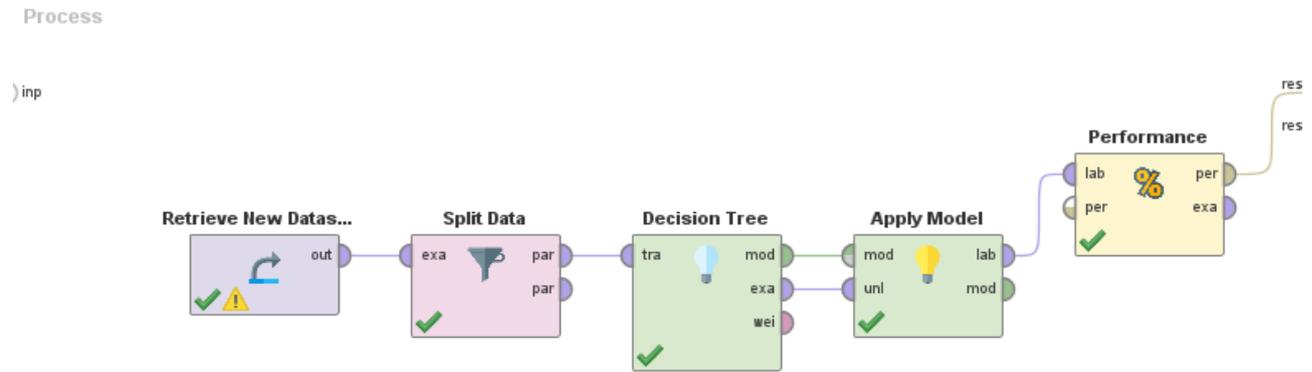
Setelah menyelesaikan beberapa langkah dalam proses klasifikasi menggunakan metode algoritma C4.5, hasilnya adalah pembentukan model melalui proses pengklasifikasiannya. Model ini berupa pohon keputusan, sebagaimana yang ditunjukkan dalam gambar di bawah ini.



Gambar 4.3. Tampilan Pohon Keputusan Decision Tree C4.5

2. Algoritma Decision Tree C4.5 Dengan Split Validation

Penerapan metode Decision Tree dalam perangkat RapidMiner dengan penerapan validasi split 80:20 untuk data training dan data testing, dilakukan dengan tujuan untuk mengukur akurasi, presisi, dan matriks kebingungan. Diagram alur operasi pada RapidMiner tersedia pada gambar di bawah ini:



Gambar 4. 4 Proses *Decision Tree C4.5* dan *Split Validation*

Pengujian dilakukan dengan confusion matrix yang terdiri dari accuracy, precision recall dan AUC dilakukan pada dataset data yang diolah dengan menggunakan Decision Tree. Pengujian confusion matrix untuk dataset yang diolah menggunakan Decision Tree dapat dilihat pada tabel dibawah ini.

Hasil Akurasi Decision Tree C4.5 dan Split Validation, pada pembagian data sampling 80% dan data testing 20% maka didapatkan akurasi sebesar 93,67%

accuracy: 93.67%

	true Sukses	true Gagal	class precision
pred. Sukses	48	3	94.12%
pred. Gagal	2	26	92.86%
class recall	96.00%	89.66%	

Gambar 4.5 Nilai Akurasi Algoritma Decision Tree

$$\begin{aligned}
 \text{Akurasi} &= \frac{(TN + TP)}{(TN + FN + FP + TP)} \\
 \text{Akurasi} &= \frac{(48 + 25)}{(48 + 25 + 3 + 2)} \\
 &= \frac{73}{78} \\
 &= 93,67\%
 \end{aligned}$$

Berikut adalah hasil precision yaitu mendapatkan 91,86% seperti gambar dibawah ini.

precision: 92.86% (positive class: Gagal)

	true Sukses	true Gagal	class precision
pred. Sukses	48	3	94.12%
pred. Gagal	2	26	92.86%
class recall	96.00%	89.66%	

Gambar 4.6 Nilai Precision Algoritma Decision Tree

$$P:precision = \frac{TP}{(TP+FP)}$$

$$\begin{aligned} :p(1) &= \frac{26}{(26+2)} \\ &= \frac{26}{28} \\ &= 92,86\% \end{aligned}$$

$$\begin{aligned} :p(0) &= \frac{48}{(48+3)} = \\ &= \frac{48}{51} \\ &= 94,12\% \end{aligned}$$

Berikut adalah hasil confusion matrix yaitu mendapatkan 89,66% seperti gambar dibawah ini

recall: 89.66% (positive class: Gagal)

	true Sukses	true Gagal	class precision
pred. Sukses	48	3	94.12%
pred. Gagal	2	26	92.86%
class recall	96.00%	89.66%	

Gambar 4.7 Nilai Reccal Algoritma Decision Tree

$$: Recall = \frac{TP}{(TP+FN)}$$

$$R(1) = \frac{26}{(26+3)}$$

$$= \frac{26}{29}$$

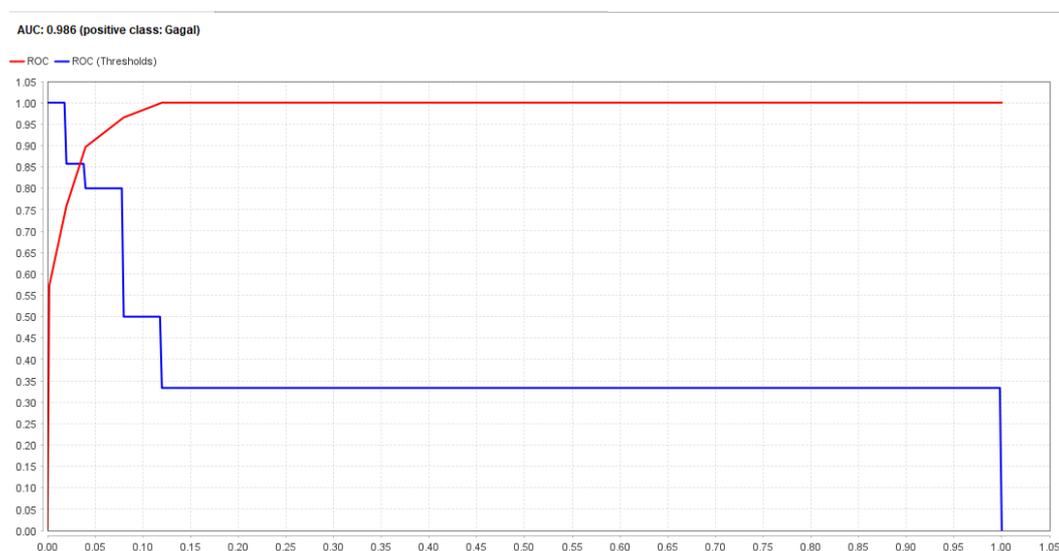
$$= 89,66\%$$

$$R(0) = \frac{48}{(48+2)}$$

$$= \frac{48}{50}$$

$$= 96,00\%$$

Berikut adalah hasil kurva AUC yaitu mendapatkan nilai 0,986 seperti gambar dibawah ini



Gambar 4.8 Nilai AUC Algoritma Decision Tree

IV. SIMPULAN

Berdasarkan hasil evaluasi model Decision Tree yang telah disajikan, dapat diambil beberapa kesimpulan dan saran sebagai berikut

1. Kinerja Model yang Unggul: Model Pohon Keputusan menunjukkan performa yang tinggi berdasarkan evaluasi metrik yang telah diukur. Tingkat akurasi, presisi, recall, dan AUC semuanya berada pada level yang memuaskan, menunjukkan kapabilitas model untuk melakukan klasifikasi dengan baik.
2. Potensi Penerapan Model: Dengan tingkat akurasi, presisi, dan recall yang tinggi, model Pohon Keputusan mungkin menjadi pilihan yang optimal untuk tugas klasifikasi pada dataset yang relevan dengan konteks yang sama.
3. Tingkat AUC yang Tinggi: AUC mendekati 1, menandakan bahwa model memiliki kemampuan yang sangat baik dalam memisahkan antara kelas positif dan negatif, memberikan keyakinan tambahan terhadap kualitas prediksinya.

DAFTAR PUSTAKA

- [1] P. Mata, P. Matematika, D. Z. Azhari, I. S. Damanik, and D. Suhendro, "Penerapan Algoritma C4.5 Untuk Klasifikasi Tingkat Pemahaman Siswa," vol. 1, no. 1, pp. 11–20, 2022.
- [2] A. H. Wicaksono *et al.*, "Klasifikasi Siswa Slow Learner Untuk Mendukung Sekolah Algoritma Naïve Bayes Classification of Slow Learner Students To Support Schools in Improving Student Understanding Using the Naïve Bayes," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, pp. 589–596, 2022, doi: 10.25126/jtiik.202295509.
- [3] D. Michael, R. E. Indrajit, and E. Dazki, "Implementation of Enterprise Architecture in Cloud Computing Companies," *Sinkron*, vol. 7, no. 2, pp. 549–559, 2022, doi: 10.33395/sinkron.v7i2.11407.
- [4] M. Solehuddin, W. A. Syafei, and R. Gernowo, "Metode Decision Tree untuk Meningkatkan Kualitas Rencana Pelaksanaan Pembelajaran dengan Algoritma C4.5," *J. Penelit. dan Pengemb. Pendidik.*, vol. 6, no. 3, pp. 510–519, 2022, doi: 10.23887/jppp.v6i3.52840.
- [5] L. Kartikawati, K. Kusrini, and E. T. Luthfi, "Algoritma K-Means pada Pengelompokan Pembelajaran Tatap Muka Terbatas Sesudah Vaksinasi COVID-19," *J. Eksplora Inform.*, vol. 11, no. 1, pp. 20–28, 2022, doi: 10.30864/eksplora.v11i1.560.
- [6] S. S. P. Luka, I. M. Candiasa, and K. Y. E. Aryanto, "Analisis Pembentukan Kelompok Diskusi Panel Siswa Menggunakan Algoritma Fuzzy C-Means Dan K-Means," *J. Pendidik. Teknol. dan Kejuru.*, vol. 16, no. 2, pp. 267–277, 2019, [Online]. Available: <https://ejournal.undiksha.ac.id/index.php/JPTK/article/download/18913/11418>
- [7] D. Anggraeni and T. Christy, "Analisa Kinerja Algoritma C4.5 Dalam Menentukan Pola Dominasi Mainstream Mahasiswa," *Jurikom*, vol. 6, no. 4, pp. 333–339, 2019, [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/jurikom%7CPage333>
- [8] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.
- [9] O. Nurdiawan, R. Herdiana, and S. Anwar, "Komparasi Algoritma Naïve Bayes dan Algoritma K-Nearest Neighbor terhadap Evaluasi Pembelajaran Daring," *Smatika J.*, vol. 11, no. 02, pp. 126–135, 2021, doi: 10.32664/smatika.v11i02.621.
- [10] Y. Azhar, A. K. Firdausy, and P. J. Amelia, "Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Penyakit Stroke," *SINTECH (Science Inf. Technol. J.)*, vol. 5, no. 2, pp. 191–197, 2022, doi: 10.31598/sintechjournal.v5i2.1222.
- [11] A. R. Damanik, S. Sumijan, and G. W. Nurcahyo, "Prediksi Tingkat Kepuasan dalam Pembelajaran Daring Menggunakan Algoritma Naïve Bayes," *J. Sistim Inf. dan Teknol.*, vol. 3, pp. 88–94, 2021, doi: 10.37034/jsisfotek.v3i3.49.