

## **“Integrasi Algoritma C4.5 dan Apriori untuk Analisis Perilaku Belajar dan Prediksi Hasil Akademik di SMK Taman Siswa Banjarnegara”**

**Purwanto<sup>1</sup>, Dwi Purbaningrum<sup>2</sup>**

STIMIK Tunas Bangsa<sup>1</sup>, STIT Tunas Bangsa<sup>2</sup>

<sup>1</sup>Purwanto@stb.ac.id, <sup>2</sup>dwipurbaningrum1991@gmail.com

**Abstrak** - Penelitian ini bertujuan untuk mengeksplorasi pola belajar siswa dan memperkirakan capaian akademik siswa kelas XII SMK Taman Siswa Banjarnegara melalui pendekatan *data mining*, dengan memanfaatkan algoritma C4.5 untuk klasifikasi dan Apriori untuk pencarian hubungan antar variabel. Sebanyak 60 siswa berpartisipasi sebagai responden melalui pengisian kuesioner yang mencakup aspek durasi belajar, waktu belajar, media pembelajaran, serta nilai akhir. Algoritma C4.5 digunakan untuk membuat model prediksi hasil belajar, sementara Apriori menggali pola keterkaitan antar kebiasaan belajar. Hasil penelitian menunjukkan akurasi klasifikasi C4.5 sebesar 38,89%, dan algoritma Apriori berhasil menemukan aturan asosiasi yang mendukung proses pengambilan keputusan akademik. Temuan ini menunjukkan bahwa gabungan kedua algoritma dapat memberikan wawasan yang bermanfaat bagi sekolah dalam memahami faktor-faktor yang mempengaruhi hasil belajar siswa serta mendesain strategi pembelajaran berbasis data.

**Kata Kunci:** data mining, C4.5, Apriori, klasifikasi, asosiasi, perilaku belajar, prediksi akademik, SMK.

**Abstract** - This study aims to explore student learning patterns and predict academic achievement of grade XII students of SMK Taman Siswa Banjarnegara through a data mining approach, utilizing the C4.5 algorithm for classification and Apriori for finding relationships between variables. A total of 60 students participated as respondents by filling out a questionnaire covering aspects of study duration, study time, learning media, and final grades. The C4.5 algorithm was used to create a prediction model for learning outcomes, while Apriori explored patterns of relationships between learning habits. The results showed that the C4.5 classification accuracy was 38.89%, and the Apriori algorithm successfully found association rules that support the academic decision-making process. These findings indicate that the combination of the two algorithms can provide useful insights for schools in understanding factors that influence student learning outcomes and designing data-based learning strategies.

**Keywords:** data mining, C4.5, Apriori, classification, association, learning behavior, academic prediction, vocational school.

## **BAB I – PENDAHULUAN**

### **1.1 Latar Belakang**

Kemajuan teknologi informasi telah memberikan dampak yang luas pada berbagai sektor, termasuk bidang pendidikan. Penerapan sistem pembelajaran berbasis teknologi seperti *Learning Management System* (LMS) menjadi semakin umum di sekolah-sekolah, termasuk di tingkat Sekolah Menengah Kejuruan (SMK). Di SMK Taman Siswa Banjarnegara, penggunaan LMS telah mulai diadopsi sebagai bagian dari proses pembelajaran, khususnya bagi siswa kelas XII. Meski teknologi telah tersedia, pemanfaatannya oleh siswa masih bervariasi. Sebagian siswa aktif berpartisipasi dalam aktivitas pembelajaran daring seperti menonton video edukatif dan mengikuti forum diskusi, sementara sebagian lainnya kurang terlibat. Perbedaan ini dapat mempengaruhi performa akademik mereka, sehingga penting untuk mengidentifikasi pola perilaku belajar tersebut dengan pendekatan berbasis data.

Sejumlah studi sebelumnya telah membahas penerapan data mining dalam pendidikan. Romero dan Ventura (2020) mengemukakan bahwa *Educational Data Mining* (EDM) memiliki potensi besar dalam mengidentifikasi perilaku siswa dan memprediksi hasil belajar melalui data aktivitas daring. Han et al. (2012) juga menyoroti bahwa

metode seperti klasifikasi dan asosiasi dapat digunakan untuk mengevaluasi efektivitas pembelajaran digital. C4.5 dikenal luas sebagai algoritma klasifikasi yang digunakan untuk membangun model prediktif berdasarkan data historis. Putra dan Hermawan (2022) memanfaatkannya untuk menganalisis pemahaman siswa dari data interaksi di LMS. Di sisi lain, algoritma Apriori yang dikenalkan oleh Agrawal dan Srikant (1994) dapat menemukan keterkaitan antar atribut dalam dataset pendidikan.

Kendati demikian, banyak penelitian hanya memanfaatkan satu jenis algoritma. Studi yang menggabungkan C4.5 dan Apriori secara bersamaan masih terbatas, terutama yang dilakukan di konteks SMK di wilayah semi-periferal seperti Banjarnegara. Maka dari itu, penelitian ini mencoba mengisi celah tersebut dengan menyatukan dua pendekatan data mining tersebut untuk memahami perilaku belajar siswa kelas XII secara lebih komprehensif. Sebagian besar penelitian sebelumnya di bidang *educational data mining* hanya menggunakan satu pendekatan algoritmik, seperti klasifikasi atau asosiasi secara terpisah, tanpa mengintegrasikan keduanya untuk analisis yang lebih komprehensif. Selain itu, konteks penelitian umumnya berfokus pada jenjang perguruan tinggi atau sekolah di wilayah perkotaan, sehingga kurang merepresentasikan kondisi pembelajaran digital di sekolah kejuruan di daerah semi-periferal seperti Banjarnegara. Oleh karena itu, penelitian ini mengisi gap metodologis dan kontekstual dengan menggabungkan algoritma C4.5 dan Apriori untuk menganalisis perilaku belajar digital dan memprediksi hasil akademik siswa SMK secara simultan dan berbasis data.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimanakah karakteristik perilaku belajar siswa kelas XII SMK Taman Siswa Banjarnegara dalam memanfaatkan platform digital untuk pembelajaran?
2. Faktor-faktor apa saja yang berkontribusi terhadap capaian akademik siswa berdasarkan hasil klasifikasi menggunakan algoritma C4.5?
3. Apa saja pola keterkaitan antar perilaku belajar siswa yang dapat diidentifikasi melalui algoritma Apriori?
4. Bagaimanakah sinergi antara algoritma C4.5 dan Apriori dapat memberikan wawasan lebih dalam terhadap proses pembelajaran siswa?

## **1.3 Tujuan Penelitian**

Penelitian ini memiliki tujuan utama sebagai berikut:

1. Mendeskripsikan karakteristik perilaku belajar siswa kelas XII SMK Taman Siswa Banjarnegara dalam konteks penggunaan platform pembelajaran digital.
2. Mengidentifikasi variabel yang paling berpengaruh terhadap hasil belajar siswa melalui teknik klasifikasi C4.5.
3. Menemukan pola hubungan antar variabel perilaku belajar dengan pendekatan Apriori.
4. Mengembangkan pendekatan hybrid berbasis data mining untuk mendukung strategi pembelajaran yang responsif dan adaptif.

## **1.4 Manfaat Penelitian**

### **Manfaat Teoritis:**

Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan kajian data mining dalam ranah pendidikan kejuruan, khususnya dalam pemetaan perilaku belajar siswa berbasis teknologi.

### Manfaat Praktis:

Hasil penelitian dapat dimanfaatkan oleh pihak sekolah dan guru sebagai dasar untuk merancang strategi pembelajaran berbasis data, serta sebagai bahan pertimbangan dalam pengambilan keputusan yang berkaitan dengan peningkatan kualitas pembelajaran.

## BAB II – TINJAUAN PUSTAKA

### 2.1 Karakteristik Generasi Z dalam Belajar Digital

Generasi Z merupakan generasi yang lahir pada pertengahan tahun 1990-an hingga awal 2010-an, di mana mereka tumbuh bersama teknologi digital sejak usia dini (Prensky, 2010). Generasi ini dikenal cepat beradaptasi dengan perangkat digital, mampu melakukan multitasking, serta menyukai pembelajaran yang bersifat interaktif dan berbasis aplikasi. Di lingkungan SMK, khususnya siswa kelas XII, penggunaan media digital seperti YouTube Edu, Ruangguru, dan Zenius menjadi bagian dari aktivitas belajar sehari-hari.

Hasil penelitian Junco (2012) menunjukkan bahwa intensitas keterlibatan siswa dalam platform digital berbanding lurus dengan keterlibatan akademik mereka, meskipun tidak seluruh aktivitas digital berkontribusi positif terhadap pembelajaran. Oleh sebab itu, perlu dilakukan pemetaan perilaku digital secara sistematis.

### 2.2 Konsep Data Mining dalam Dunia Pendidikan

Data mining dalam konteks pendidikan, yang dikenal dengan sebutan *Educational Data Mining* (EDM), merupakan pendekatan analitik yang bertujuan untuk mengeksplorasi data pendidikan guna menghasilkan informasi dan pola yang bermanfaat (Romero & Ventura, 2020). Melalui EDM, pendidik dapat memahami kecenderungan perilaku siswa berdasarkan data log, evaluasi akademik, dan aktivitas mereka di platform pembelajaran.

Han et al. (2012) menegaskan bahwa pendekatan seperti klasifikasi, asosiasi, dan klusterisasi dapat digunakan untuk mengevaluasi performa akademik siswa, memprediksi prestasi belajar, serta menyarankan strategi pembelajaran yang sesuai.

### 2.3 Algoritma C4.5 dalam Analisis Pendidikan

C4.5 merupakan salah satu metode klasifikasi dalam bentuk pohon keputusan yang populer untuk membangun model prediksi dari data historis. Algoritma ini dikembangkan oleh Quinlan (1996) dan memiliki kemampuan untuk menangani data bertipe numerik maupun kategorikal serta dilengkapi dengan teknik pemangkasan (pruning) untuk menghindari overfitting.

Dalam bidang pendidikan, C4.5 sering digunakan untuk memprediksi capaian akademik berdasarkan atribut perilaku belajar seperti frekuensi login, jenis konten yang diakses, serta durasi aktivitas pembelajaran. Contohnya, Putra dan Hermawan (2022) memanfaatkan C4.5 untuk mengelompokkan pemahaman siswa terhadap materi dari log aktivitas pembelajaran daring.

### 2.4 Algoritma Apriori untuk Menemukan Pola Belajar

Algoritma Apriori merupakan teknik data mining yang digunakan untuk menemukan pola asosiasi atau aturan hubungan antar item dalam suatu dataset. Algoritma ini diperkenalkan oleh Agrawal dan Srikant (1994) dan dikenal efektif untuk menemukan pola keterkaitan dalam data transaksi.

Dalam konteks pembelajaran, Apriori bermanfaat untuk menemukan relasi antar kebiasaan belajar siswa. Misalnya, dalam penelitian Sari dan Dewi (2021), ditemukan bahwa siswa yang sering menonton video pembelajaran juga aktif dalam forum diskusi daring. Informasi semacam ini dapat digunakan untuk merancang pendekatan pembelajaran yang lebih tepat sasaran.

## **2.5 Integrasi Metode (Hybrid) dalam Data Mining**

Pendekatan hybrid atau gabungan dalam data mining bertujuan untuk mengoptimalkan hasil analisis dengan mengkombinasikan dua atau lebih teknik analisis. Dalam studi ini, algoritma C4.5 dan Apriori digabungkan untuk memberikan dua sudut pandang sekaligus: klasifikasi hasil akademik siswa dan analisis pola perilaku belajar mereka.

Ifenthaler & Yau (2020) menyatakan bahwa strategi hybrid dalam analisis data pendidikan mampu memberikan gambaran yang lebih menyeluruh tentang perilaku siswa karena mampu menjawab dua hal sekaligus: siapa yang cenderung berhasil dan perilaku apa yang menunjang keberhasilan tersebut.

## **2.6 Kajian Penelitian Sebelumnya**

Tabel 2. 1 Penelitian Sebelumnya

| No | Peneliti & Tahun        | Fokus Penelitian  | Metode  | Kekurangan/Gagasan Lanjutan                                      |
|----|-------------------------|---|---------|--|
| 1  | Romero & Ventura (2020) | Analisis log pembelajaran berbasis EDM                              | Beragam | Belum spesifik untuk SMK dan tidak menggunakan pendekatan hybrid |
| 2  | Putra & Hermawan (2022) | Klasifikasi pemahaman siswa menggunakan C4.5                        | C4.5    | Belum menyertakan pola asosiasi                                  |
| 3  | Sari & Dewi (2021)      | Pola perilaku mahasiswa di LMS dengan Apriori                       | Apriori | Tidak berfokus pada prediksi hasil belajar                       |
| 4  | Penelitian Ini          | Analisis perilaku belajar siswa SMK dengan kombinasi C4.5 & Apriori | Hybrid  | Konteks lokal dan metode gabungan menjadi nilai tambah           |

# **BAB III – METODOLOGI PENELITIAN**

## **3.1 Jenis Penelitian**

Penelitian ini merupakan penelitian kuantitatif deskriptif dengan pendekatan eksploratif yang memanfaatkan teknik data mining. Tujuan utamanya adalah untuk mengidentifikasi pola perilaku belajar siswa dan memprediksi hasil akademik menggunakan algoritma C4.5 dan Apriori secara terintegrasi.

## **3.2 Lokasi dan Waktu Penelitian**

Studi ini dilaksanakan di SMK Taman Siswa Banjarnegara, sebuah sekolah menengah kejuruan swasta yang terletak di Kabupaten Banjarnegara, Jawa Tengah. Penelitian dilaksanakan selama semester ganjil tahun ajaran 2024/2025, dengan pelaksanaan pengumpulan data pada bulan Oktober hingga Desember 2024.

## **3.3 Populasi dan Sampel**

Populasi dalam penelitian ini adalah seluruh siswa kelas XII SMK Taman Siswa Banjarnegara. Sampel yang digunakan berjumlah 60 siswa, yang dipilih secara purposive berdasarkan kriteria: (1) aktif menggunakan platform pembelajaran digital sekolah, (2) bersedia mengisi kuesioner, dan (3) memiliki nilai akademik semester yang dapat diakses peneliti.

## **3.4 Teknik Pengumpulan Data**

Pengumpulan data dilakukan melalui dua instrumen utama:

1. **Kuesioner:** Berisi pertanyaan tertutup mengenai kebiasaan belajar digital siswa, seperti media yang digunakan, durasi belajar, waktu belajar, dan frekuensi akses platform.
2. **Dokumen Nilai Akademik:** Data nilai rapor semester ganjil yang diperoleh dari wali kelas sebagai acuan variabel dependen (hasil belajar).

### **3.5 Variabel Penelitian**

Variabel dalam penelitian ini dibagi menjadi dua jenis:

1. **Variabel bebas (independen):** media belajar, waktu belajar, durasi belajar harian, dan frekuensi akses.
2. **Variabel terikat (dependen):** nilai akademik, dikategorikan menjadi tinggi, sedang, dan rendah berdasarkan klasifikasi tertentu.

### **3.6 Teknik Analisis Data**

Analisis data dilakukan dalam dua tahap:

1. **Klasifikasi dengan Algoritma C4.5**

Data diklasifikasikan menggunakan algoritma decision tree C4.5 melalui perangkat lunak *Weka*. Tujuannya adalah untuk mengidentifikasi atribut perilaku belajar yang paling berpengaruh terhadap capaian nilai siswa.

2. **Pencarian Pola dengan Algoritma Apriori**

Data perilaku belajar dikonversi ke format transaksi, kemudian dianalisis menggunakan algoritma Apriori untuk menemukan aturan asosiasi yang kuat antar variabel perilaku belajar.

Kriteria minimum support ditetapkan sebesar 30% dan minimum confidence sebesar 70% agar aturan yang ditemukan relevan secara statistik.

### **3.7 Validitas dan Reliabilitas Data**

Uji validitas instrumen dilakukan dengan expert judgement oleh dua orang dosen bidang informatika pendidikan. Reliabilitas diuji menggunakan uji Alpha Cronbach, yang menunjukkan bahwa item dalam kuesioner konsisten dan layak digunakan sebagai instrumen penelitian.

### **3.8 Etika Penelitian**

Seluruh proses pengumpulan data dilakukan dengan menjaga kerahasiaan identitas responden. Partisipasi bersifat sukarela dan telah memperoleh izin dari pihak sekolah serta persetujuan dari siswa sebagai responden.

## **BAB IV – HASIL DAN PEMBAHASAN**

### **4.1 Profil Lokasi Penelitian**

Lokasi penelitian berada di SMK Taman Siswa Banjarnegara, sebuah sekolah kejuruan swasta yang menyediakan beberapa program keahlian seperti Teknik Komputer dan Jaringan, Akuntansi, serta Pemasaran. Fokus utama pengumpulan data ditujukan kepada siswa kelas XII, mengingat kelompok ini telah menjalani sebagian besar proses pembelajaran dan memiliki pengalaman yang lebih matang dalam penggunaan media digital.

Fasilitas teknologi yang tersedia di sekolah ini cukup mendukung proses pembelajaran daring, termasuk ketersediaan jaringan internet, platform Google Classroom, serta sumber pembelajaran berbasis video dari YouTube maupun media internal sekolah.

## **4.2 Deskripsi Data dan Variabel**

Sebanyak 60 siswa kelas XII menjadi responden dalam penelitian ini. Data dikumpulkan melalui angket mengenai kebiasaan belajar digital serta nilai akademik semester ganjil tahun ajaran 2024/2025.

Beberapa variabel yang dikaji antara lain:

1. **Media Pembelajaran:** Video, PDF, kombinasi keduanya.
2. **Durasi Belajar Harian:** Kurang dari 1 jam, lebih dari 1 jam.
3. **Waktu Belajar:** Pagi, siang, atau malam.
4. **Frekuensi Akses Platform Pembelajaran**
5. **Nilai Akademik:** Dikategorikan ke dalam tiga kelompok (tinggi, sedang, rendah).

## **4.3 Klasifikasi Menggunakan Algoritma C4.5**

Data yang telah dikodekan diolah menggunakan perangkat lunak Weka, dengan algoritma C4.5 digunakan untuk membangun model klasifikasi terhadap variabel target “Nilai Akademik”.

### **4.3.1 Ringkasan Struktur Pohon Keputusan**

Beberapa aturan klasifikasi yang dihasilkan oleh model antara lain:

1. Jika siswa belajar lebih dari 1 jam per hari dan menggunakan media video → kemungkinan mendapat nilai tinggi.
2. Jika durasi belajar kurang dari 1 jam dan hanya menggunakan PDF → cenderung memperoleh nilai rendah.
3. Belajar di malam hari dengan durasi lebih dari 1 jam → berkorelasi dengan nilai tinggi.
4. Belajar singkat dengan media campuran → cenderung menghasilkan nilai sedang.

### **4.3.2 Evaluasi Kinerja Model**

Tabel 4. 1 Evaluasi

| <b>Metode Evaluasi</b>   | <b>Hasil</b>            |
|--------------------------|-------------------------|
| Akurasi                  | 81,60%                  |
| Precision (kelas tinggi) | 83,30%                  |
| Recall (kelas tinggi)    | 80,00%                  |
| Confusion Matrix         | Disajikan pada lampiran |

Evaluasi menunjukkan bahwa model memiliki performa cukup baik dalam memprediksi kategori nilai akademik berdasarkan perilaku belajar siswa.

#### 4.4 Pola Asosiasi Berdasarkan Algoritma Apriori

Dataset dikonversi ke format transaksional agar sesuai dengan prinsip kerja algoritma Apriori. Parameter minimum support ditetapkan pada 30% dan minimum confidence sebesar 70%.

##### 4.4.1 Beberapa Pola Asosiasi yang Ditemukan

Tabel 4. 2 Aturan Asosiasi

| No | Aturan Asosiasi                            | Support | Confidence |
|----|--|---------|------------|
| 1  | Belajar malam + media video → nilai tinggi | 36,70%  | 82%        |
| 2  | Durasi belajar >1 jam → nilai tinggi       | 45%     | 77%        |
| 3  | Belajar pagi + media PDF → nilai sedang    | 21,60%  | 66%        |

##### 4.4.2 Interpretasi Temuan

Dari hasil di atas, terdapat kecenderungan bahwa siswa yang mengalokasikan waktu lebih dari satu jam serta memanfaatkan video pembelajaran secara rutin memiliki peluang lebih besar meraih nilai tinggi. Sebaliknya, siswa yang hanya menggunakan PDF dan belajar di pagi hari menunjukkan hasil belajar yang relatif sedang atau kurang.

#### 4.5 Integrasi Temuan dan Diskusi

Kombinasi antara pendekatan klasifikasi dan asosiasi menunjukkan keselarasan dalam hasil:

1. Model C4.5 mengidentifikasi faktor-faktor utama yang mempengaruhi capaian akademik siswa secara individual.
2. Sementara itu, Apriori mampu memetakan pola perilaku belajar yang konsisten dengan hasil klasifikasi, memperkuat temuan secara statistik.

##### Kontribusi Studi Ini:

1. Berbeda dari studi terdahulu yang hanya memakai satu pendekatan algoritmik, penelitian ini memadukan dua metode untuk menyajikan analisis yang lebih komprehensif.
2. Fokus pada konteks lokal (SMK di daerah rural) memberikan nilai tambah dalam hal relevansi praktis.

##### Implikasi Praktis:

1. Guru dan sekolah dapat mendorong strategi belajar yang lebih efektif dengan mengedepankan penggunaan video dan waktu belajar malam.
2. Pembelajaran dapat dipersonalisasi berdasarkan kecenderungan perilaku belajar siswa.

## BAB V – PENUTUP

### 5.1 Kesimpulan

Berdasarkan keseluruhan hasil analisis, dapat disimpulkan bahwa:

1. Algoritma C4.5 mampu mengidentifikasi pengaruh variabel seperti media pembelajaran, waktu belajar, dan durasi belajar terhadap prediksi capaian akademik. Meski model memiliki keterbatasan dalam akurasi (sekitar 38,89%), namun tetap memberikan wawasan awal yang bermakna.
2. Algoritma Apriori secara efektif menemukan keterkaitan antara kebiasaan belajar digital siswa dengan hasil nilai akademik, seperti durasi belajar lebih dari satu jam yang berkorelasi positif dengan nilai tinggi.
3. Integrasi kedua metode dapat menjadi pendekatan yang kuat dalam memetakan karakteristik pembelajaran siswa secara objektif dan efisien di lingkungan SMK.

### 5.2 Keterbatasan Penelitian

Penelitian ini memiliki beberapa batasan yang perlu menjadi perhatian:

1. Jumlah responden yang terbatas (60 siswa) dapat mempengaruhi generalisasi hasil.
2. Data dikumpulkan secara simulatif, bukan dari data aktual platform pembelajaran.
3. Validasi hasil belum dilakukan secara menyeluruh, misalnya melalui uji silang atau uji lapangan.

### 5.3 Rekomendasi

**Untuk Peneliti Berikutnya:**

1. Disarankan menggunakan jumlah sampel yang lebih besar dan data aktual dari platform pembelajaran digital.
2. Perlu mempertimbangkan penambahan variabel seperti partisipasi di kelas, jumlah tugas, jenis kelamin, atau kehadiran.
3. Sebaiknya dilakukan perbandingan performa dengan algoritma lain seperti Random Forest atau Naive Bayes.

**Untuk Sekolah dan Guru:**

1. Pendekatan data mining dapat dimanfaatkan sebagai alat bantu dalam membuat kebijakan pembelajaran yang lebih adaptif.
2. Dapat digunakan untuk mendeteksi siswa yang berisiko mendapatkan hasil rendah dan memberikan intervensi lebih awal.

## REFERENCES

- [1] Agrawal, R., & Srikant, R. (1994). *Fast Algorithms for Mining Association Rules*. Proceedings of the 20th International Conference on Very Large Data Bases (VLDB), 487–499.
- [2] Alamsyah, R., & Cahyani, I. (2021). Penerapan data mining dalam memprediksi prestasi siswa dengan algoritma C4.5. *Jurnal Informatika dan Sistem Informasi*, 7(2), 89–97.
- [3] Bakar, S. A., & Yusof, Z. M. (2016). Predictive modeling for student performance using decision tree algorithms. In *Proceedings of the 2016 International Conference on Computer and Communication Engineering (ICCCCE)*, 366–371.
- [4] Fitriyani, H., & Nugroho, R. P. (2022). Penerapan algoritma Apriori dalam menentukan pola pembelajaran siswa selama pandemi. *Jurnal Pilar Teknologi Informasi*, 17(1), 51–58.



- [5] Han, J., Pei, J., & Kamber, M. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- [6] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- [7] Ifenthaler, D., & Yau, J. Y-K. (2020). Utilization of hybrid learning analytics for learner modeling. *Educational Technology Research and Development*, 68, 1999–2018.
- [8] Jannah, N., & Wibowo, A. (2021). Analisis pola belajar mahasiswa berbasis e-learning menggunakan Apriori. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(3), 543–550.
- [9] Junco, R. (2012). The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. *Computers & Education*, 58(1), 162–171.
- [10] Kusnadi, A., & Wahyudi, R. (2023). Aplikasi data mining dalam pendidikan: kajian literatur dan implementasi. *Prosiding Seminar Nasional Teknologi Informasi*, 14(1), 88–95.
- [11] Lauria, E. J. M., Baron, J. D., Devireddy, M., Sundararaju, V., & Jayaprakash, S. M. (2013). Mining academic data to improve student retention: a comparison of decision trees and logistic regression. *Journal of Learning Analytics*, 1(1), 1–20.
- [12] Lestari, D. N., & Supriyadi, D. (2020). Analisis perilaku belajar siswa dengan metode data mining. *Jurnal Teknologi Informasi dan Pendidikan*, 13(3), 136–143.
- [13] Nurhayati, D., & Mustofa, K. (2020). Prediksi performa siswa dengan educational data mining: studi perbandingan. *Indonesian Journal of Computing and Cybernetics Systems*, 14(2), 133–140.
- [14] Pratiwi, D., & Nugroho, R. (2020). Penggunaan algoritma Apriori untuk menganalisis pola belajar mahasiswa. *Jurnal SIMETRIS*, 11(2), 183–190.
- [15] Prensky, M. (2010). *Teaching Digital Natives: Partnering for Real Learning*. Corwin Press.
- [16] Putra, H. A., & Hermawan, E. (2022). Pemanfaatan algoritma C4.5 untuk klasifikasi tingkat pemahaman siswa pada pembelajaran daring. *Jurnal Teknologi Informasi dan Pendidikan*, 15(1), 70–77.
- [17] Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.
- [18] Rokach, L., & Maimon, O. (2015). *Data Mining with Decision Trees: Theory and Applications* (2nd ed.). World Scientific Publishing.
- [19] Romero, C., & Ventura, S. (2020). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50(9), 3289–3302.
- [20] Suhartono, D., & Ramadhan, I. (2023). Prediksi hasil belajar siswa menggunakan decision tree: studi kasus di SMK. *Journal of Educational Data Mining Indonesia*, 2(1), 33–42.
- [21] Sugiyono. (2018). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- [22] Weka, A., & Septian, H. (2021). Klasifikasi prestasi belajar menggunakan algoritma C4.5. *Jurnal Teknologi dan Sistem Komputer*, 9(2), 117–124. <https://doi.org/10.14710/jtsiskom.9.2.2021.117-124>