

Evaluasi Kinerja Akademik Mahasiswa ITS NU Sriwijaya Menggunakan Data Mining

Eriene Dheanda Absharina¹, Dila Novela^{2*}, Tania Galuh Parwati³, Rinda Antika⁴

^{1,2,3,4}Program Studi Informatika, Institut Teknologi dan Sains Nahdlatul Ulama Sriwijaya, Palembang, Sumatera Selatan, Indonesia

Email: ¹eriene.dheanda@itsnusriwijaya.ac.id, ^{2*}fatzah1031@gmail.com, ³taniaparwati39@gmail.com, ⁴antikarinda6@gmail.com

(* Email Corresponding Author: fatzah1031@gmail.com)

Received: 4 Juni 2025 | Revision: 10 Juni 2025 | Accepted: 10 Juni 2025

Abstrak

Di era saat ini, harapan masyarakat terhadap lembaga pendidikan semakin besar, terutama dalam mencetak generasi yang cerdas, terampil, dan mampu bersaing secara global. Di lingkungan ITS NU Sriwijaya Sumatera Selatan, Indeks Prestasi Kumulatif (IPK) menjadi tolok ukur penting dalam menilai pencapaian akademik mahasiswa. Salah satu faktor yang sering dikaitkan dengan IPK adalah durasi waktu belajar. Namun, dalam praktiknya, cukup banyak mahasiswa yang justru lebih banyak menghabiskan waktu untuk berselancar di internet, khususnya media sosial, yang secara tidak langsung dapat mengganggu fokus belajar mereka. Penelitian ini menggunakan pendekatan data mining dengan algoritma klasifikasi *Decision Tree* untuk mengelompokkan mahasiswa berdasarkan IPK yang diperoleh pada semester sebelumnya. Selain itu, metode statistik *chi-square* digunakan untuk menelusuri faktor-faktor yang berpengaruh terhadap capaian IPK. Hasil penelitian menunjukkan bahwa 84% mahasiswa berada dalam kategori IPK tinggi, sementara 16% sisanya masuk kategori sedang. Aktivitas di media sosial terbukti memiliki hubungan yang signifikan dengan IPK, sedangkan durasi belajar ternyata tidak menunjukkan pengaruh yang berarti.

Kata kunci: Klasifikasi, Prediksi, IPK, Media Sosial, Decision Tree, Chi-square

Abstract

In the current era, public expectations of educational institutions are increasingly high, especially in producing a generation that is intelligent, skilled, and capable of competing globally. In the environment of ITS NU Sriwijaya South Sumatra, the Cumulative Grade Point Average (CGPA) has become an important benchmark in assessing students' academic achievements. One of the factors often associated with GPA is the duration of study time. However, in practice, quite a few students actually spend more time surfing the internet, especially on social media, which can indirectly disrupt their study focus. This research uses a data mining approach with the Decision Tree classification algorithm to group students based on the GPA obtained in the previous semester. Additionally, the chi-square statistical method is used to trace the factors that influence GPA achievement. The research results show that 84% of students fall into the high GPA category, while the remaining 16% fall into the moderate category. Social media activity has been proven to have a significant relationship with GPA, while study duration does not show a meaningful impact.

Keywords: Classification, Prediction, GPA, Social Media, Decision Tree, Chi-square

1. PENDAHULUAN

Kinerja akademik mahasiswa menjadi salah satu indikator utama dalam menilai keberhasilan proses pendidikan tinggi. Hal ini tidak hanya mencerminkan efektivitas sistem pembelajaran, tetapi juga kesiapan mahasiswa dalam menyelesaikan studi tepat waktu sesuai dengan target yang ditetapkan. Indeks Prestasi Kumulatif (IPK) kerap dijadikan tolok ukur utama untuk mengukur capaian akademik mahasiswa. Namun, mengelola data akademik yang sangat besar dan beragam menjadi tantangan tersendiri. Oleh karena itu, diperlukan metode analisis yang efektif untuk mengelompokkan mahasiswa berdasarkan kinerja akademik serta mengidentifikasi faktor-faktor yang berpengaruh terhadapnya. Penyelesaian studi tepat waktu juga penting untuk menghindari penumpukan mahasiswa di semester akhir yang berpotensi menurunkan kualitas pendidikan dan menurunkan rasio kelulusan [1].

Dalam konteks ini, *data mining* menjadi solusi yang tepat untuk mengekstrak pola-pola penting dari data akademik yang besar dan kompleks. Salah satu teknik yang banyak digunakan adalah klasifikasi, yang berfungsi mengelompokkan data berdasarkan karakteristik tertentu. Algoritma *Decision Tree* merupakan salah satu metode klasifikasi yang efektif dan mudah dipahami, karena menghasilkan pohon keputusan yang memandu proses pengambilan keputusan berdasarkan atribut data [2]. Selain itu, analisis statistik seperti uji *chi-square* juga sering dipakai untuk menelusuri hubungan antara berbagai faktor potensial dengan variabel target, dalam hal ini capaian IPK mahasiswa. [3]

Perkembangan teknologi informasi dan perubahan pola belajar mahasiswa turut memberikan pengaruh signifikan terhadap kinerja akademik. Teknologi, terutama media sosial, kini telah menjadi bagian penting dalam kehidupan sehari-hari mahasiswa dan dapat berdampak pada aktivitas belajar mereka. Berdasarkan survei APJII tahun 2016, penggunaan internet dan smartphone di Indonesia terus meningkat pesat, dengan Instagram sebagai salah satu platform media sosial paling populer di kalangan generasi muda [4]. Oleh karena itu, penting untuk memahami sejauh mana aktivitas di media sosial serta faktor lain seperti durasi belajar memengaruhi prestasi akademik mahasiswa. [3]

Berbagai penelitian sebelumnya telah mengaplikasikan *data mining* untuk evaluasi kinerja akademik. Misalnya, Nasir et al. [1] menggunakan algoritma Naive Bayes untuk mengklasifikasikan kinerja mahasiswa, sementara Pratama et al. [5] menerapkan metode *data mining* untuk klasifikasi penerima program bantuan pendidikan. Widaningsih [6] membandingkan beberapa algoritma klasifikasi termasuk *Decision Tree* dan menyoroti kelebihan metode ini dalam konteks prediksi nilai dan kelulusan mahasiswa. Namun, masih sedikit penelitian yang menggabungkan penggunaan algoritma *Decision Tree* dengan analisis statistik chi-square secara bersamaan untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap capaian IPK.

Berdasarkan hal tersebut, penelitian ini bertujuan untuk mengelompokkan mahasiswa Institut Teknologi dan Sains Nahdlatul Ulama Sriwijaya (ITS NU) berdasarkan IPK pada semester sebelumnya menggunakan algoritma *Decision Tree*, sekaligus menelusuri faktor-faktor yang memengaruhi capaian IPK dengan uji chi-square. Diharapkan hasil penelitian ini dapat memberikan wawasan yang bermanfaat untuk pengembangan strategi akademik dan kebijakan pendidikan di institusi tersebut

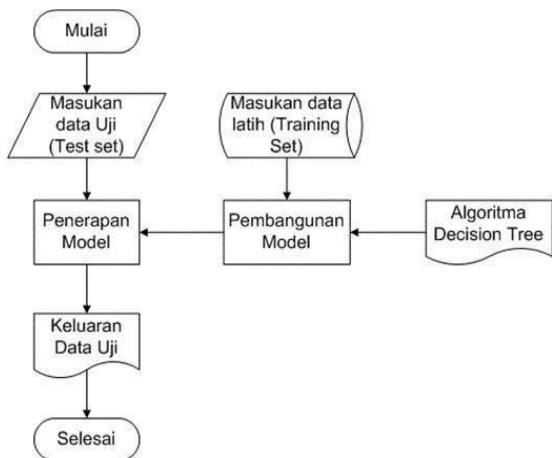
2. METODOLOGI PENELITIAN

2.1 Metode dan Pendekatan Penelitian

Metode penelitian ini menggunakan pendekatan klasifikasi dalam teknik data mining yaitu *Decision Tree* untuk menganalisis pola dan menghasilkan rule yang menjelaskan pencapaian Indeks Prestasi Kumulatif (IPK) mahasiswa [6]. Teknik pengumpulan data yang dilakukan melalui metode survei untuk mendapatkan data primer dari para responden. Sementara itu, untuk mengetahui pengaruh signifikan dari beberapa faktor terutama waktu belajar dan durasi penggunaan media sosial terhadap IPK, digunakan uji statistic chi-square [5]. Dalam penelitian ini, IPK mahasiswa dalam penelitian ini dikelompokkan menjadi tiga tingkat, yaitu :

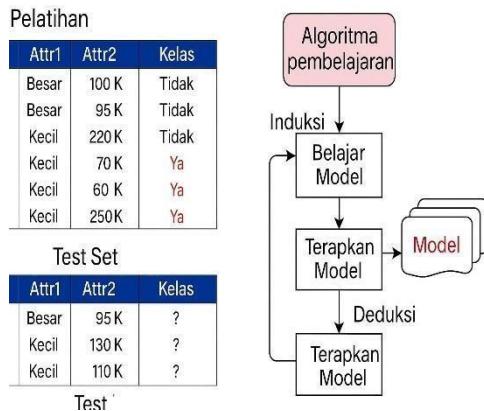
- rendah (<2.50),
- sedang (≥ 2.50 dan <3.50), dan
- tinggi (≥ 3.50).

Proses klasifikasi yang digunakan untuk memprediksi IPK mahasiswa diilustrasikan pada gambar 1. Proses ini dimulai dengan memasukkan data uji yang diperoleh dari data mahasiswa, kemudian diterapkan model klasifikasi yang telah dibentuk menggunakan data latih melalui algoritma *Decision Tree*. Setelah pemodelan, hasil prediksi diperoleh sebagai output [7].



Gambar 1. Flowchart Pengerjaan Klasifikasi Data

Adapun gambar 2 memperlihat alur lengkap proses klasifikasi, mulai dari data latih yang diproses oleh leraning algoritma, membentuk model prediktif, hingga proses implementasi model terhadap data uji.



Gambar 2. Proses Pengklasifikasian Data

2.2 Prosedur Penelitian

a. Survei Responden

Mahasiswa yang menjadi responden dalam penelitian di minta secara sukarela untuk mengisi survei dalam bentuk kuisioner. Pengisian kuisioner dapat dilakukan melalui dua cara, yaitu secara daring daring melalui tautan yang disediakan, atau secara luring dengan mengisi lembar kuisioner yang dibagikan secara langsung. Responden diminta untuk masuk ke link website yang akan disediakan atau mengisi kuisioner secara manual melalui lembaran lembaran yang akan dibagikan.

b. Struktur Kuesioner

Responden diminta untuk menjawab seluruh pertanyaan yang tercantum dalam kuesioner. Agar mempermudah pemahaman dan pengisian, kuesioner dibagi menjadi 2 bagian, yaitu :

1. Bagian pertama, berisi pertanyaan terkait data demografis responden, meliputi jenis kelamin, durasi penggunaan internet, status keaktifan dalam menggunakan media sosial, semester yang ditempuh, IPK terakhir, Keaktifan dalam Unit Kegiatan Mahasiswa (UKM), serta waktu yang dialokasikan untuk belajar.
2. Bagian kedua, berfokus pada pengumpulan data melalui daftar pertanyaan yang harus dijawab oleh responden yang telah ditentukan sebelumnya.

2.3 Analisis Data

Metode analisis yang digunakan dalam penelitian ini adalah algoritma *decision tree c4.5*, yang berfungsi untuk melakukan klasifikasi data, membangun model prediksi, mengidentifikasi pola, dan menghasilkan aturan untuk memprediksi nilai IPK mahasiswa [6], [8] Sementara itu, analisis chi-square diterapkan untuk mengukur kekuatan hubungan antara variabel waktu belajar dan penggunaan media social terhadap nilai IPK mahasiswa.

Jika nilai $p < 0,05$ maka hipotesis alternatif (H_1) diterima. Artinya, terhadap hubungan yang signifikan antara variabel independen (waktu belajar dan penggunaan media sosial) dan variabel independen (IPK). Sebaliknya, jika nilai $p \geq 0,05$ maka hipotesis nol (H_0) diterima. Artinya, tidak terdapat hubungan yang signifikan antara variabel independen dan dependen.

2.4 Hipotesis Penelitian

Dalam penelitian ini menguji dua hipotesis utama :

a. Hipotesis pertama

Hipotesis pertama menyatakan bahwa penggunaan media sosial, khususnya aplikasi Instagram, berdampak negatif terhadap nilai Indeks Prestasi Kumulatif (IPK) mahasiswa. Hal ini didukung oleh penelitian oleh C. Romero dan S. Ventura yang menunjukkan bahwa penggunaan media sosial yang tidak terkontrol dapat mengganggu fokus belajar mahasiswa dan berdampak negatif terhadap performa akademik mereka [7].

b. Hipotesis Kedua

Hipotesis kedua menyatakan adanya hubungan positif antara durasi belajar dengan IPK mahasiswa. Studi yang dilakukan oleh M. Nasir *et al.* menunjukkan bahwa faktor kebiasaan belajar termasuk durasi dan intensitas belajar memiliki pengaruh signifikan terhadap capaian akademik mahasiswa [1].

Variabel independen dalam penelitian ini meliputi penggunaan media sosial (dengan fokus pada platform instagram) dan durasi waktu belajar, sedangkan variabel dependennya adalah IPK mahasiswa. Analisis data akan difokuskan pada pengujian kedua hipotesis tersebut untuk menentukan hubungan kausal antara variabel-variabel yang telah diidentifikasi.

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Dataset

Penelitian ini memanfaatkan dataset yang diperoleh melalui penyebaran angket secara acak kepada mahasiswa semester 3 di Institut Teknologi dan Sains Nahdlatul Ulama Sriwijaya angkatan 2023. Data latih yang digunakan adalah sebanyak 100 data mahasiswa yang terdiri dari 26 mahasiswa memiliki IPK (Indeks Prestasi Kumulatif) kategori Rendah, 29 mahasiswa memiliki IPK kategori Sedang, dan 45 mahasiswa memiliki IPK kategori Tinggi. Sementara itu, 50 set data mahasiswa semester 1 (satu) digunakan sebagai data uji untuk memprediksi kategori IPK-nya [9].

3.2 Proses Preprocessing Data

Tahapan selanjutnya adalah proses pra-pemrosesan data, yaitu mencakup beberapa langkah berikut :

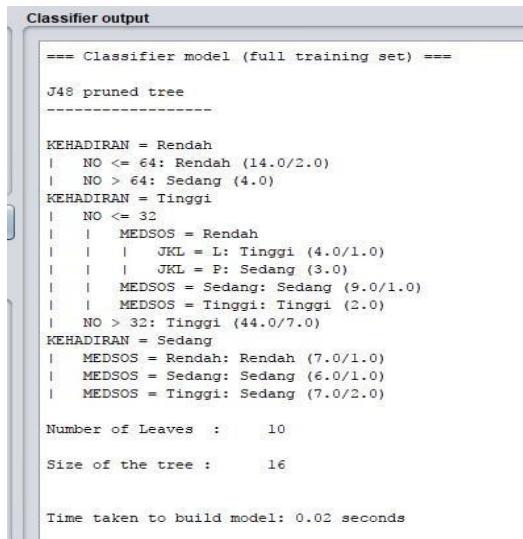
- Pembersih Data (Data Cleaning): Langkah awal dalam preprocessing adalah menghapus data yang mengandung *outlier, noise*, inkonsistensi dan nilai yang hilang (missing values). Dalam dataset pelatihan dan pengujian, atribut seperti Program Studi, Fakultas dan Jenis Media Sosial memiliki nilai hilang, sehingga dihapus untuk mencegah penurunan akurasi model saat evaluasi [10]
- Transformasi Data (Data Information): Pada tahap ini, nilai beberapa atribut numerik diubah menjadi kategori kualitatif seperti Rendah, Sedang, dan Tinggi. Proses ini membantu dalam menyederhanakan data dan mempermudah analisis.
- Pengurangan data (Data Diskritisasi): Semua atribut numerik akan menjalani proses reduksi dan diskritisasi, mengubah data numerik menjadi data kategorikal.

Setelah melalui tahap pra-pemrosesan data, diperoleh sejumlah atribut yang akan digunakan dalam penerapan data minig pada penelitian ini. Atribut-atribut tersebut mencakup jenis kelamin, nilai kehadiran, jam berinternet, waktu belajar, waktu penggunaan media sosial, dan Indeks Prestasi Kumulatif (IPK)[11]

3.3 Pembangunan Model dan Evaluasi

Dalam penelitian ini, model klasifikasi dibangun menggunakan algoritma J48, yang merupakan implementasi dari algoritma C4.5, dalam perangkat lunak WEKA [12]. Data penelitian yang digunakan telah dilengkapi dengan label kelas pada atribut IPK, yang dikategorikan menjadi rendah, sedang, dan tinggi. Setelah model klasifikasi selesai dibangun, tahap selanjutnya adalah dilakukan pengujian dengan menggunakan data uji yang belum memiliki nilai pada atribut IPK. Tahap evaluasi kemudian dilakukan untuk mengetahui tingkat akurasi model dalam memprediksi kategori IPK data uji tersebut [1] [11].

3.3.1 Hasil Klasifikasi Menggunakan Weka



```

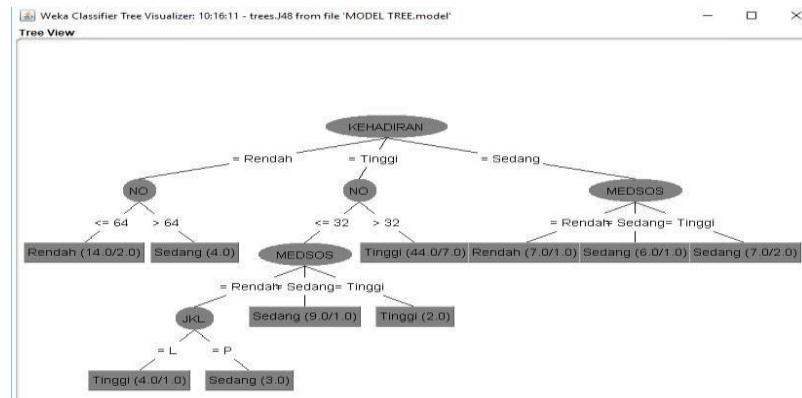
Classifier output
=====
==== Classifier model (full training set) ====
J48 pruned tree
-----
KEHADIRAN = Rendah
| NO <= 64: Rendah (14.0/2.0)
| NO > 64: Sedang (4.0)
KEHADIRAN = Tinggi
| NO <= 32
| | MEDSOS = Rendah
| | | JKL = L: Tinggi (4.0/1.0)
| | | JKL = P: Sedang (3.0)
| | MEDSOS = Sedang: Sedang (9.0/1.0)
| | MEDSOS = Tinggi: Tinggi (2.0)
| NO > 32: Tinggi (44.0/7.0)
KEHADIRAN = Sedang
| MEDSOS = Rendah: Rendah (7.0/1.0)
| MEDSOS = Sedang: Sedang (6.0/1.0)
| MEDSOS = Tinggi: Sedang (7.0/2.0)

Number of Leaves : 10
Size of the tree : 16
Time taken to build model: 0.02 seconds
  
```

Gambar 3. Hasil Klasifikasi Menggunakan WEKA

$$IF KEHADIRAN = RENDAH \wedge NO \leq 64 \rightarrow KINERJA = Rendah \quad (1)$$

Seperi diilustrasikan pada Gambar 3, analisis pengklasifikasi WEKA hanya mengidentifikasi atribut 'Kehadiran', 'Media Sosial', dan 'Gender' untuk dimasukkan dalam pohon keputusan, sedangkan atribut lainnya dihilangkan seluruhnya dari pohon keputusan. Hal ini menunjukkan bahwa berdasarkan kuantitas dan sifat data yang dapat diakses, hanya sejumlah atribut yang diperlukan untuk memperoleh hasil kelas dari kumpulan data. Pohon keputusan yang telah dibuat secara visual digambarkan pada Gambar 4.



Gambar 4. Hasil Analisis WEKA

Akurasi hasil dievaluasi selama fase penilaian. Uji analisis pohon keputusan dengan penyelarasan catatan menunjukkan bahwa kinerja algoritma C4.5 [13]. dalam kasus ini tidak sepenuhnya akurat. Ilustrasi hasil yang diringkas dalam Gambar 5 digunakan untuk menentukan tingkat akurasi klasifikasi. Bagian "Instansi Terklasifikasi" berisi hasil pemrosesan data yang memerlukan perhatian khusus karena menunjukkan apakah hasil pemodelan memuaskan atau tidak.

== Summary ==			
Correctly Classified Instances	85	85	%
Incorrectly Classified Instances	15	15	%
Kappa statistic	0.7643		
Mean absolute error	0.1666		
Root mean squared error	0.2886		
Relative absolute error	38.6597 %		
Root relative squared error	62.2053 %		
Total Number of Instances	100		

Gambar 5. Hasil Summary

Seperti yang dapat dilihat dari ringkasan hasil pengembangan model pada Gambar 7, terdapat 100 contoh secara keseluruhan. Ini menunjukkan bahwa semua data dari dataset uji telah berhasil diproses dan tidak ada informasi yang hilang. Hasil model dari 100 contoh pelatihan yang digunakan untuk membangun pohon dari 50 contoh pengujian telah mencapai tingkat akurasi yang sangat tinggi.

Berdasarkan hasil ringkasan dari proses yang telah dilakukan, diperoleh tingkat akurasi dari algoritma klasifikasi C4.5 sebesar 85%, yang menunjukkan persentase Correctly Classified Instance [11]. Correctly Classified Instance merujuk pada jumlah data yang berhasil diklasifikasikan dengan oleh model yang digunakan. Nilai ini tergolong tinggi dan dapat dianggap cukup baik karena mendekati angka 90%. Sementara itu, 15% sisanya merupakan Incorrectly Classified Instances, yaitu data yang tidak berhasil diklasifikasikan dengan tepat atau mengalami kesalahan dalam proses klasifikasi oleh model yang telah dibangun sebelumnya. Persentase kesalahan ini masih relatif rendah, sehingga hasil klasifikasinya tetap dapat diterima.

Hasil data statistik menunjukkan bahwa 84% mahasiswa memiliki IPK dalam kategori tinggi, sementara 16% lainnya berada dalam kategori sedang. Tidak terdapat mahasiswa dengan IPK pada kategori rendah. Dari keseluruhan responden, 49% berjenis kelamin perempuan dan 51% laki-laki.

3.4 Analisis Hubungan Variabe

Tahapan analisis selanjutnya adalah menguji tingkat signifikansi hubungan antara beberapa variabel, seperti pengaruh media sosial terhadap IPK serta pengaruh waktu belajar terhadap IPK. Dalam penelitian ini, terdapat dua hipotesis yang diuji menggunakan metode Chi-kuadrat :

- a. Hipotesis pertama (H1) menyatakan bahwa penggunaan media sosial berdampak negatif terhadap prestasi akademik (IPK) mahasiswa Institut Teknologi dan Sains Nahdlatul Ulama Sriwijaya Sumsel. Berdasarkan tabel 1 yang menampilkan hasil uji chisquare, diperoleh nilai signifikansi (p-value) sebesar 0,003. Karena nilai tersebut lebih kecil dari 0,05, maka hipotesis H1 diterima. Hal ini menunjukkan bahwa terdapat hubungan antara penggunaan media sosial dengan pencapaian IPK mahasiswa.
- b. Hipotesis kedua (H2) menyatakan bahwa waktu belajar mahasiswa memberikan pengaruh positif terhadap IPK. Masih pada tabel 1, ditunjukkan bahwa nilai p-value untuk variabel waktu belajar adalah sebesar 0,465. Karena nilai ini lebih besar dari 0,05 maka hipotesis nol (H0) diterima. Artinya, tidak terdapat ubungan yang signifikan antara waktu belajar mahasiswa dengan perolehan IPK mereka.

Tabel 1. Hasil uji untuk medkos terhadap IPK

Item	Value	df	Asymp. Sig. (2-sided)
<i>Pearson Chi-Square</i>	15.866	4	.003
<i>Likelihood Ratio</i>	19.917	4	.001
<i>N of Valid Cases</i>		150	

Dalam N kasus yang dapat diterima yang ditemukan di kedua tabel, terdapat 150 titik data yang terlihat jelas, yang menunjukkan bahwa setiap informasi yang dimasukkan telah ditangani secara akurat dan tidak ada titik data yang hilang (*missing values*)

Tabel 2. Waktu belajar terhadap IPK

Item	Value	Df	Asymp. Sig. (2-sided)
<i>Pearson Chi-Square</i>	3.586	4	.465
<i>Likelihood Ratio</i>	3.502	4	.478
<i>N of Valid Cases</i>		150	

Dari hasil studi yang dilakukan, terlihat kalau perkembangan teknologi informasi-terutama media sosial sudah mulai memengaruhi cara belajar anak muda. Pengaruh ini cukup mengganggu, karena fokus mereka yang seharusnya untuk belajar malah terbagi karena terlalu tertarik sama media sosial. Kondisi seperti ini perlu jadi perhatian khusus, terutama bagi pendidik dan orang tua dalam menyikapi perkembangan teknologi saat ini. Data hasil analisis angka yang mengindikasikan adanya pengaruh yang signifikan dan patut diwaspadai. Fenomena serupa kemungkinan besar akan terjadi dilingkungan pendidikan lainnya, mengingat dampak dari perkembangan teknologi saat ini sudah mulai dirasakan secara luas dan memerlukan kewaspadaan dari pihak-pihak di sekitarnya [12].

Dari hasil studi yang dilakukan, terlihat kalau perkembangan teknologi informasi-terutama media sosial sudah mulai memengaruhi cara belajar anak muda. Pengaruh ini cukup mengganggu, karena fokus mereka yang seharusnya untuk belajar malah terbagi karena terlalu tertarik sama media sosial [14]. Kondisi seperti ini perlu jadi perhatian khusus, terutama bagi pendidik dan orang tua dalam menyikapi perkembangan teknologi saat ini. Data hasil analisis angka yang mengindikasikan adanya pengaruh yang signifikan dan patut diwaspadai [15]. Fenomena serupa kemungkinan besar akan terjadi dilingkungan pendidikan lainnya, mengingat dampak dari perkembangan teknologi saat ini sudah mulai dirasakan secara luas dan memerlukan kewaspadaan dari pihak-pihak di sekitarnya.

4. KESIMPULAN

Penggunaan teknik data mining dapat membantu dalam meramalkan klasifikasi IPK mahasiswa. Dalam penelitian ini, hasilnya menunjukkan bahwa 84% mahasiswa memperoleh IPK tinggi, sedangkan 16% lainnya memperoleh IPK sedang. Metrik kinerja dari algoritma yang diterapkan menunjukkan bahwa algoritma C4.5 atau J48 dari perangkat lunak WEKA mencapai akurasi 85%. Hal ini menunjukkan bahwa masih terdapat margin kesalahan sebesar 15% saat

menggunakan algoritma ini. Analisis data yang dilakukan melalui metode Chi-Square menunjukkan adanya hubungan yang signifikan antara waktu yang dihabiskan di media sosial dengan IPK siswa, dengan tingkat probabilitas sebesar 0,003. Sebaliknya, skor probabilitas sebesar 0,465 menunjukkan bahwa tidak ada korelasi yang signifikan antara lama belajar siswa dan nilai IPK mereka. Hasil ini memberikan gambaran bahwa waktu penggunaan media sosial memiliki kontribusi nyata terhadap pencapaian akademik mahasiswa, sementara waktu belajar yang dihabiskan belum tentu berbanding lurus dengan hasil akademik yang dicapai. Temuan ini menjadi dasar untuk memperhatikan kebiasaan digital mahasiswa sebagai salah satu indikator penting dalam proses pembelajaran. Untuk penelitian yang selanjutnya, disarankan agar algoritma C4.5 digunakan untuk memeriksa atribut tambahan dan untuk mengeksplorasi hubungan antara berbagai variabel dengan hipotesis tambahan, sehingga memungkinkan pandangan yang lebih jelas tentang perbedaan dan variasi

REFERENCES

- [1] M. Nasir, V. Verawaty, and V. Sahfitri, "Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier (Studi Kasus: Universitas Bina Darma)," *JIKO (Jurnal Informatika dan Komputer)*, vol. 5, no. 2, pp. 81–88, 2021.
- [2] M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan data mining untuk evaluasi kinerja akademik mahasiswa menggunakan algoritma naive bayes classifier," *jurnal EECCIS*, vol. 7, no. 1, pp. 59–64, 2013.
- [3] A. Nurfaizin, N. Dengan, and H. S. Pakpahan, "Analisis Evaluasi Kinerja Akademik Mahasiswa Program Studi Teknik Informatika Universitas Mulawarman menggunakan Algoritma CART," *Jurnal Rekayasa Teknologi Informasi (JURTI)*, vol. 4, no. 2, pp. 116–124, 2020.
- [4] A. Indonesia, "Penetrasi dan perilaku pengguna internet Indonesia," *APJII-MASTEL (2016). Konklusi survey Ekosistem Device, Network & Applications (DNA)*, 2016.
- [5] F. D. Pratama, I. Zufria, and T. Triase, "Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar," *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 7, no. 1, pp. 77–84, 2022.
- [6] S. Widaningsih, "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*, vol. 13, no. 1, pp. 16–25, 2019.
- [7] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 10, no. 3, p. e1355, 2020.
- [8] R. P. S. Putri and I. Waspada, "Penerapan Algoritma C4. 5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika," *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, vol. 4, no. 1, pp. 1–7, 2018.
- [9] M. S. Mustafa, M. R. Ramadhan, and A. P. Thenata, "Implementasi data mining untuk evaluasi kinerja akademik mahasiswa menggunakan algoritma naive bayes classifier," *Creative Information Technology Journal*, vol. 4, no. 2, pp. 151–162, 2018.
- [10] C. R. P. Mk, S. J. A. Sumarauw, and V. E. Regar, "Data Mining Model Klasifikasi Untuk Ketepatan Waktu Kelulusan Mahasiswa," *Dharmas Education Journal (DE_Journal)*, vol. 5, no. 1, pp. 555–562, 2025.
- [11] A. Ridwan *et al.*, *Pengantar Sistem Digital: Konsep Dasar dan Penerapannya*. PT. Green Pustaka Indonesia, 2024.
- [12] A. K. Virgiawan, E. D. Absharina, and F. Fenando, "PERAN BIG DATA DALAM MENINGKATKAN DAYA SAING BISNIS DI ERA DIGITAL," *Simtek: jurnal sistem informasi dan teknik komputer*, vol. 10, no. 1, pp. 15–18, 2025.
- [13] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, and M. Data, "Practical machine learning tools and techniques," in *Data mining*, Elsevier Amsterdam, The Netherlands, 2005, pp. 403–413.
- [14] E. D. Absharina, F. K. Wijaya, and I. D. Jaya, "Analisis Kesuksesan Sistem Informasi E-KKN LP2M UIN Raden Fatah Palembang dengan Menggunakan Model Delone dan Mclean," *Jurnal Intelektualita: Keislaman, Sosial, Dan Sains*, vol. 7, no. 2, pp. 217–228, 2018.
- [15] N. W. Purnawati *et al.*, *Sistem Informasi: Teori dan Implementasi Sistem Informasi di berbagai Bidang*. PT. Sonpedia Publishing Indonesia, 2024