



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol.4 No.4 (2025) pp: 53-67

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Penerapan Data Mining Menggunakan Algoritma Decision Tree Untuk Menganalisis Penggunaan Media Sosial Dengan Konsentrasi Belajar Mahasiswa

Kharisma Istiqomah¹, Verra Sofica²

^{1,2}Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

¹kharismaistiqomah19@gmail.com, ²Verra.vsc@bsi.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis antara penggunaan media sosial dengan konsentrasi belajar. Metode penelitian menggunakan pendekatan kuantitatif dengan teknik data mining berbasis algoritma Decision Tree. melibatkan 200 mahasiswa aktif berusia 18–25 tahun yang dipilih secara acak, mengenai frekuensi, durasi, platform media sosial, waktu belajar, IPK, dan konsentrasi belajar. Data kemudian diolah menggunakan perangkat lunak RapidMiner melalui tahapan preprocessing, pemodelan Decision Tree, serta evaluasi model menggunakan cross-validation. Hasil penelitian menunjukkan bahwa frekuensi dan durasi penggunaan media sosial merupakan faktor dominan yang memengaruhi konsentrasi belajar mahasiswa. Mahasiswa dengan frekuensi akses tinggi (>5 kali/hari) dan durasi penggunaan lama (>3 jam/hari) cenderung memiliki konsentrasi belajar rendah. Model Decision Tree yang dibangun menghasilkan akurasi sebesar 63%, dengan pola prediksi yang menunjukkan keterkaitan signifikan antara kebiasaan penggunaan media sosial dan tingkat konsentrasi belajar. Kesimpulannya, algoritma Decision Tree dapat menjadi metode yang efektif untuk mengidentifikasi pola-pola digital yang berpengaruh terhadap konsentrasi mahasiswa, serta memberikan rekomendasi strategis bagi institusi pendidikan dalam mengelola kebiasaan digital mahasiswa. Kesimpulan berikut dapat diambil berdasarkan temuan penelitian tentang penggunaan penambangan data dan algoritma Pohon Keputusan untuk menganalisis antara penggunaan media sosial dan kemampuan siswa untuk fokus selama kelas, bahwa Analisis Penggunaan Media Sosial dengan Konsentrasi Belajar. Temuan analisis Pohon Keputusan menunjukkan bahwa variabel utama yang memengaruhi kemampuan siswa untuk fokus saat belajar adalah frekuensi dan durasi penggunaan media sosial mereka.

Kata kunci: Data Mining, Decision Tree, Media Sosial, Konsentrasi Belajar, Mahasiswa, RapidMiner.

1. Latar Belakang

Kemajuan teknologi informasi telah memberikan dampak yang mendalam terhadap pendidikan serta aspek-aspek lain dalam kehidupan sehari-hari. Sebagai anggota generasi digital native, siswa menggunakan situs media sosial seperti Instagram, TikTok, dan Twitter untuk berkomunikasi dan belajar, selain untuk bersenang-senang. Di sisi lain, penggunaan media sosial yang berlebihan dapat mengganggu fokus.

Besaran penggunaan media sosial TikTok dan tingkat fokus belajar mahasiswa Fakultas Ushuluddin Adab dan Dakwah IAIN Parepare saling berkaitan, menurut penelitian Herman dan Tasya Safitri. (Safitri, 2024). Temuan penelitian menunjukkan bahwa tingkat konsentrasi belajar siswa menurun seiring meningkatnya intensitas penggunaan TikTok.

Penggunaan media sosial ditemukan berkorelasi negatif dengan produktivitas belajar siswa dalam studi lain oleh Rara Indiani dkk., khususnya dalam hal manajemen waktu yang kurang efisien. (Indiani, 2025). Hal ini menunjukkan bahwa meskipun media sosial dapat menjadi alat pendidikan, penggunaannya yang berlebihan dapat mengganggu fokus dan hasil akademis.

Kemampuan untuk memfokuskan perhatian pada tugas tertentu selama jangka waktu tertentu dikenal sebagai konsentrasi. Dalam konteks pembelajaran, konsentrasi sangat penting untuk memahami materi, menyelesaikan tugas,

dan mencapai prestasi akademik yang optimal. Faktor-faktor yang mempengaruhi konsentrasi belajar antara lain lingkungan belajar, kondisi fisik dan mental, serta penggunaan teknologi.

Penelitian oleh St. Syakirah Halim et al menunjukkan bahwa kualitas tidur yang buruk, yang seringkali disebabkan oleh penggunaan media sosial yang berlebihan, berhubungan dengan penurunan konsentrasi belajar pada siswa (Halim, 2024). Selain itu, penggunaan smartphone yang berlebihan juga ditemukan memiliki hubungan negatif dengan konsentrasi belajar siswa, seperti yang ditunjukkan dalam studi oleh Luh Putu Marhaeni Hal ini menegaskan bahwa faktor-faktor eksternal seperti penggunaan media sosial dan smartphone dapat mempengaruhi konsentrasi belajar mahasiswa (Marhaeni, Luh Putu, Putu Budi Adnyana, 2020).

Praktik mengekstrak informasi dan pola yang berharga dari kumpulan data besar dikenal sebagai penambangan data. Dalam konteks pendidikan, data mining dapat digunakan untuk menganalisis perilaku belajar mahasiswa, memprediksi prestasi akademik, dan mengidentifikasi faktor-faktor yang mempengaruhi konsentrasi belajar. Salah satu teknik penambangan data terbaik untuk klasifikasi dan prediksi dalam data pendidikan adalah algoritma pohon keputusan, khususnya C4.5.

Penelitian oleh Uci Suriani (2024) menunjukkan bahwa algoritma decision tree C4.5 dapat memprediksi tingkat kelulusan mahasiswa dengan akurasi sebesar 99,64%. Hal ini menunjukkan potensi besar penggunaan data mining dalam menganalisis dan meningkatkan proses pembelajaran di perguruan tinggi.

Studi lain oleh Daniel Sinaga menerapkan algoritma decision tree C4.5 untuk mengklasifikasikan mahasiswa berpotensi drop out di Universitas Advent Indonesia (Sinaga, Daniel, Edwin J. Solaiman, 2021). Hasil penelitian menunjukkan bahwa metode ini dapat membantu institusi pendidikan dalam mengidentifikasi mahasiswa yang berisiko tinggi untuk tidak menyelesaikan studi tepat waktu.

Penelitian yang secara tepat menganalisis hubungan antara penggunaan media sosial dan konsentrasi belajar siswa menggunakan pendekatan penambangan data masih kurang, meskipun faktanya banyak penelitian telah mengkaji dampak media sosial terhadap prestasi akademik. Mengingat seberapa sering siswa menggunakan media sosial, penting untuk memahami bagaimana hal ini memengaruhi kemampuan mereka untuk fokus saat belajar..

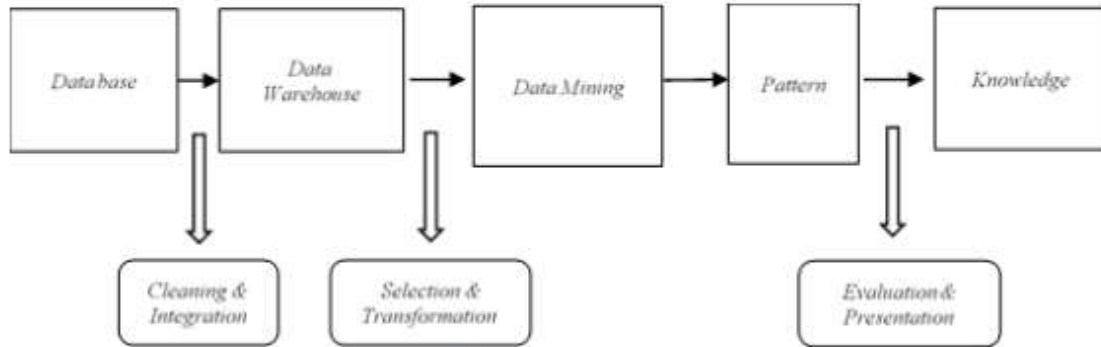
Saat data penggunaan media sosial dan konsentrasi belajar dianalisis, algoritma pohon keputusan dapat digunakan untuk menemukan tren yang terlewatkan oleh analisis tradisional. Hal ini dapat memberikan wawasan baru bagi institusi pendidikan dalam merancang strategi untuk meningkatkan konsentrasi dan prestasi belajar mahasiswa Sistem Informasi, sehingga dengan ini peneliti merumuskan judul penelitian yaitu: Penerapan Data Mining Dengan Algoritma Decision Tree Untuk Menganalisis Hubungan Penggunaan Media Sosial Dan Konsentrasi Belajar Mahasiswa

2. Kajian Teoritis

2.1 Data Mining

Proses pemeriksaan sejumlah besar data untuk mengidentifikasi korelasi bermakna yang memungkinkan penarikan kesimpulan praktis dikenal sebagai penambangan data. Ketika jumlah data sangat besar, penerapan data mining menjadi penting guna memperoleh informasi baru yang nantinya bisa dimanfaatkan sebagai dasar pengambilan keputusan. Data mining juga dikenal dengan berbagai istilah lain, seperti pattern analysis, data archaeology, data dredging, maupun knowledge discovery.

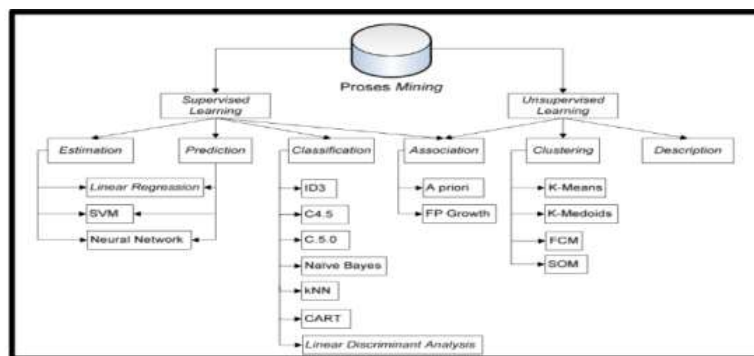
Menurut Han, alur penambangan data adalah prosedur berulang dengan beberapa fase utama, seperti terlihat pada gambar di bawah. (Han, 2012).



Gambar 1. Langkah-Langkah Pengolahan data

Data mining menyediakan sejumlah metode pengolahan data yang dirancang untuk tujuan serta kebutuhan yang berbeda. Beberapa teknik yang sering dimanfaatkan dalam proses data mining dapat dijelaskan sebagai berikut:

1. **Teknik Regression/Estimation** digunakan untuk memperkirakan nilai target dengan tipe numerik dari sebuah titik data. Di antara algoritma yang sering diterapkan pada metode ini yaitu *Linear Regression* serta *Logistic Regression*.
2. **Pendekatan penambangan data** yang disebut prediksi menggunakan variabel-variabel yang sudah diketahui dapat mengungkap pola-pola tersembunyi dalam data. Nilai atau klasifikasi variabel yang tidak memiliki informasi yang tersedia kemudian diprediksi menggunakan pola-pola ini.
3. **Dalam penambangan data**, klasifikasi adalah metode prediktif untuk mengelompokkan data ke dalam klasifikasi yang telah ditentukan sebelumnya. K-Nearest Neighbor (KNN), model Bayesian, jaringan saraf tiruan, dan pohon keputusan adalah beberapa contoh metode kategorisasi yang populer.
4. **Analisis asosiasi** pada data mining digunakan untuk menggali keterkaitan antar item dalam sekumpulan data transaksi. Teknik ini umumnya didukung oleh algoritma *FP-Growth* maupun *Apriori*.
5. Metode **klasterisasi** digunakan untuk membagi dataset menjadi sejumlah kelompok (cluster) yang memiliki karakteristik serupa berdasarkan atributnya. Contoh algoritma yang banyak diterapkan adalah *K-Means* dan *Density-Based Clustering*.
6. **Description** dalam data mining dipakai untuk memperdalam pemahaman terhadap data sehingga perilaku dan karakteristiknya dapat diketahui dengan lebih baik. Ridwan dkk. menyebutkan sejumlah metode yang dapat dimanfaatkan dalam pendekatan deskriptif ini, sebagaimana ditunjukkan pada gambar berikut.



Gambar 2. Teknik Pengolahan Data Mining

2.2 Pohon Keputusan (Decision Tree)

Pohon keputusan adalah metode analisis data berbasis kecerdasan buatan yang menyerupai pohon. Daun pohon melambangkan pilihan akhir, sementara cabang-cabangnya melambangkan berbagai pilihan. Kemampuan pohon keputusan untuk menggabungkan pola data atau informasi ke dalam struktur pohon yang sederhana telah membuatnya populer. Komponen utamanya terdiri dari simpul-simpul (node) yang dihubungkan oleh cabang-cabang, yang dapat diklasifikasikan menjadi tiga kategori.:

1. Root node, merupakan simpul yang tidak memiliki input tetapi memiliki output yang lebih dari satu.
2. Internal node, memiliki satu input dan memiliki output lebih dari dua
3. Leaf atau terminal node, mempunyai satu input dan tidak mempunyai output.

Setiap daun pada pohon keputusan mewakili kelas yang berbeda, dan simpul internal dan akar berfungsi untuk memuat aturan yang membedakan data dengan berbagai kualitas.

2.3 Algoritma Decision Tree

Pohon keputusan dapat dibuat menggunakan berbagai metode, dengan C4.5 menjadi salah satu yang paling banyak digunakan. Ross Quinlan menciptakan metode ini sebagai penyempurnaan dari ID3. Prosedur C4.5 menciptakan struktur pohon keputusan menggunakan data pelatihan. Metode ini memiliki kelebihan, antara lain dapat menangani *missing value*, bekerja dengan data numerik maupun kategorikal, serta mampu mengolah data yang mengandung noise. Tahapan pembentukan Decision Tree dengan C4.5 dijelaskan sebagai berikut .

1. Langkah awal dalam algoritma C4.5 adalah menyiapkan **data training**, yaitu data historis yang sebelumnya telah dikumpulkan dan dikelompokkan ke dalam kelas-kelas tertentu sesuai kebutuhan analisis.
2. Langkah berikutnya adalah menentukan **root (akar) pohon keputusan**. Akar ditentukan berdasarkan atribut yang memiliki nilai *gain* tertinggi dari hasil perhitungan semua atribut. Sebelum memperoleh nilai *gain*, terlebih dahulu dihitung nilai *entropy* dari atribut dengan menggunakan rumus tertentu:

$$\text{Entropy (N)} = \sum_i^p = 1 \text{ pi} * \log_2 \text{ pi}$$

Dimana :

N : Himpunan kasus
 P : Jumlah partisi N
 Pi : Proporsi dari Ni terhadap N

3. Setelah menghitung entropy setiap kasus maka selanjutnya menghitung nilai gain menggunakan rumus sebagai berikut :

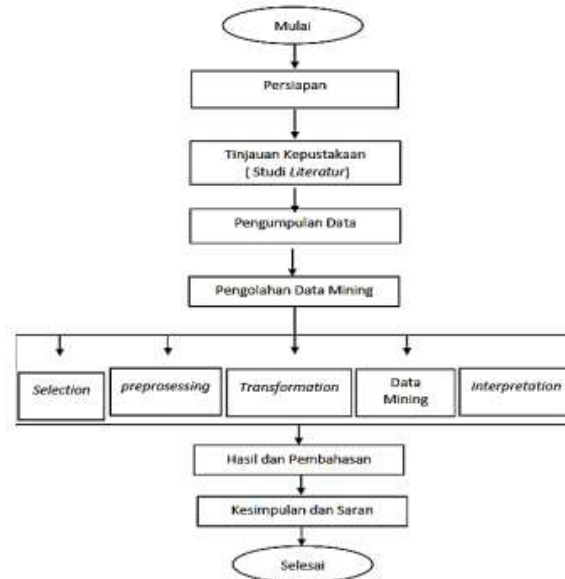
$$\text{Gain (N,a)} = \text{Entropy (N)} - \sum_i^p \frac{N_i}{N} * \text{Entropy (N}_i)$$

Dimana :

N : Himpunan kasus
 A : Atribut
 P : Jumlah partisi atribut A
 |Ni| : jumlah kasus pada partisi ke i
 |N| : Jumlah kasus dalam N

4. Setelah selesai lalu ulangi langkah ke 2 dan ke 3 sampai semua atribut dihitung dan memiliki informasi dan yang akan dijadikan sebagai node (akar)

3. Metode Penelitian



Gambar 3
Alur Penelitian

3.1 Pendekatan Penelitian

Studi ini mengkaji hubungan antara penggunaan media sosial dan kemampuan siswa untuk fokus pada studi mereka menggunakan metodologi kuantitatif dan teknik penambangan data, khususnya algoritma Pohon Keputusan. Metode ini dipilih karena kemampuannya menemukan pola dalam kumpulan data yang luas dan rumit, sehingga secara metodis dan tidak memihak menentukan elemen-elemen yang memengaruhi kemampuan siswa untuk fokus selama pembelajaran. (Han, J., Kamber, M., & Pei, 2023).

3.2 Populasi dan Sampel

Populasi penelitian ini adalah data acak seluruh mahasiswa yang terdaftar pada tahun ajaran 2025. Pengambilan sampel dilakukan secara purposif, dan persyaratan berikut terpenuhi: mahasiswa harus terdaftar di kelas, memiliki akun media sosial, dan siap menjawab pertanyaan tentang kebiasaan belajar, tingkat fokus, serta frekuensi dan lama penggunaan media sosial. Jumlah sampel ditetapkan sebanyak **200 mahasiswa**, sesuai kapasitas pengolahan data pada aplikasi RapidMiner dan untuk memastikan hasil analisis dapat merepresentasikan populasi secara memadai (Sugiyono, 2019).

3.3 Variabel Penelitian

Variabel dalam penelitian ini terdiri dari:

3.3.1 Variabel bebas (independen):

1. Frekuensi penggunaan media sosial (jumlah akses per hari)
2. Durasi penggunaan media sosial (lama penggunaan per hari)
3. Platform favorit media sosial (misal: Instagram, YouTube, Facebook)

3.3.2 Variabel terikat (dependen):

Konsentrasi belajar mahasiswa (dikelompokkan dalam kategori rendah, sedang, tinggi)

3.3.3 Variabel kontrol:

IPK mahasiswa, waktu belajar per hari

Pengelompokan data ini bertujuan agar algoritma Decision Tree dapat mengenali pola dan membuat aturan keputusan berdasarkan kombinasi variabel input terhadap output yang diinginkan (H. et Al, n.d.)

3.4 Instrumen Penelitian

Instrumen penelitian berupa **kuesioner daring** yang berisi pertanyaan tertutup mengenai:

1. Frekuensi harian penggunaan media sosial.
2. Platform media sosial favorit.
3. Durasi harian penggunaan media sosial.
4. Waktu belajar harian.
5. IPK mahasiswa.

Data ini kemudian dikodekan ke dalam format **dataset numerik dan kategorikal**, siap untuk diolah menggunakan perangkat lunak **RapidMiner**.

3.5 Prosedur Pengumpulan Data

1. **Pengumpulan data:** Data yang masuk dikompilasi dalam format Excel (.xlsx) dengan kolom: ID Mahasiswa, Frekuensi_Harian, Platform_Favorit, Durasi_Harian, Waktu_Belajar, IPK, dan Konsentrasi_Belajar.
2. **Pembersihan data:** Data yang tidak lengkap atau tidak valid dihapus atau diperbaiki sebelum dimasukkan ke RapidMiner.

3.6 Teknik Analisis Data

Analisis data dilakukan menggunakan **data mining** dengan algoritma **Decision Tree** di aplikasi **RapidMiner**. Tahapan analisis meliputi:

1. **Import data:** Dataset Excel dimasukkan ke dalam RapidMiner.
2. **Preprocessing:** Melakukan konversi data kategorikal menjadi numerik jika diperlukan, menangani missing value, dan normalisasi data.
3. **Pemodelan Decision Tree:** Menggunakan operator **Decision Tree** di RapidMiner untuk membuat pohon keputusan.
4. **Evaluasi model:** Menggunakan metode **cross-validation** untuk mengevaluasi akurasi model dan menentukan kekuatan prediksi variabel bebas terhadap konsentrasi belajar mahasiswa.
5. **Interpretasi hasil:** Pohon keputusan yang terbentuk dianalisis untuk mengetahui pola penggunaan media sosial yang mempengaruhi konsentrasi belajar.

3.7 Validitas dan Reliabilitas Data

1. **Validitas:** Kuesioner diuji validitasnya menggunakan metode **Pearson Correlation** untuk memastikan setiap item pertanyaan relevan terhadap variabel konsentrasi belajar.
2. **Reliabilitas:** Uji reliabilitas menggunakan **Cronbach's Alpha**, dengan nilai $\geq 0,7$ dianggap memadai untuk penelitian kuantitatif.

Diyakini bahwa dengan mengambil pendekatan ini, penelitian akan memberikan gambaran menyeluruh tentang bagaimana penggunaan media sosial oleh siswa memengaruhi kemampuan mereka untuk berkonsentrasi di kelas dan memberikan rekomendasi berbasis bukti untuk strategi guna meningkatkan konsentrasi belajar.

4. Hasil Dan Pembahasan

4.1 Hasil Analisis Data

Dalam penelitian ini, perancangan model dilakukan dengan memanfaatkan perangkat lunak RapidMiner. Proses pembangunan model pada RapidMiner melibatkan penggunaan sejumlah operator, di antaranya sebagai berikut.

1. File Excel yang berisi kumpulan data yang telah diproses sebelumnya dapat dibaca menggunakan operator Baca Excel.
2. Proses pemisahan data pelatihan dan pengujian melalui perbandingan dan dengan menerapkan teknik pembagian sampel linear dikenal sebagai validasi operator.
3. Pengolahan data dilakukan dengan menggunakan Algoritma Operator Pohon Keputusan.
4. Model algoritma diterapkan menggunakan operator Terapkan Model.
5. Kinerja pemrosesan data algoritma Pohon Keputusan diputar dan dilihat menggunakan Kinerja Operator..

4.1.1 Deskripsi Dataset

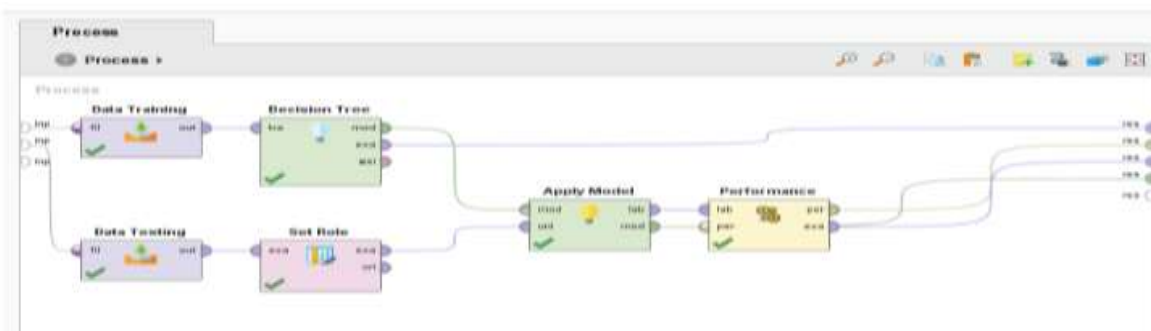
Dataset yang digunakan dalam penelitian ini terdiri dari 200 mahasiswa aktif yang memenuhi kriteria inklusi. Variabel yang dikumpulkan meliputi:

1. **Frekuensi Penggunaan Media Sosial:** Jumlah akses per hari.
2. **Durasi Penggunaan Media Sosial:** Lama penggunaan per hari.
3. **Platform Media Sosial Favorit:** Platform yang paling sering digunakan.
4. **Waktu Belajar Harian:** Durasi waktu yang dihabiskan untuk belajar.

Data ini diolah menggunakan **RapidMiner Studio** versi 9.10.8, dengan tahapan sebagai berikut (Muhammad Firdaus Mustapha, 2023):

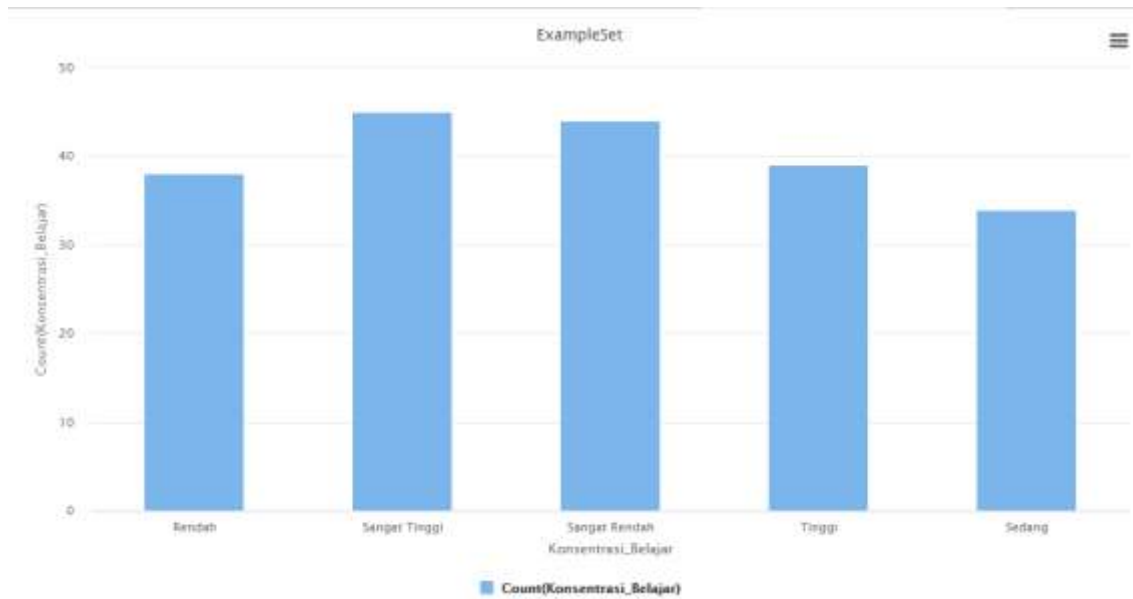
1. **Import Data:** Mengimpor dataset dalam format Excel.
2. **Preprocessing:** Menangani missing values dan normalisasi data.
3. **Modeling:** Membangun model Decision Tree menggunakan operator *Decision Tree*.
4. **Evaluation:** Mengevaluasi model menggunakan cross-validation.

4.1.2. Proses Olah Data



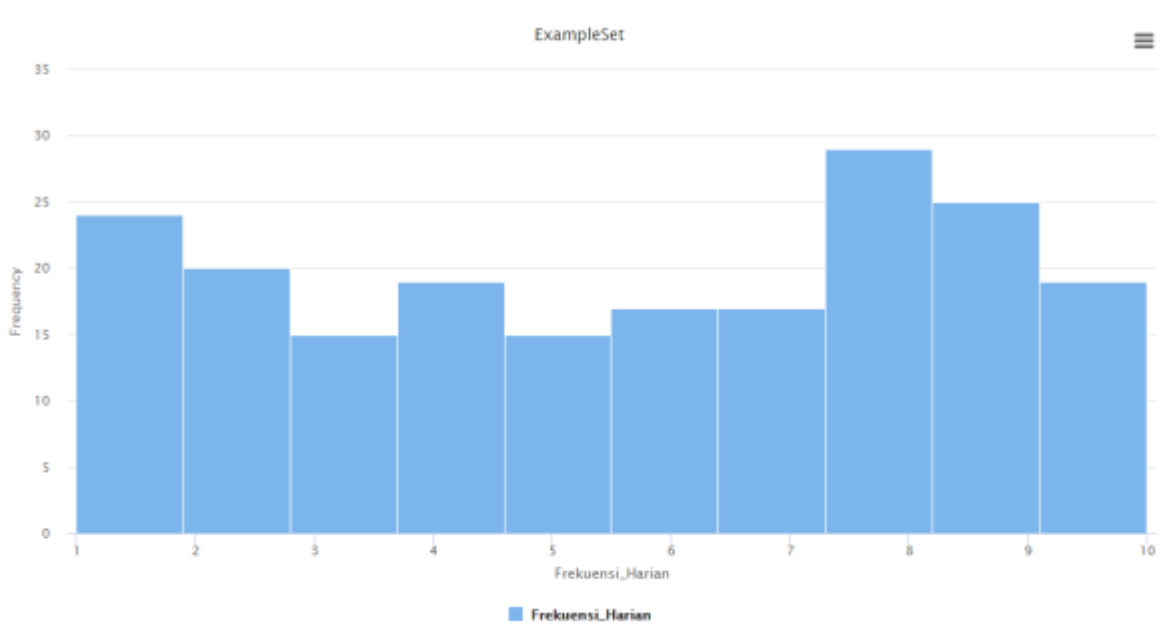
Gambar 4. Proses Olah Data dengan Repidminer

Proses olah data sebagaimana gambar di atas menghasilkan pohon keputusan dan persentase klasifikasi. Ini konsisten dengan alur dasar di RapidMiner: training model → apply model → performance evaluation (confusion matrix, metrik) .



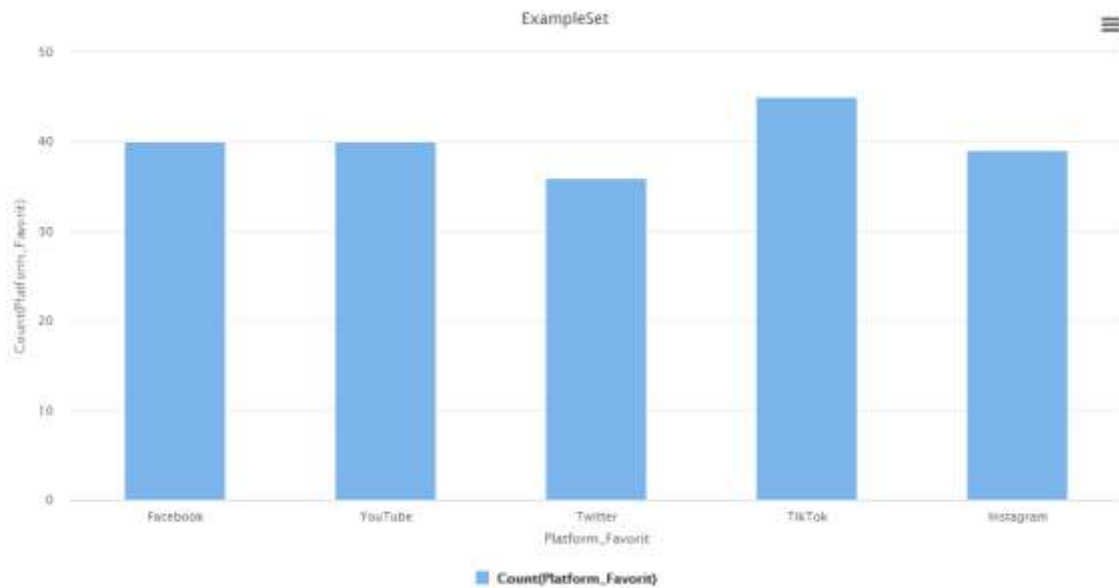
Gambar 5. Visualisasi Hasil olah Data pada Konsentrasi Belajar

Gambar di atas menunjukkan bahwa pada variable konsentrasi belajar diperoleh angka sangat tinggi (45), sangat rendah (44), tinggi (39), rendah (38), sedang (34). Dengan demikian, dapat disimpulkan bahwa paling tinggi untuk konsentrasi belajar pada mahasiswa sebanyak 45.



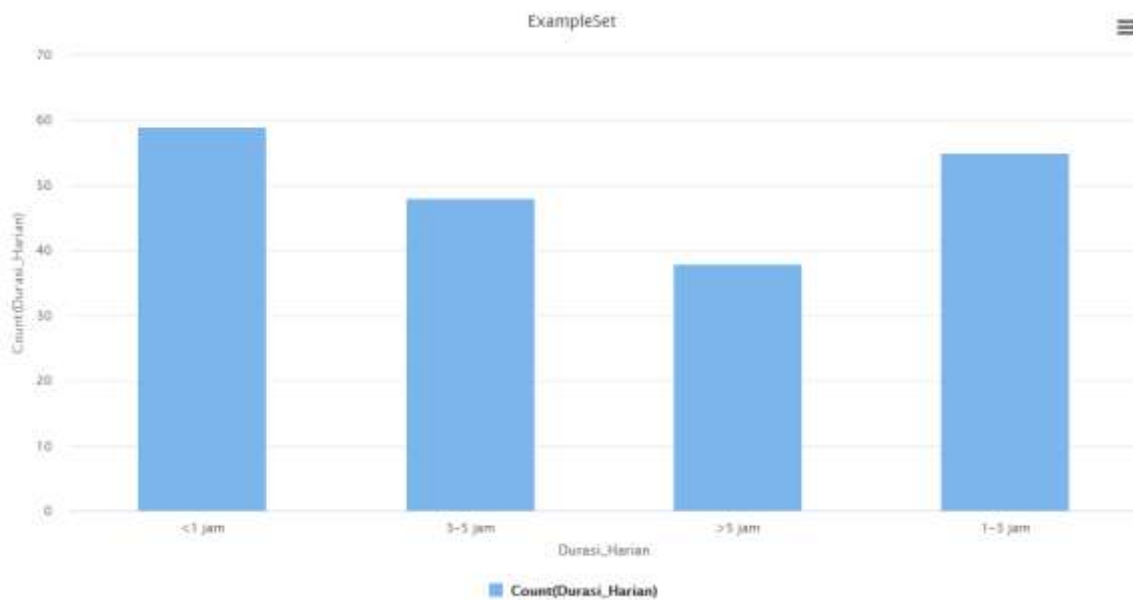
Gambar 6. Visualisasi Hasil olah Data pada Frekuensi Harian

Gambar di atas menunjukkan bahwa pada variable Frekuensi Harian diperoleh hasil minimal 15 kali dan paling banyak 29 kali.



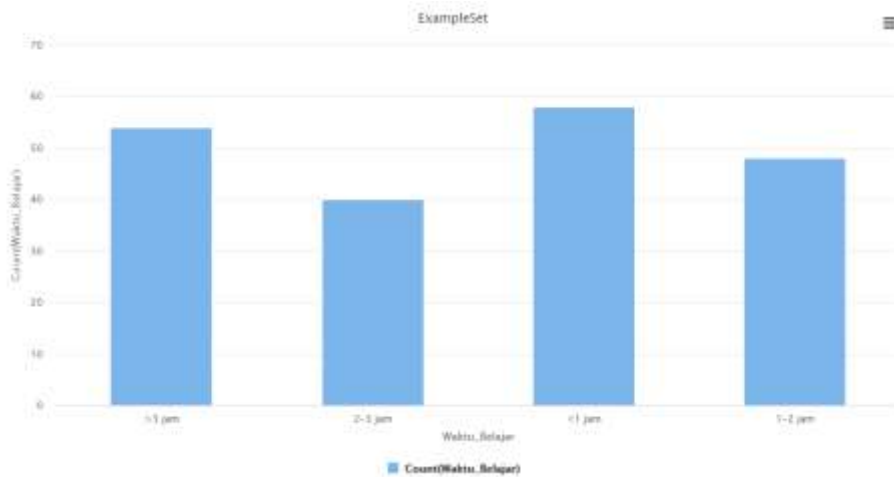
Gambar 7. Visusalisasi Hasil olah Data pada Platform Favorit

Gambar di atas menunjukkan bahwa pada variable Platform Favorit diperoleh hasil jumlah kunjungan sosmed yaitu Tiktok (45), Facebook (40), You Tube (40), Instagram (39), Twitter (36). Dengan demikian sosmed paling sering dikunjungi adalah Tiktok.



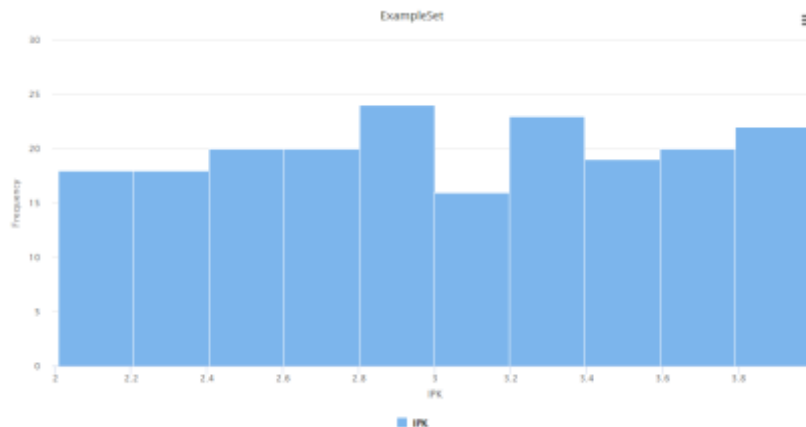
Gambar 8. Visusalisasi Hasil olah Data pada Durasi Harian

Gambar di atas menunjukkan bahwa pada variable Durasi Harian diperoleh perkiraan waktu 1 jam (59), 1-3 jam (55), 3-5 jam (48), >5 jam (38).



Gambar 9. Visualisasi Hasil olah Data pada Waktu Belajar

Gambar di atas menunjukkan bahwa pada variable Waktu Belajar diperoleh perkiraan waktu 1 jam (58), 3 jam (54), 1-2 jam (48), 2-3 jam (40).

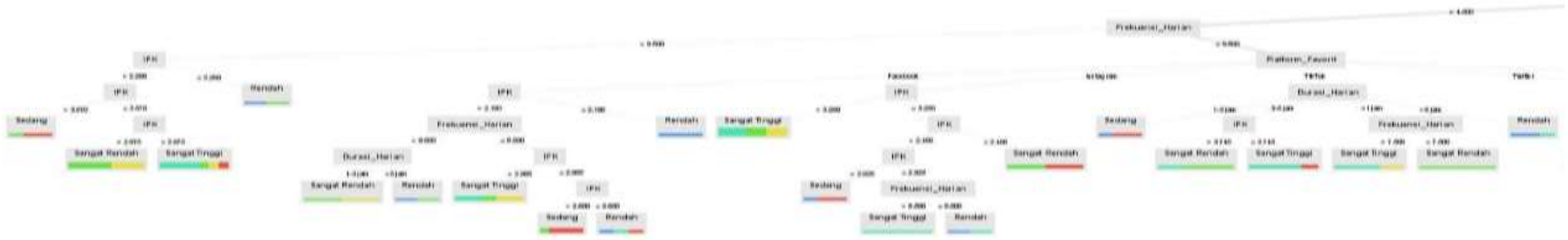


Gambar 10. Visualisasi Hasil olah Data pada IPK

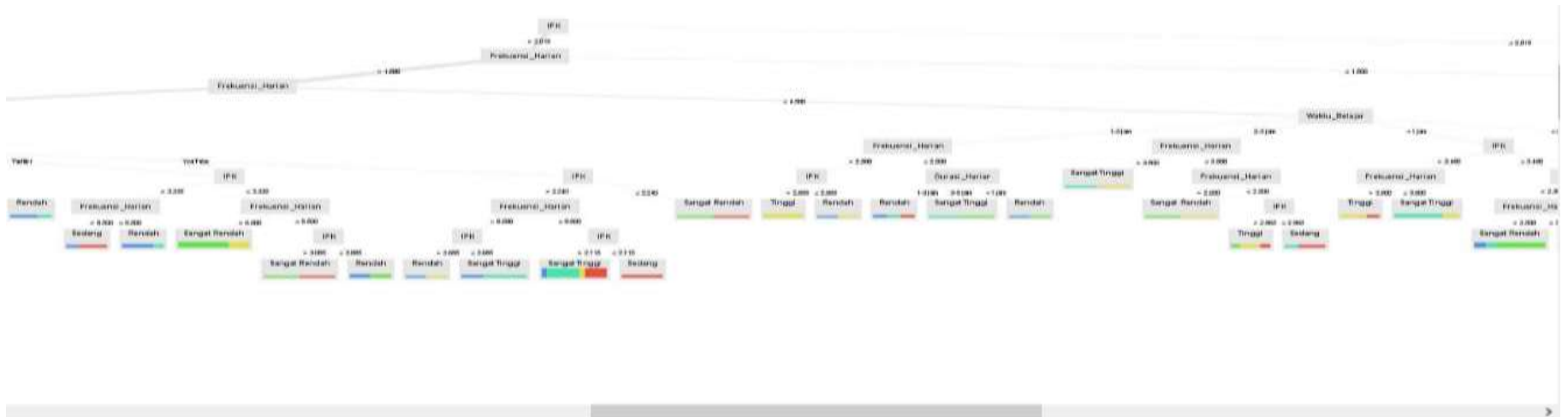
Gambar di atas menunjukkan bahwa pada variable IPK diperoleh mahasiswa paling sedikit mendapat IPK 2.0 sedangkan untuk ipk paling tinggi di IPK 3.9.

4.1.3 Struktur Pohon Keputusan

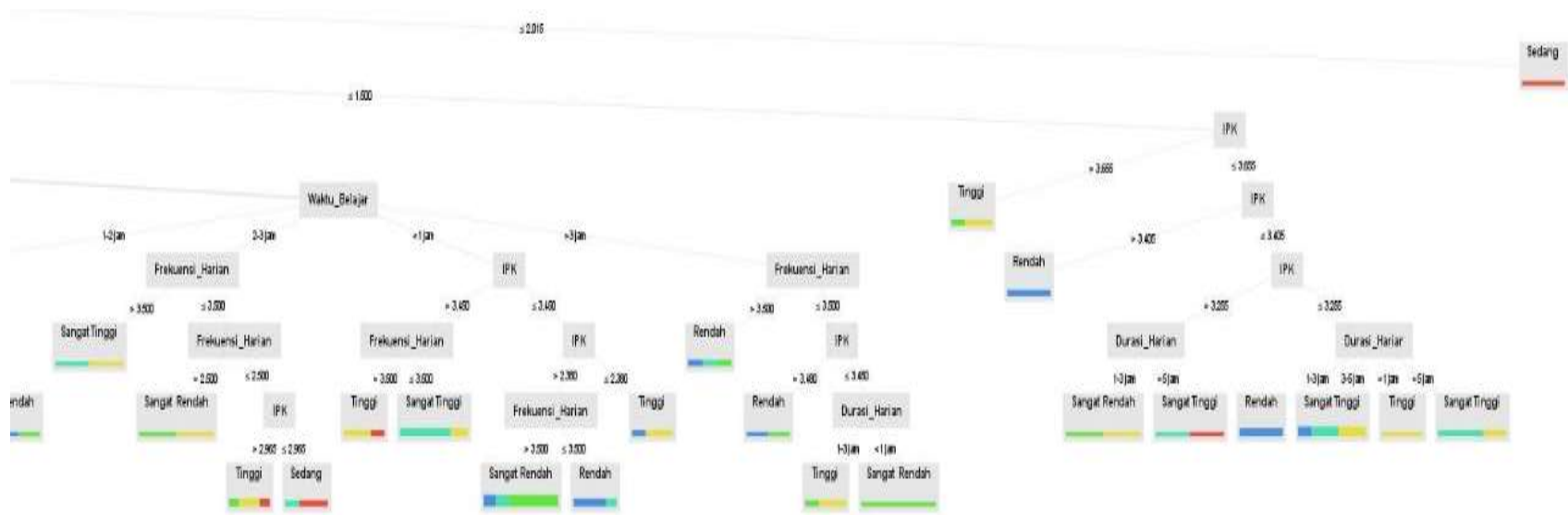
Pohon keputusan yang dihasilkan menunjukkan bahwa **frekuensi penggunaan media sosial** merupakan variabel yang paling signifikan dalam memprediksi tingkat konsentrasi belajar mahasiswa (Mostafa Alizade-Harakiyan, 2025). Mahasiswa dengan frekuensi akses media sosial tinggi (>5 kali/hari) cenderung memiliki konsentrasi belajar yang rendah. Sebaliknya, mahasiswa dengan frekuensi akses rendah (<3 kali/hari) menunjukkan konsentrasi belajar yang lebih baik. Melalui penerapan algoritma Decision Tree C4.5 pada RapidMiner, model yang diuji menghasilkan output berupa Decision Tree beserta aturan atau rule yang menyertainya. Representasi dari pohon keputusan dan rule tersebut dapat dilihat pada gambar di bawah ini..



Gambar 11
Pohon Keputusan



Gambar 12
Pohon Keputusan



Gambar 13
Pohon Keputusan

Berdasarkan gambar di atas dapat dilihat bahwa tingkat konsentrasi mahasiswa masih terbilang cukup rendah dikarenakan hasil penelitian menemukan bahwa konsentrasi mahasiswa pada tahap sedang ke rendah, hal ini menjadi bukti bahwa penggunaan media sosial yang terlalu lama dapat berpengaruh signifikan terhadap konsentrasi belajar.

```

IFH > 2.015
|
| Platform_Favorit = Facebook
|
| Frekuensi_Harian > 1.700
|
| | Frekuensi_Harian > 4.500
|
| | | IFH > 2.240
|
| | | | IFH > 2.570: Sedang (Bendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=1)
|
| | | | IFH < 2.570
|
| | | | IFH > 2.975: Sangat Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=0)
|
| | | | IFH < 2.975: Sangat Tinggi (Bendah=0, Sangat Tinggi=1, Sangat Rendah=1, Tinggi=1, Sedang=0)
|
| | | | IFH < 2.160: Rendah (Bendah=1, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=0)
|
| | Platform_Favorit = Instagram
|
| | Frekuensi_Harian > 9.500
|
| | | Platform_Favorit = Facebook
|
| | | | IFH > 2.100
|
| | | | Frekuensi_Harian > 9.500
|
| | | | Durasi_Harian = 1-5 jam: Sangat Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=1, Sedang=1)
|
| | | | Durasi_Harian = >5 jam: Rendah (Bendah=1, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=0)
|
| | | | Frekuensi_Harian < 8.500
|
| | | | | IFH > 2.800: Sangat Tinggi (Bendah=0, Sangat Tinggi=1, Sangat Rendah=1, Tinggi=1, Sedang=0)
|
| | | | | IFH < 2.800
|
| | | | | IFH > 2.600: Sedang (Bendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=1)
|
| | | | | IFH < 2.600: Rendah (Bendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=1)
|
| | | | IFH < 2.180: Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | Platform_Favorit = Twitter
|
| | | IFH > 3.250: Sangat Tinggi (Bendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | IFH < 3.200
|
| | | | IFH > 2.460
|
| | | | | IFH > 2.930: Sedang (Bendah=1, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=1)
|
| | | | | IFH < 2.930
|
| | | | | Frekuensi_Harian > 8.500: Sangat Tinggi (Bendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | | | Frekuensi_Harian < 8.000: Rendah (Bendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | | | IFH < 2.660: Sangat Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=1)
|
| | Platform_Favorit = TikTok
|
| | | Durasi_Harian = 1-2 jam: Sedang (Bendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=1)
|
| | | Durasi_Harian = 3-5 jam
|
| | | | IFH > 2.740: Sangat Rendah (Bendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | | IFH < 2.740: Sangat Tinggi (Bendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=1)
|
| | | Durasi_Harian < 1 jam
|
| | | | Frekuensi_Harian > 7.000: Sangat Tinggi (Bendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | | Frekuensi_Harian < 7.500: Sangat Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | | Durasi_Harian > 1 jam: Rendah (Bendah=1, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | Platform_Favorit = YouTube
|
| | | IFH > 3.000
|
| | | | Frekuensi_Harian > 8.500: Sedang (Bendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=1)
|
| | | | Frekuensi_Harian < 8.500: Rendah (Bendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | | | IFH < 2.100
|
| | | | Frekuensi_Harian > 6.500: Sangat Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=1, Sedang=0)
|
| | | Frekuensi_Harian < 6.000
|
| | | | IFH > 3.000: Sangat Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=1)
|
| | | | IFH < 3.000: Rendah (Bendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
|
| | Platform_Favorit = YouTube

```

```

Platform_Favorite = YouTube
IEK > 2.250
    Frekuensi_Harian > 0.500
        IFK > 3.655: Rendah (Rendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=1, Sedang=0)
        IFK < 3.655: Sangat Tinggi (Rendah=1, Sangat Tinggi=2, Sangat Rendah=0, Tinggi=0, Sedang=0)
    Frekuensi_Harian < 0.500
        IFK > 2.745: Sangat Tinggi (Rendah=1, Sangat Tinggi=4, Sangat Rendah=0, Tinggi=1, Sedang=1)
        IFK < 2.745: Sedang (Rendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=3)
    IEK < 2.250: Sangat Rendah (Rendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=1)
Frekuensi_Harian < 4.500
    Waktu_Belajar < 1.2 jam
        Frekuensi_Harian > 3.500
            IFK > 2.555: Tinggi (Rendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=1, Sedang=0)
            IFK < 2.555: Rendah (Rendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=1, Sedang=0)
        Frekuensi_Harian < 3.500
            Durasi_Harian = 1-3 jam: Rendah (Rendah=1, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=1)
            Durasi_Harian = 3-5 jam: Sangat Tinggi (Rendah=0, Sangat Tinggi=1, Sangat Rendah=1, Tinggi=0, Sedang=0)
            Durasi_Harian < 1 jam: Rendah (Rendah=1, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=0)
        Waktu_Belajar = 3-3 jam
            Frekuensi_Harian > 3.500: Sangat Tinggi (Rendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=1, Sedang=0)
            Frekuensi_Harian < 3.500
                IFK > 2.500: Sangat Rendah (Rendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=1, Sedang=1)
                IFK < 2.500: Sangat Tinggi (Rendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=2)
            Waktu_Belajar < 1 jam
                IFK > 3.450
                    Frekuensi_Harian > 3.500: Tinggi (Rendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=1, Sedang=1)
                    Frekuensi_Harian < 3.500: Sangat Tinggi (Rendah=0, Sangat Tinggi=3, Sangat Rendah=0, Tinggi=1, Sedang=0)
                IFK < 3.450
                    IFK > 3.300
                        Frekuensi_Harian > 3.500: Sangat Rendah (Rendah=1, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
                        Frekuensi_Harian < 3.500: Rendah (Rendah=3, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=0)
                    IFK < 2.350: Tinggi (Rendah=1, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=2, Sedang=0)
                Waktu_Belajar = >3 jam
                    Frekuensi_Harian > 3.500: Rendah (Rendah=1, Sangat Tinggi=1, Sangat Rendah=1, Tinggi=0, Sedang=0)
                    Frekuensi_Harian < 3.500
                        IFK > 3.450: Rendah (Rendah=1, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=0, Sedang=0)
                        IFK < 3.450
                            Durasi_Harian = 1-3 jam: Tinggi (Rendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=2, Sedang=0)
                            Durasi_Harian = >3 jam: Sangat Rendah (Rendah=0, Sangat Tinggi=0, Sangat Rendah=2, Tinggi=0, Sedang=0)
                    Frekuensi_Harian < 1.500
                        IFK > 3.555: Tinggi (Rendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=2, Sedang=0)
                        IFK < 3.655
                            IFK > 3.105: Rendah (Rendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
                            IFK < 3.105
                                IFK > 3.255
                                    Durasi_Harian = 1-3 jam: Sangat Rendah (Rendah=0, Sangat Tinggi=0, Sangat Rendah=1, Tinggi=1, Sedang=0)
                                    Durasi_Harian = >3 jam: Sangat Tinggi (Rendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=0, Sedang=1)
                                IFK < 3.255
                                    Durasi_Harian = 1-3 jam: Rendah (Rendah=4, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=0)
                                    Durasi_Harian = 3-5 jam: Sangat Tinggi (Rendah=1, Sangat Tinggi=2, Sangat Rendah=0, Tinggi=2, Sedang=0)
                                    Durasi_Harian < 1 jam: Tinggi (Rendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=2, Sedang=0)
                                    Durasi_Harian = >3 jam: Sangat Tinggi (Rendah=0, Sangat Tinggi=1, Sangat Rendah=0, Tinggi=1, Sedang=0)
                    IFK < 2.015: Sedang (Rendah=0, Sangat Tinggi=0, Sangat Rendah=0, Tinggi=0, Sedang=2)
    
```

Gambar 14. Aturan atau Rule Decision Tree pada Rapid Miner

3. Evaluasi Model

Model Decision Tree yang dibangun memiliki **akurasi sebesar 63%**, dengan class **precision** rendah 62,50%, sangat tinggi 54,698%, sangat rendah 63,41%, tinggi 72,7%, dan sedang 76,00%, sedangkan untuk clas recall true rendah 78,95%, sangat tinggi, 77,78%, sangat rendah 59,09%, tinggi 41,05% dan sedang 55,89% Evaluasi ini menunjukkan bahwa model cukup efektif dalam memprediksi tingkat konsentrasi belajar berdasarkan variabel yang ada.

accuracy: 63.00%

	true Rendah	true Sangat Tinggi	true Sangat Rendah	true Tinggi	true Sedang	class precision
pred. Rendah	30	7	7	2	2	62.50%
pred. Sangat Tinggi	3	35	6	13	7	54.69%
pred. Sangat Rendah	1	2	26	8	4	63.41%
pred. Tinggi	1	0	3	16	2	72.73%
pred. Sedang	3	1	2	0	19	76.00%
class recall	78.95%	77.78%	59.09%	41.03%	55.88%	

Gambar 18. Evaluasi Model Keakurasaan Data

Accuracy: 63.00%: Ini menunjukkan akurasi keseluruhan model, yaitu persentase prediksi yang benar dari total prediksi. Kolom tue Rendah, tue Sangat Tinggi, tue Sangat Rendah, tue Tinggi, bus Sedang: Ini adalah label kelas aktual (nilai sebenarnya) dari data. Baris pred Rendah, pred Sangat Tingg, pred Sangat Rendah, pred Tinggi, gred tetang: Ini adalah label kelas prediksi (hasil klasifikasi) oleh model. Angka di

dalam sel (misal: 30, 7, 20): Ini menunjukkan jumlah instance (data) yang termasuk dalam kombinasi kelas aktual dan kelas prediksi tertentu. Diagonal utama (misal: 30, 15, 20, 15, 19): Angka-angka ini mewakili jumlah prediksi yang benar (True Positives/True Negatives), di mana kelas prediksi sesuai dengan kelas aktual. Sel lainnya: Angka-angka ini mewakili jumlah prediksi yang salah (False Positives/False Negatives), di mana kelas prediksi tidak sesuai dengan kelas aktual. class precision adalah "Class Precision" atau presisi untuk setiap kelas. Presisi mengukur seberapa banyak dari instance yang diprediksi sebagai kelas tertentu benar-benar termasuk dalam kelas tersebut. Baris terakhir dengan persentase (misal: 77.78%, 59.09%): Ini kemungkinan adalah "Class Recall" atau recall untuk setiap kelas. Recall mengukur seberapa banyak dari instance aktual dalam suatu kelas berhasil diprediksi oleh mode.

5. Kesimpulan

Kesimpulan berikut dapat diambil berdasarkan temuan penelitian tentang penggunaan penambahan data dan algoritma Pohon Keputusan untuk menganalisis antara penggunaan media sosial dan kemampuan siswa untuk fokus selama kelas, bahwa **Analisis Penggunaan Media Sosial dengan Konsentrasi Belajar. Temuan analisis Pohon Keputusan menunjukkan bahwa variabel utama yang memengaruhi kemampuan siswa untuk fokus saat belajar adalah frekuensi dan durasi penggunaan media sosial mereka.** Mahasiswa dengan frekuensi penggunaan tinggi (>5 kali/hari) atau durasi penggunaan lama (>3 jam/hari) cenderung memiliki konsentrasi belajar rendah, sedangkan mahasiswa dengan frekuensi dan durasi rendah menunjukkan konsentrasi belajar lebih tinggi. Algoritma Decision Tree yang diterapkan menggunakan **RapidMiner Studio** berhasil membentuk model prediksi dengan akurasi sebesar 63.00%: Hal ini menunjukkan seberapa baik metode Pohon Keputusan bekerja untuk menemukan pola dalam hubungan antara karakteristik penggunaan media sosial dan tingkat konsentrasi belajar siswa. **Pola Prediksi Konsentrasi Belajar.** Pohon keputusan yang terbentuk memberikan insight bahwa kombinasi **frekuensi, durasi, platform media sosial, dan waktu belajar** dapat memprediksi tingkat konsentrasi belajar mahasiswa. Dengan kata lain, model ini dapat digunakan untuk memberikan rekomendasi penggunaan media sosial yang optimal agar konsentrasi belajar tetap terjaga.

Daftar Pustaka

1. A, E. K. C., & Begum, dan R. (2025). The academic implications of social media use and multitasking: Evidence from university students in a developing nation. *Social Sciences & Humanities Open*, 11, 101623.
2. Al, H. et. (n.d.). *Data Mining: Concepts and Techniques*.
3. Al, R. J. S. et. (2024). Night Screen Time is Associated with Cognitive Function in Healthy Young Adults: A Cross-Sectional Study. *Journal of Multidisciplinary Healthcare*, 17, 2093–2104.
4. Halim, S. (2024). Hubungan Intensitas Penggunaan Internet dan Kualitas Tidur dengan Konsentrasi Belajar pada Siswa Madrasah Aliyah Darul Ulum Amessangeng Maros. *Window of Public Health Journal*, 5(1), 160–169.
5. Han, J., Kamber, M., & Pei, J. (2023). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
6. Han, J. (2012). *Data Mining Concepts and Techniques* (Third Edit).
7. Indiani, R. (2025). *Penggunaan Media Sosial dan Produktivitas Belajar Mahasiswa*. Cendikia Pendidikan.
8. Jones, Chloe N., et al. (2024). Investigating the links between objective social media use, attentional control, and psychological distress. *Social Science & Medicine*, 361, 117400.
9. Kaewpradit, Kamollada, Pitchayanont Ngamchaliew, and N. B. (2025). Digital screen time usage, prevalence of excessive digital screen time, and its association with mental health, sleep quality, and academic performance among Southern University students. *Frontiers in Psychiatry*, 16, 1535631.
10. Marhaeni, Luh Putu, Putu Budi Adnyana, and N. L. P. M. W. (2020). Hubungan Penggunaan Smartphone dengan Konsentrasi dan Minat Belajar Biologi Siswa SMA. *Jurnal Pendidikan Biologi Undiksha*, 7(3), 137–147.
11. Mostafa Alizade-Harakiyan, et al. (2025). Decision tree-based machine learning algorithm for prediction of acute radiation esophagitis. *Biochemistry and Biophysics Reports*, 42, 101991.
12. Muhammad Firdaus Mustapha, dkk. (2023). The prediction of student's academic performance using RapidMiner. *Indonesian Journal of Electrical Engineering and Computer Science*, 32(1), 363–371.
13. Safitri, H. dan T. (2024). *Hubungan antara Konsentrasi Belajar dengan Tingkat Penggunaan Media Sosial TikTok pada Mahasiswa Fakultas Ushuluddin Adab dan Dakwah IAIN Parepare*. IAIN Parepare.
14. Sinaga, Daniel, Edwin J. Solaiman, and F. J. K. (2021). Penerapan algoritma decision tree C4. 5 untuk klasifikasi mahasiswa berpotensi drop out di universitas advent indonesia. *TeKa*, 11(2), 167–173.
15. Sugiyono. (2019). *Metode penelitian pendidikan: kuantitatif, kualitatif, kombinasi, R&D dan penelitian pendidikan*. Alfabeta.
16. Trendeline Haliti-Sylaja, A. S. (2024). Impact of Short Reels on Attention Span and Academic Performance of Undergraduate Students. *Eurasian Journal of Applied Linguistics*, 10(3), 60–68