

Penerapan Data Mining Dalam Pemilihan Laptop Mahasiswa Menggunakan Metode Random Forest (Studi Kasus : Universitas Methodist Indonesia)

Smansal Hartawan Totonafu Gulo¹, Humuntal Rumapea², Harlen Gilbert Simanullang³
^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, Sep 9, 2025
Revised, May 20, 2025
Accepted, Jun 11, 2025

Keywords:

Data Mining,
Random Forest,
Pemilihan Laptop,
Rapid Miner,
Evaluasi Model.

ABSTRAK

Pemilihan laptop yang sesuai menjadi kebutuhan mendasar bagi mahasiswa di Universitas Methodist Indonesia, mengingat pentingnya peran perangkat ini dalam mendukung aktivitas akademik. Penelitian ini bertujuan untuk mengklasifikasikan kecocokan spesifikasi laptop terhadap kebutuhan akademik mahasiswa dari lima fakultas menggunakan algoritma *Random Forest*. Dataset yang digunakan mencakup 326 data yang didapat melalui kuesioner dan wawancara, dengan atribut seperti RAM, prosesor, penyimpanan, daya tahan baterai, harga, dan ukuran layar. Algoritma *Random Forest* dipilih karena kemampuannya menghasilkan prediksi akurat melalui kombinasi pohon keputusan. Implementasi dilakukan menggunakan perangkat lunak RapidMiner, yang mempermudah proses pengolahan data, pelatihan model, dan evaluasi. Berdasarkan hasil evaluasi, model ini mencapai akurasi sebesar 89,23%, *precision* sebesar 91,49%, *recall* sebesar 93,48%, dan *F1-Score* sebesar 92,45%, menunjukkan performa yang andal dalam memberikan rekomendasi laptop. Penelitian ini memberikan kontribusi signifikan dalam membantu mahasiswa dan orang tua memilih laptop yang sesuai dengan kebutuhan akademik dan kondisi ekonomi

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Koresponden:

Smansal Hartawan Totonafu Gulo,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: smansalgulo06@gmail.com

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat menjadikan laptop sebagai kebutuhan esensial bagi mahasiswa dalam menunjang aktivitas akademik. Di Universitas Methodist Indonesia (UMI), kebutuhan laptop bervariasi sesuai dengan program studi, seperti pengembangan aplikasi dan analisis data di Fakultas Ilmu Komputer, penelitian berbasis sains di Kedokteran, analisis statistik di Ekonomi, studi literatur di Sastra, serta penggunaan *software* agribisnis di Pertanian. Pemilihan laptop yang tidak sesuai dapat berdampak negatif, baik dalam bentuk pemborosan anggaran maupun penurunan produktivitas akademik. Mahasiswa yang memilih spesifikasi terlalu rendah mungkin mengalami keterbatasan dalam menjalankan aplikasi yang dibutuhkan, sementara mereka yang memilih spesifikasi terlalu tinggi dapat mengalami pemborosan biaya.

Selain itu, perbedaan antara kebutuhan teknis dan keinginan pribadi sering kali menjadi tantangan dalam proses pemilihan laptop. Banyak mahasiswa cenderung memilih laptop berdasarkan tren, merek populer, atau desain yang menarik tanpa mempertimbangkan apakah spesifikasinya mendukung kebutuhan akademik mereka. Di sisi lain, kurangnya pemahaman orang tua tentang spesifikasi laptop yang sesuai serta adanya keterbatasan finansial semakin memperumit

pengambilan keputusan. Oleh karena itu, diperlukan pendekatan objektif yang mampu membantu mahasiswa memilih laptop berdasarkan aspek teknis yang relevan dengan jurusan mereka.

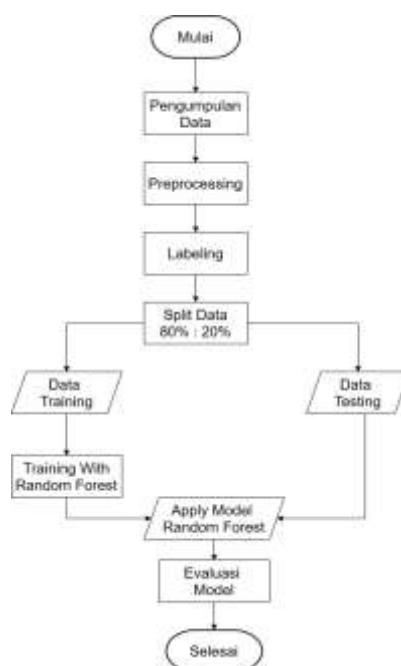
Data mining menjadi salah satu solusi dalam menangani permasalahan ini karena mampu mengekstraksi pola dari dataset besar untuk menghasilkan rekomendasi yang lebih akurat [1]. Dalam penelitian ini, metode Random Forest dipilih karena kemampuannya menghasilkan prediksi yang akurat dengan menggabungkan hasil dari beberapa pohon keputusan[2]. Metode ini sangat efektif dalam menangani data dengan banyak variabel serta dapat memberikan model yang andal dalam pengambilan keputusan [3].

Untuk mengimplementasikan metode ini, digunakan perangkat lunak RapidMiner yang mampu melakukan pengolahan data, pelatihan model, serta evaluasi dengan efisien tanpa perlu melakukan perhitungan manual yang kompleks [4]. Dengan demikian, penerapan metode Random Forest dalam pemilihan laptop mahasiswa di Universitas Methodist Indonesia diharapkan dapat memberikan rekomendasi yang lebih objektif, akurat, dan sesuai dengan kebutuhan akademik serta kemampuan finansial mahasiswa.

2. METODE PENELITIAN

2.1 Framework Penelitian

Framework penelitian dalam penelitian ini dapat dilihat pada Gambar 1 berikut.



Gambar 1. *Framework penelitian*

1. *Pengumpulan Data*
Penelitian ini diawali dengan mengumpulkan data kepada mahasiswa melalui kuesioner, wawancara dan studi pustaka.
2. *Preprocessing*
Data yang terkumpul kemudian dibersihkan dengan mengatasi nilai yang hilang, normalisasi, serta menghilangkan duplikasi atau outlier agar kualitas data tetap terjaga
3. *Labeling*
Setiap sampel dalam dataset diberikan label sesuai dengan kategori atau kelas yang ditentukan.
4. *Split Data*
Dataset dibagi menjadi 80% untuk data *training* dan 20% untuk data *testing* agar model dapat belajar dari sebagian besar data dan diuji pada data baru.

5. Data Training

Data *training* digunakan untuk membangun model menggunakan algoritma Random Forest, yang merupakan metode ensemble berbasis pohon keputusan.

6. Apply Model

Model yang telah dilatih kemudian diterapkan pada data testing untuk menguji performanya dalam melakukan prediksi.

7. Evaluasi Model

Kinerja model diukur menggunakan metrik seperti akurasi, *precision*, *recall*, *F1-score*, serta *confusion matrix* untuk melihat distribusi prediksi yang benar dan salah.

2.2 Data Mining

Data *mining* adalah proses yang digunakan untuk menemukan pola dan informasi berharga dari kumpulan data besar [5]. Proses ini melibatkan teknik statistik dan algoritma pembelajaran mesin untuk menganalisis data dan menghasilkan wawasan yang dapat digunakan dalam pengambilan keputusan. Dalam konteks bisnis, data *mining* dapat membantu perusahaan dalam memahami perilaku pelanggan serta meningkatkan strategi pemasaran yang diterapkan [6].

Dalam penelitian terbaru, data *mining* telah diterapkan dalam berbagai bidang, termasuk kesehatan, keuangan, dan pemasaran [7]. Dengan menggunakan teknik seperti klasifikasi dan klustering, peneliti dapat mengidentifikasi tren dan pola yang sebelumnya tidak terlihat, yang dapat meningkatkan efisiensi operasional dan pengambilan keputusan [8]. Aplikasi data *mining* di berbagai sektor, seperti keuangan, pemasaran, kesehatan, dan *e-commerce*, menunjukkan kemampuannya dalam mendeteksi penipuan, segmentasi pasar, analisis hasil perawatan, dan rekomendasi produk, menjadikannya alat yang sangat berharga dalam menghadapi tantangan bisnis modern [9].

2.3 Klasifikasi

Klasifikasi dalam data mining adalah proses mengelompokkan data ke dalam kategori atau kelas tertentu berdasarkan pola yang ditemukan dari dataset pelatihan [10]. Tujuannya adalah untuk memetakan pola-pola yang ada dalam data pelatihan ke dalam kelompok yang sesuai berdasarkan atribut-atribut yang beragam [11]. Dengan analisis ini, data yang belum diketahui kelasnya dapat diklasifikasikan, memungkinkan prediksi atau pengelompokan yang lebih akurat. Contoh metode klasifikasi yang umum digunakan mencakup algoritma seperti *Decision Trees*, *Naive Bayes*, dan *Support Vector Machines*, yang diaplikasikan dalam berbagai bidang, mulai dari deteksi email spam hingga diagnosis medis. Klasifikasi ini mendukung pengambilan keputusan yang lebih berbasis informasi dengan menganalisis pola yang diperoleh dari data historis [12].

Metode klasifikasi dalam data *mining* merupakan pendekatan untuk mengategorikan data ke dalam kelompok yang sudah ditentukan berdasarkan pola yang ditemukan dari data pelatihan. Langkah pertama dalam proses ini adalah pengumpulan dan pembersihan data, diikuti dengan pemilihan fitur yang relevan. Selanjutnya, algoritma klasifikasi seperti *Decision Trees*, *Naive Bayes*, *K-nearest Neighbors*, dan *Support Vector Machines* digunakan untuk membangun model prediktif. Teknik ini sangat berguna dalam berbagai bidang, mulai dari identifikasi spam hingga diagnosis medis, dan berperan penting dalam pengambilan keputusan yang didasarkan pada data [13].

2.4 Algoritma Random Forest

Metode *Random Forest* (RF) merupakan sebuah teknik yang dapat meningkatkan akurasi hasil, karena proses pembentukan simpul anak pada setiap *node* dilakukan secara acak. Metode ini

membangun pohon keputusan yang terdiri dari *root node*, *internal node*, dan *leaf node*, dengan pemilihan atribut dan data yang diambil secara acak berdasarkan aturan tertentu. *Root node* adalah simpul yang berada di posisi paling atas, yang berfungsi sebagai akar dari pohon keputusan. *Internal node* berfungsi sebagai simpul percabangan, dengan setidaknya dua output dan hanya satu input. Sementara itu, *leaf node* atau *terminal node* adalah *node* terakhir yang hanya memiliki satu input tanpa menghasilkan output. Proses pembuatan pohon keputusan dimulai dengan menghitung nilai *Entropy* untuk menilai tingkat ketidakmurnian atribut, serta menghitung nilai *information gain*. Rumus untuk menghitung *Entropy* dijelaskan dalam persamaan 1, sedangkan perhitungan *information gain* menggunakan persamaan 2 [14].

$$\text{Entropy}(S) = \sum_{i=1}^n - p_i * \log_2 p_i \dots\dots\dots (1)$$

Dimana :

S = Himpunan Kasus.

n = Jumlah Partisi S.

p_i = Proporsi dari S_i terhadap S

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i) \quad (2)$$

Dimana :

S = Himpunan Kasus.

A = Atribut

n = Jumlah Partisi Atribut A

$|S_i|$ = Jumlah kasus pada partisi ke-i

$|S|$ = Jumlah kasus dalam S

2.5 Confusion Matrix

Confusion matrix merupakan salah satu alat evaluasi yang digunakan untuk mengukur kinerja suatu *classifier* dalam mengklasifikasikan data. Alat ini membantu menentukan apakah *classifier* mampu mengenali data dari berbagai kelas dengan baik atau tidak. Jika *classifier* berhasil mengklasifikasikan data dengan benar, maka nilai *True Positive* (TP) dan *True Negative* (TN) akan menunjukkan keberhasilannya. Sebaliknya, jika terjadi kesalahan dalam klasifikasi, maka nilai *False Positive* (FP) dan *False Negative* (FN) akan memberikan informasi mengenai tingkat kesalahan yang terjadi [15].

Tabel 1. *Confusion Matrix*

		Kelas Prediksi	
		<i>Class = Yes</i>	<i>Class = No</i>
Kelas Sebenarnya	<i>Class = Yes</i>	TP	FN
	<i>Class = No</i>	FP	TN

Keterangan :

- TP (*true positive*) = Prediksi dikategorikan sebagai positif, dan kelas aslinya memang positif.
- TN (*true negative*) = Prediksi dikategorikan sebagai negatif, dan kelas aslinya juga negatif.
- FP (*false positive*) = Prediksi dikategorikan sebagai positif, namun kelas aslinya sebenarnya negatif.
- FN (*false negative*) = Prediksi dikategorikan sebagai negatif, padahal kelas aslinya adalah positif.

Rumus *confusion matrix* untuk menghitung *accuracy*, *precision*, dan *recall* sebagai berikut.

$$accuracy = \frac{TP + TN}{Total} \dots\dots\dots (3)$$

$$precision = \frac{TP}{TP + FP} \dots\dots\dots (4)$$

$$recall = \frac{TP}{TP + FN} \dots\dots\dots (5)$$

2.6 Rapid Miner

RapidMiner adalah perangkat lunak analisis data yang dikembangkan oleh perusahaan yang sama, menyediakan platform terintegrasi untuk berbagai keperluan seperti pembelajaran mesin, pembelajaran mendalam, penambangan teks, dan analisis prediktif. Perangkat ini mendukung seluruh tahapan dalam proses pembelajaran otomatis, mulai dari persiapan data, visualisasi hasil, hingga validasi dan pengoptimalan model [16]. RapidMiner termasuk dalam kategori perangkat lunak pengolahan data yang memungkinkan pembuatan model dari dataset besar dengan menggunakan prinsip-prinsip dan algoritma data *mining*. Pendekatan yang digunakan menggabungkan teknik statistika, kecerdasan buatan, dan pengelolaan basis data. Dengan menggunakan operator yang disediakan, RapidMiner memudahkan pengguna dalam melakukan komputasi pada data dalam skala besar [17]. Selain itu, perangkat ini memiliki antarmuka pengguna yang ramah, memungkinkan pemrograman dan analisis data tanpa memerlukan keterampilan pengkodean yang mendalam, sehingga mempercepat proses pembuatan model dan analisis data yang kompleks

3. HASIL DAN PEMBAHASAN

3.1 Data Yang Digunakan

Berdasarkan penelitian yang dilakukan di Universitas Methodist Indonesia, terdapat 326 data terkait pengalaman penggunaan laptop oleh mahasiswa dari lima fakultas. Sampel data mengenai penggunaan laptop oleh mahasiswa dapat dilihat pada Tabel 2 berikut.

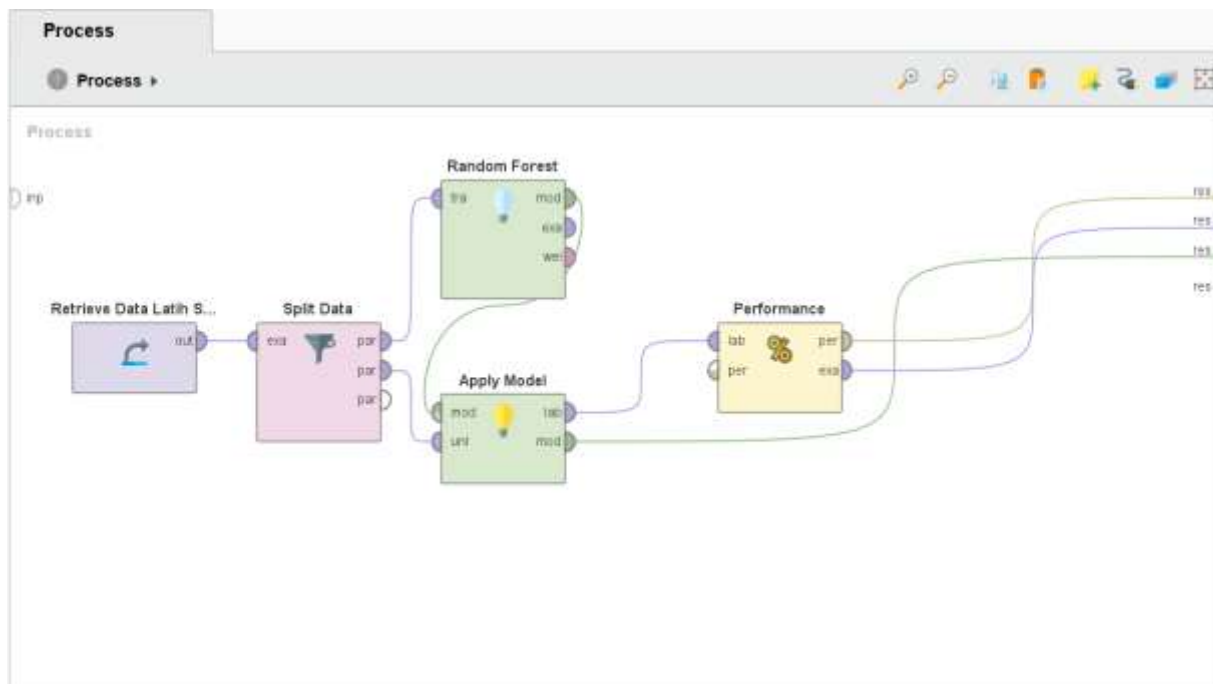
Tabel 2. Dataset laptop mahasiswa

No	Fakultas	Ram	Prosesor	Daya Tahan Baterai	Penyimpanan	Harga	Ukuran	Kebutuhan
1	Ilmu Komputer	8 GB	Amd Ryzen 3 Gen 10	< 4 jam	1 TB SSD	Rp 5.000.000 - Rp 10.000.000	> 16 inci	Memenuhi

2	Ilmu		Amd			Rp		
	Komputer	16	Ryzen 5	6 - 8	512 GB SSD	5.000.000 -	13 - 14	Memenuhi
		GB	Series	jam		Rp	inci	
			7000			10.000.000		
3	Ilmu		Amd			Rp		
	Komputer	16	Ryzen 5	> 8 jam	256 GB SSD	5.000.000 -	13 - 14	Memenuhi
		GB	Series			Rp	inci	
			5000			10.000.000		
4	Ilmu					Rp		
	Komputer	4	Intel Core	4 - 6	256 GB HDD	5.000.000 -	13 - 14	Tidak
		GB	i3 Gen 10	jam		Rp	inci	Memenuhi
						10.000.000		
5	Ilmu					Rp		
	Komputer	8	Intel Core	4 - 6	256 GB SSD	5.000.000 -	13 - 14	Memenuhi
		GB	i5 Gen 12	jam		Rp	inci	
						10.000.000		
...								
...								
...								
326	Ilmu					Rp		
	Komputer	4	Intel Core	< 4 jam	256 GB SSD	5.000.000 -	13 - 14	Memenuhi
		GB	i5 Gen 10			Rp	inci	
						10.000.000		

3.2 Hasil Pengujian Data

Konfigurasi hasil pengujian data menggunakan RapidMiner dapat dilihat pada Gambar 2 berikut.



Gambar 2. Konfigurasi Hasil Pengujian *RapidMiner*

1. *Retrieve* adalah operator yang berfungsi untuk memuat data dari spreadsheet Microsoft Excel.
2. *Split Data* merupakan operator yang digunakan untuk membagi data menjadi data latih dan data uji. Dalam penelitian ini, data dibagi dengan rasio 80% untuk data latih dan 20% untuk data uji.
3. *Random Forest* adalah metode yang digunakan dalam penelitian ini. Model Random Forest dibangun dengan 100 pohon keputusan (trees), di mana setiap pohon memiliki kedalaman maksimum 5.
4. *Apply Model* merupakan operator yang digunakan untuk menerapkan model yang telah dilatih pada data uji guna melakukan prediksi.
5. *Performance* adalah operator yang digunakan untuk mengukur tingkat akurasi dari model yang telah dibangun.

Sampel hasil akhir dari metode *Random Forest* yang dihasilkan dalam RapidMiner dapat dilihat pada Gambar 3 berikut.

...	Kebutuhan	prediction(K...	confidence(...	confidence(...	Fakultas	RAM	Prosesor	Days Tahan...	Penyimpanan	Ukuran Layar	Harga
1	Memenuhi	Memenuhi	1	0	Ilmu Komputer	Besar	Sedang	Hemat	Kecil	Kecil	Sedang
2	Memenuhi	Tidak Memen...	0.380	0.620	Ilmu Komputer	Sedang	Rendah	Tidak Hemat	Kecil	Sedang	Sedang
3	Memenuhi	Memenuhi	1	0	Ilmu Komputer	Sedang	Sedang	Hemat	Sedang	Kecil	Sedang
4	Memenuhi	Memenuhi	0.990	0.010	Ilmu Komputer	Besar	Sedang	Hemat	Sedang	Sedang	Sedang
5	Memenuhi	Memenuhi	0.960	0.040	Pertanian	Kecil	Rendah	Sedang	Kecil	Kecil	Sedang
6	Memenuhi	Memenuhi	1	0	Ekonomi	Sedang	Rendah	Sedang	Kecil	Kecil	Sedang
7	Tidak Memen...	Memenuhi	0.890	0.110	Ekonomi	Kecil	Rendah	Sedang	Sedang	Kecil	Murah
8	Memenuhi	Memenuhi	0.880	0.120	Ilmu Komputer	Sedang	Rendah	Sedang	Sedang	Kecil	Sedang
9	Memenuhi	Memenuhi	1	0	Ilmu Komputer	Sedang	Sedang	Hemat	Sedang	Kecil	Sedang
10	Tidak Memen...	Tidak Memen...	0.350	0.650	Ilmu Komputer	Kecil	Rendah	Sedang	Sedang	Kecil	Sedang
11	Tidak Memen...	Memenuhi	0.930	0.070	Kedokteran	Besar	Tinggi	Tidak Hemat	Kecil	Sedang	Mahal
12	Memenuhi	Memenuhi	1	0	Pertanian	Sedang	Sedang	Sedang	Sedang	Kecil	Sedang
13	Memenuhi	Memenuhi	0.740	0.260	Ekonomi	Besar	Rendah	Hemat	Sedang	Sedang	Sedang

Gambar 3. Hasil Metode *Random Forest*

Berdasarkan banyak data uji sebesar 65 data didapatkan 58 data yang sesuai dan 7 data yang tidak sesuai. Maka nilai akurasi yang didapatkan sebesar 89,23%.

3.3 Evaluasi Model

Evaluasi model bertujuan untuk mengukur kinerja model yang telah dibangun dengan menggunakan data uji. Beberapa metrik yang digunakan untuk evaluasi antara lain akurasi, *precision*, *recall*, dan *F1 score*. Metrik ini membantu menilai seberapa tepat model dalam memprediksi hasil yang benar, sehingga dapat memberikan gambaran tentang kekuatan dan kelemahan model yang digunakan. Hasil akhir klasifikasi dari 261 data latih dan 65 data uji menggunakan 100 pohon pada metode *Random Forest*, dapat dilihat pada tabel berikut ini.

Tabel 3. *Confusion matrix Random Forest*

	True Memenuhi	True Tidak Memenuhi
Pred. Memenuhi	43	4
Pred. Tidak Memenuhi	3	15

Berdasarkan hasil tabel di atas, terlihat bahwa 43 laptop diprediksi dengan benar masuk dalam kategori memenuhi, sementara 15 laptop diprediksi dengan benar masuk dalam kategori tidak memenuhi. Terdapat 4 laptop yang diprediksi masuk dalam kategori memenuhi, namun sebenarnya termasuk dalam kategori tidak memenuhi. Sedangkan 3 laptop diprediksi sebagai Tidak Memenuhi, tetapi sebenarnya memenuhi. Tingkat akurasi, presisi, *recall*, dan skor F1 juga dapat dievaluasi berdasarkan tabel di atas sebagai berikut.

1. Akurasi

$$akurasi = \frac{43 + 15}{65} = 0.8923$$

2. Presisi

$$presisi = \frac{43}{43 + 4} = 0.9149$$

3. *Recall*

$$Recall = \frac{43}{43 + 3} = 0.9348$$

4. *F1 Score*

$$F1\ Score = 2 \left(\frac{0.9347 * 0.9148}{0.9347 + 0.9148} \right) = 0.9245$$

4. KESIMPULAN

Berdasarkan hasil pengujian dan penelitian penerapan metode algoritma *Random Forest* dalam pemilihan kebutuhan laptop mahasiswa Universitas Methodist Indonesia yang telah dilakukan, maka dapat disimpulkan sebagai berikut:

1. Berdasarkan hasil pengujian, algoritma *Random Forest* terbukti mampu melakukan analisis dengan baik dan berhasil menyelesaikan permasalahan mahasiswa dalam pemilihan kebutuhan laptop sesuai dengan fakultas di Universitas Methodist Indonesia.
2. Berdasarkan hasil evaluasi model menggunakan metode Random menunjukkan kinerja sangat baik dalam memprediksi klasifikasi kebutuhan laptop mahasiswa, dengan akurasi 89,23%, presisi 91,49%, recall 93,48%, dan F1 score 92,45%. Hasil ini mencerminkan kemampuan model dalam meminimalkan kesalahan prediksi dan menghasilkan klasifikasi yang andal, sehingga layak digunakan sebagai alat efektif untuk mendukung rekomendasi pemilihan laptop mahasiswa.
3. Aplikasi *RapidMiner* digunakan sebagai aplikasi pendukung untuk implementasi *Random Forest* algoritma klasifikasi dari aplikasi RapidMiner dapat menghasilkan prediksi yang membantu mahasiswa dalam menentukan kebutuhan laptop.

REFERENSI

- [1] C. Zai, "Implementasi Data Mining Sebagai Pengolahan Data," *J. Portal Data*, vol. 2, no. 3, pp. 1–12, 2022.
- [2] H. G. Simanullang, A. P. Silalahi, and M. I. Hutapea, *Pendukung Keputusan Seleksi Indeks LQ-45*. Malang: Madza Media, 2023.
- [3] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 9, pp. 4305–4313, 2022.
- [4] S. Marzukhi, N. Awang, S. N. Alsagoff, and H. Mohamed, "RapidMiner and Machine Learning Techniques for Classifying Aircraft Data," *J. Phys. Conf. Ser.*, vol. 1997, no. 1, 2021, doi: 10.1088/1742-6596/1997/1/012012.
- [5] Y. Syach Putra, R. Kurniawan, and Y. Arie Wijaya, "Penerapan Data Mining Menggunakan Algoritma Fp-Growth Pada Data Penjualan Sembako," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 561–567, 2024, doi: 10.36040/jati.v8i1.8391.
- [6] Sumiyatun, Y. Cahyadi, and E. Iskandar, "Data Mining Untuk Memprediksi Status Kelulusan Mahasiswa," *J. Inform. Komputer, Bisnis dan Manaj.*, vol. 21, no. 3, pp. 11–19, 2023, doi: 10.61805/fahma.v21i3.3.
- [7] I. M. D. Rahayu, P. W., Sudipa, I. G. I., Suryani, S., Surachman, A., Ridwan, A., Darmawiguna, I. G. M., ... & Maysanjaya, *BUKU AJAR DATA MINING*. PT. Sonpedia Publishing Indonesia, 2024., 2024.
- [8] V. Loja, G. K. Pati, A. P. Setiawi, U. Stella, and M. Sumba, "Klasifikasi Data Mining dalam Memprediksi Kinerja Karyawan dengan Metode Algoritma C4 . 5 pada Toko Merpati Simpang," vol. 2, no. 4, 2024.
- [9] H. S. Suci and Sugiyono, "Pemodelan Data Mining Transaksi Penjualan Menggunakan Algoritma Apriori (Studi Kasus: Kedai Ngodeng & Smoothies)," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no, 2024.
- [10] C. Armayani, A. Fauzi, and H. Sembiring, "Implementasi Data Mining Pengelompokan Jumlah Data Produktivitas Ubinan Tanaman Pangan Berdasarkan Jenis Ubinan Dengan Metode Clustering Dikab Langkat (Studi Kasus : Badan Pusat Statistik Langkat)," *J. Inform. Kaputama*, vol. 5, no. 1, pp. 185–196, 2021, doi: 10.59697/jik.v5i1.318.
- [11] A. P. Silalahi and H. G. Simanullang, "Supervised Learning Metode K-Nearest Neighbor Untuk Prediksi Diabetes Pada Wanita," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 7, no. 1, pp. 144–149, 2023, doi: 10.46880/jmika.vol7no1.pp144-149.
- [12] R. Gustrianda and D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids," *J. Media Inform.*

- Budidarma*, vol. 6, no. 1, p. 27, 2022, doi: 10.30865/mib.v6i1.3294.
- [13] M. Melladia, D. E. Putra, and L. Muhelni, "Penerapan Data Mining Pemasaran Produk Menggunakan Metode Clustering," *J. Tek. Inf. dan Komput.*, vol. 5, no. 1, p. 160, 2022, doi: 10.37600/tekinkom.v5i1.458.
 - [14] Y. Chlarasasti and I. K. D. Nuryana, "Komparasi Algoritma Naïve Bayes dan Random Forest Dalam Mengetahui Kepuasan Pengguna Aplikasi Jenius," *J. Emerg. Inf. ...*, vol. 04, no. 03, pp. 64–72, 2023.
 - [15] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <https://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369/348>
 - [16] F. Meila Azzahra Sofyan, A. Voutama, and Y. Umaidah, "Penerapan Algoritma C4.5 Untuk Prediksi Penyakit Paru-Paru Menggunakan Rapidminer," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 2, pp. 1409–1415, 2023, doi: 10.36040/jati.v7i2.6810.
 - [17] K. F. Irnanda, A. P. Windarto, and I. S. Damanik, "Optimasi Particle Swarm Optimization Pada Peningkatan Prediksi dengan Metode Backpropagation Menggunakan Software RapidMiner," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 1, p. 122, 2022, doi: 10.30865/jurikom.v9i1.3836.