

Implementasi Metode Decission Tree Dalam Mengklasifikasi Depresi Menggunakan Rapidminer

Syariful Abrori^{1,*}, Zaehol Fatah¹

* Korespondensi: e-mail: syarifulabrori2307@gmail.com

¹ Teknologi Informasi; Universitas Ibrahimy; Jl. KHR. Syamsul Arifin No. 1-2, Sukorejo, Sumberejo, Banyuputih, Situbondo 68374; e-mail: syarifulabrori2307@gmail.com, zaeholfatah@gmail.com

Submitted : **10 September 2024**
Revised : **11 Oktober 2024**
Accepted : **9 November 2024**
Published : **30 November 2024**

Abstract

Depression has become a serious mental health problem with a significant impact on quality of life and work productivity. This study aims to develop a depression classification model using the Decision Tree method implemented through RapidMiner software. The dataset consists of 2054 data with 11 variables covering demographic aspects, working conditions, and mental health. Data preprocessing is carried out through several stages, including data format conversion, categorical variable transformation using Nominal to Binominal, and numeric data normalization with Z-transformation. Implementing the Decision Tree uses the gain ratio parameter as the criterion, maximum depth 10, and confidence 0.1, and activates the pruning and prepruning features for model optimization. The results showed excellent performance with an accuracy of 97.50%, a weighted mean recall of 92.29%, and a weighted mean precision of 93.49%. The confusion matrix shows that the model successfully identified 1463 non-depression cases and 139 depression cases correctly, with a low misclassification rate.

Keywords: *Decision Tree, Depression Classification, Machine Learning, Mental Health, RapidMiner*

Abstrak

Depresi telah menjadi masalah kesehatan mental yang serius dengan dampak signifikan terhadap kualitas hidup dan produktivitas kerja. Penelitian bertujuan untuk mengembangkan model klasifikasi depresi menggunakan metode *Decision Tree* yang diimplementasikan melalui *software RapidMiner*. *Dataset* yang digunakan terdiri dari 2054 data dengan 11 variabel yang mencakup aspek demografis, kondisi kerja, dan kesehatan mental. *Preprocessing* data dilakukan melalui beberapa tahap, termasuk konversi format data, transformasi variabel kategorikal menggunakan Nominal to Binominal, dan normalisasi data numerik dengan *Z-transformation*. Implementasi *Decision Tree* menggunakan parameter *gain ratio* sebagai *criterion*, *maximal depth* 10, dan *confidence* 0.1, serta mengaktifkan fitur *pruning* dan *prepruning* untuk optimalisasi model. Hasil penelitian menunjukkan performa yang sangat baik dengan akurasi mencapai 97.50%, *weighted mean recall* 92.29%, dan *weighted mean precision* 93.49%. *Confusion matrix* menunjukkan model berhasil mengidentifikasi 1463 kasus non-depresi dan 139 kasus depresi dengan tepat, dengan tingkat kesalahan klasifikasi yang rendah.

Kata kunci: *Decision Tree, Klasifikasi Depresi, Kesehatan Mental, Machine Learning, RapidMiner*

1. Pendahuluan

Depresi telah menjadi salah satu masalah kesehatan mental yang sangat serius di era modern. Menurut data *World Health Organization (WHO)*, lebih dari 264 juta orang di seluruh dunia mengalami depresi, dengan peningkatan kasus yang signifikan setiap tahunnya. Di Indonesia, angka penderita depresi terus meningkat mencapai 6,1% dari total populasi pada tahun 2023. Kondisi yang sangat mengkhawatirkan, terutama di kalangan profesional menghadapi berbagai tekanan dalam pekerjaan (Dahlia, 2022). Dampak depresi tidak hanya terbatas pada kesehatan mental individu, tetapi juga berpengaruh besar terhadap produktivitas kerja, hubungan sosial, dan kualitas hidup secara keseluruhan. Penelitian menunjukkan bahwa pekerja yang mengalami depresi cenderung mengalami penurunan kinerja hingga 35% dan memiliki tingkat absensi yang lebih tinggi dibandingkan rekan kerja (Putri et al., 2022).

Mendiagnosis depresi bukanlah hal yang sederhana karena melibatkan berbagai faktor yang saling terkait. Beberapa faktor utama yang mempengaruhi kondisi depresi seseorang meliputi tekanan dalam pekerjaan, tingkat kepuasan terhadap pekerjaan, kualitas tidur, pola makan, dan riwayat kesehatan mental dalam keluarga. Selain itu, faktor-faktor seperti usia, jenis kelamin, dan jam kerja juga berperan penting dalam menentukan risiko depresi seseorang (Saras, 2023). Metode diagnosis tradisional yang mengandalkan konsultasi langsung dengan profesional kesehatan mental seringkali memakan waktu lama dan membutuhkan biaya yang tidak sedikit. Hal ini dapat menjadi hambatan bagi banyak orang untuk mendapatkan bantuan yang mereka butuhkan secara tepat waktu. Stigma sosial yang masih melekat pada masalah kesehatan mental membuat banyak orang enggan untuk mencari bantuan profesional secara langsung (Ariyani, 2022).

Perkembangan teknologi informasi, khususnya dalam bidang *machine learning*, membuka peluang baru dalam upaya mendeteksi dan mengklasifikasikan depresi secara lebih efektif. Metode *Decision Tree* muncul sebagai salah satu pendekatan yang menjanjikan dalam menganalisis dan mengklasifikasikan tingkat depresi berdasarkan berbagai variabel yang ada. Metode ini dipilih karena kemampuannya dalam menghasilkan model keputusan yang tidak hanya akurat tetapi juga mudah dipahami oleh pengguna awam (Rachmawati et al., 2023). *Decision Tree* bekerja dengan membuat serangkaian keputusan berdasarkan karakteristik data yang ada, mirip dengan cara seorang profesional kesehatan mental mendiagnosis pasien, namun dengan kecepatan dan konsistensi yang lebih tinggi.

Penelitian menggunakan *dataset* yang komprehensif dengan 2054 responden dan 11 variabel yang mencakup berbagai aspek kehidupan profesional dan pribadi. Implementasi metode *Decision Tree* dilakukan menggunakan platform *RapidMiner*, yang dikenal dengan kemampuannya dalam memproses data secara efisien dan menyediakan antarmuka yang mudah digunakan. Pemilihan *RapidMiner* sebagai *tool* analisis didasarkan pada kemampuannya dalam menangani *dataset* besar dan menyediakan berbagai fitur untuk optimasi model. Melalui kombinasi *dataset* yang representatif, metode *Decision Tree* yang handal, dan platform *RapidMiner* yang powerful, penelitian ini bertujuan untuk mengembangkan

sistem klasifikasi depresi yang dapat membantu dalam proses screening awal dan deteksi dini kasus depresi di kalangan profesional (Dinasty et al., 2023; Wasik et al., 2024).

2. Metode Penelitian

2.1. Data Mining

Merupakan proses pengolahan data untuk menemukan pola dan informasi penting dari kumpulan data yang besar. Proses ini melibatkan berbagai teknik analisis data yang bertujuan untuk mengubah data mentah menjadi pengetahuan yang bermanfaat. Dalam praktiknya, data mining memiliki beberapa tahapan penting mulai dari pemahaman data, persiapan data (*preprocessing*), pemodelan, hingga evaluasi hasil. *Data mining* banyak diterapkan di berbagai bidang, termasuk kesehatan, untuk mengidentifikasi pola-pola tersembunyi yang dapat membantu dalam pengambilan keputusan. Khususnya dalam konteks kesehatan mental, data mining membantu mengungkap hubungan antara berbagai faktor yang mempengaruhi kondisi mental seseorang, seperti hubungan antara pola hidup, lingkungan kerja, dan risiko depresi (Rachmawati et al., 2023).

2.2. Decision Tree

Metode klasifikasi dan prediksi yang menggunakan struktur pohon untuk merepresentasikan aturan keputusan. Metode ini bekerja dengan cara memecah data menjadi himpunan-himpunan yang lebih kecil berdasarkan pemilihan atribut yang optimal pada setiap node. Pemilihan atribut terbaik dilakukan menggunakan perhitungan *Information Gain* dan *Gain Ratio* (Pratama et al., 2024). Proses dimulai dengan menghitung *entropy* dari *dataset*, dimana S adalah himpunan kasus dan p_i adalah proporsi kelas i dalam *dataset*.

$$\text{Entropy}(S) = -\sum(p_i \log_2 p_i) \quad (1)$$

Selanjutnya, *Information Gain* dihitung untuk menentukan efektivitas suatu atribut dalam mengklasifikasikan data. Dimana A adalah atribut dan S_v adalah subset yang dihasilkan jika *dataset* dipisah berdasarkan atribut A .

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum((|S_v|/|S|) \times \text{Entropy}(S_v)) \quad (2)$$

Untuk mengatasi bias terhadap atribut dengan banyak nilai, digunakan *Gain Ratio*, dimana $\text{SplitInfo}(S,A) = -\sum((|S_v|/|S|) \times \log_2(|S_v|/|S|))$

$$\text{GainRatio}(S,A) = \text{Gain}(S,A) / \text{SplitInfo}(S,A) \quad (3)$$

2.3. Depresi

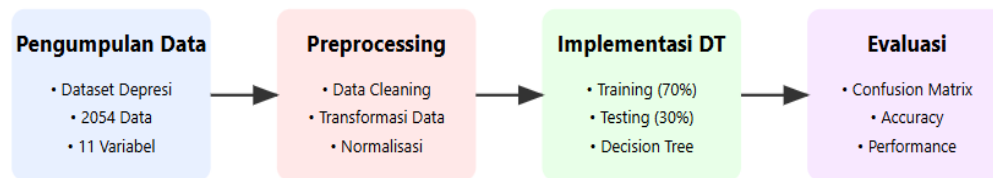
Merupakan gangguan kesehatan mental yang ditandai dengan perasaan sedih berkepanjangan, kehilangan minat terhadap aktivitas yang biasa dilakukan, dan penurunan kemampuan dalam menjalani aktivitas sehari-hari. Menurut *World Health Organization (WHO)*, depresi mempengaruhi lebih dari 264 juta orang di seluruh dunia dan menjadi penyebab utama disabilitas global. Pada kalangan profesional, depresi sering dipicu oleh berbagai faktor seperti tekanan kerja yang tinggi, ketidakpuasan kerja, gangguan pola tidur, dan masalah finansial.

Gejala depresi dapat bervariasi dari ringan hingga berat, termasuk perubahan pola tidur, nafsu makan terganggu, sulit berkonsentrasi, dan dalam kasus berat dapat muncul pikiran

tentang bunuh diri (Saras, 2023). Pemahaman yang mendalam tentang depresi sangat penting dalam konteks penelitian ini, karena membantu dalam menentukan variabel-variabel yang relevan untuk dianalisis menggunakan metode *Decision Tree*, serta memastikan hasil klasifikasi yang diperoleh memiliki makna yang berarti dalam konteks kesehatan mental.

2.4. Tahapan Penelitian

Tahap pertama dimulai dengan pengumpulan data, dimana *dataset* yang digunakan berasal dari platform Kaggle yang terdiri dari 2054 data dengan 11 variabel. Variabel-variabel ini mencakup informasi seperti jenis kelamin, usia, tekanan kerja, kepuasan kerja, durasi tidur, kebiasaan makan, riwayat pikiran bunuh diri, jam kerja, tekanan finansial, riwayat kesehatan mental keluarga, dan status depresi sebagai variabel target. *Dataset* yang awalnya berbentuk CSV dikonversi ke format Excel untuk memudahkan proses pengolahan di *RapidMiner* (Yogianto et al., 2024).



Sumber: Hasil Penelitian (2024)

Gambar 1. Metodologi Penelitian

Tahap kedua adalah *preprocessing* data yang merupakan langkah penting untuk memastikan kualitas data yang akan diproses. Pada tahap ini, dilakukan beberapa proses seperti penentuan variabel target menggunakan operator *Set Role*, dimana kolom "*Depression*" ditetapkan sebagai label. Selanjutnya, dilakukan transformasi data menggunakan operator *Nominal to Binominal* untuk mengubah data kategorikal menjadi format biner, dan *Nominal to Numerical* untuk mengkonversi variabel kategorikal lainnya menjadi bentuk numerik. Proses normalisasi data juga diterapkan menggunakan operator *Normalize* dengan metode *Z-transformation* untuk menyeragamkan skala nilai pada variabel numerik (Jalil et al., 2024).

Implementasi *Decision Tree* menjadi tahap ketiga dalam penelitian ini. Pada tahap ini, data yang telah melalui *preprocessing* dibagi menjadi dua bagian dengan proporsi 70% untuk data training dan 30% untuk data testing menggunakan operator *Split Data*. Algoritma *Decision Tree* dikonfigurasi dengan parameter yang telah dioptimalkan, dimana *criterion* menggunakan *gain_ratio*, *maximal depth* diatur pada nilai 10, dan *confidence* sebesar 0.1. Fitur *pruning* dan *prepruning* juga diaktifkan untuk mencegah *overfitting* dan mengoptimalkan struktur pohon keputusan (Wanto et al., 2020).

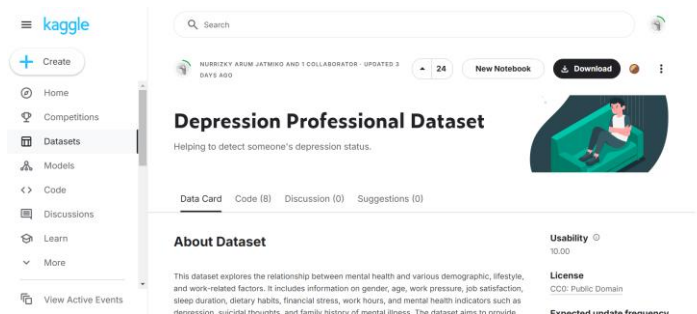
Tahap terakhir adalah evaluasi model yang dilakukan untuk mengukur performa dari model yang telah dikembangkan. Evaluasi dilakukan menggunakan beberapa metrik seperti accuracy, confusion matrix, *weighted mean recall*, dan *weighted mean precision*. Hasil evaluasi ini memberikan gambaran komprehensif tentang kemampuan model dalam mengklasifikasikan kasus depresi dan non-depresi. Setiap metrik evaluasi dianalisis untuk memastikan model tidak

hanya memiliki akurasi yang tinggi tetapi juga seimbang dalam mengidentifikasi kedua kelas target (Alrasyid et al., 2024).

3. Hasil dan Pembahasan

3.1. Implementasi Decision Tree

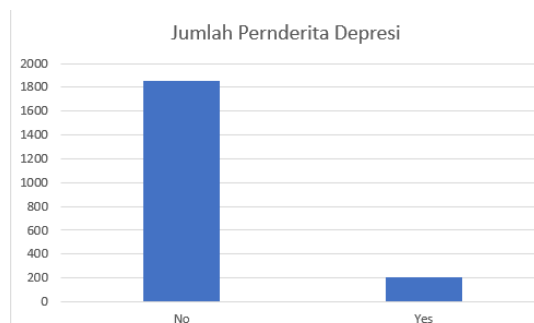
Dataset yang digunakan dalam penelitian ini merupakan *dataset* tentang depresi pada kalangan profesional yang diperoleh dari platform *Kaggle* pada link ini <https://www.kaggle.com/datasets/ikynahidwin/depression-professional-dataset>. *Dataset* ini terdiri dari 2054 data dengan 11 variabel yang mencakup berbagai aspek kehidupan profesional dan kondisi kesehatan mental individu. Berdasarkan tipe datanya, variabel-variabel dalam *dataset* ini dapat dikelompokkan menjadi tiga kategori utama.



Sumber: Hasil Penelitian (2024)

Gambar 2. Halaman *Dataset* Kaggle

Kategori pertama adalah variabel kategorikal yang terdiri dari Gender (jenis kelamin dengan nilai *Male/Female*), *Sleep Duration* (durasi tidur yang terbagi menjadi *Less than 5 hours/5-6 hours/7-8 hours/More than 8 hours*), dan *Dietary Habits* (kebiasaan makan dengan kategori *Healthy/Moderate/Unhealthy*). Kategori kedua mencakup variabel numerik seperti *Age* (usia dalam tahun), *Work Pressure* (tekanan kerja dengan skala 1-5), *Job Satisfaction* (kepuasan kerja dengan skala 1-5), *Work Hours* (jam kerja per hari), dan *Financial Stress* (tekanan finansial dengan skala 1-5). *Suicidal Thoughts* (riwayat pikiran bunuh diri dengan nilai Yes/No) dan *Family History* (riwayat kesehatan mental keluarga dengan nilai Yes/No) juga termasuk dalam variabel yang dianalisis.



Sumber: Hasil Penelitian (2024)

Gambar 3. Diagram Jumlah Penderita Depresi

Variabel target atau label dalam *dataset* ini adalah Depression yang memiliki dua nilai yaitu "Yes" untuk mengindikasikan adanya depresi dan "No" untuk kondisi tidak depresi. *Dataset* ini mencerminkan kompleksitas faktor-faktor yang dapat mempengaruhi kondisi depresi seseorang, mulai dari aspek pribadi seperti pola tidur dan kebiasaan makan, hingga aspek profesional seperti tekanan kerja dan kepuasan kerja. Keberagaman variabel ini memungkinkan analisis yang komprehensif tentang faktor-faktor yang berkontribusi terhadap depresi di kalangan profesional.

3.2. Preprocessing Data

Tahap *preprocessing* data dimulai dengan konversi format *dataset* dari CSV (*Comma Separated Values*) ke format Excel (*.xlsx*). Konversi dilakukan untuk memudahkan pembacaan dan pengolahan data di *RapidMiner*, mengingat format Excel lebih kompatibel dan menyediakan fitur yang lebih lengkap untuk manipulasi data. *Dataset* yang digunakan terdiri dari 2054 baris data dengan 11 variabel yang terkait dengan depresi.



Sumber: Hasil Penelitian (2024)

Gambar 4. Tahapan Preprocessing

Setelah format data dikonversi, dilakukan serangkaian proses *preprocessing* menggunakan beberapa operator di *RapidMiner*. Pertama, menggunakan operator *Set Role* untuk mendefinisikan kolom "Depression" sebagai label atau target variable, akan menjadi variabel yang diprediksi oleh model. Label memiliki dua nilai kategorikal yaitu "Yes" untuk mengindikasikan adanya depresi dan "No" untuk kondisi tidak depresi.

	A	B	C	D	E	F	G	H	I	J	K
	Gender	Age	Work Pressure	Job Satisfaction	Sleep Duration	Dietary Habits	Have you ever had suicidal thoughts ?	Work Hours	Financial Stress	Family History of Mental Illness	Depression
1	Female	37	2	4	4-7.8 hours	Moderate	No	6	2	No	No
2	Male	60	4	3	5.6 hours	Unhealthy	Yes	0	4	Yes	No
3	Female	42	2	2	3.5-6 hours	Moderate	No	0	2	No	No
4	Female	44	3	3	5-7.8 hours	Healthy	Yes	1	2	Yes	No
5	Male	48	4	4	3-7.8 hours	Moderate	Yes	5	5	Yes	No
6	Female	60	1	4	4-7.8 hours	Unhealthy	Yes	12	3	Yes	No
7	Female	30	4	2	More than 8 hours	Healthy	No	3	1	No	No
8	Male	30	1	2	More than 8 hours	Unhealthy	Yes	6	1	No	No
9	Male	68	1	2	More than 8 hours	Moderate	Yes	11	5	Yes	No
10	Female	35	3	4	Less than 5 hours	Moderate	No	6	4	Yes	No
11	Male	21	2	5	Less than 5 hours	Moderate	No	6	1	Yes	No
12	Male	57	1	2	Less than 5 hours	Healthy	No	10	1	No	No
13	Male	54	3	3	2-7.8 hours	Healthy	No	11	4	No	No
14	Male	51	2	3	Less than 5 hours	Healthy	No	8	5	No	No
15	Male	18	3	3	1 More than 8 hours	Unhealthy	Yes	7	3	No	Yes
16	Female	51	3	3	Less than 5 hours	Unhealthy	No	9	1	No	No
17	Male	31	3	4	7.8 hours	Healthy	No	9	6	No	No
18	Female	60	4	1	Less than 5 hours	Unhealthy	No	10	5	Yes	No
19	Male	58	1	2	More than 8 hours	Healthy	No	8	4	No	No
20	Female	51	5	5	5-6 hours	Healthy	No	11	2	No	No
21	Male	47	4	4	1.5-6 hours	Moderate	No	10	4	Yes	No
22	Male	50	4	1	Less than 5 hours	Moderate	No	8	1	Yes	No

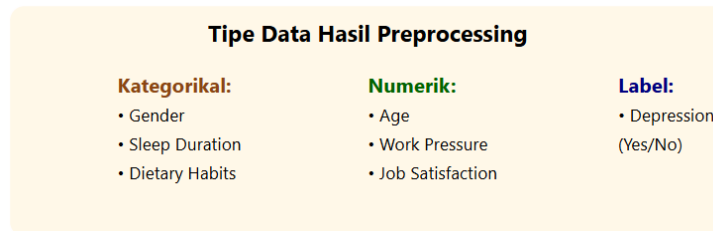
Sumber: Hasil Penelitian (2024)

Gambar 5. Dataset setelah dilakukan convert Excel

Langkah selanjutnya adalah transformasi data menggunakan operator *Nominal to Binominal* untuk mengkonversi variabel kategorikal menjadi format *biner*. Proses ini penting terutama untuk variabel-variabel seperti Gender (Male/Female), *Sleep Duration* (Less than 5 hours/5-6 hours/7-8 hours/More than 8 hours), *Dietary Habits* (Healthy/Moderate/Unhealthy), dan variabel kategorikal lainnya. Transformasi ini diperlukan karena algoritma *Decision Tree* membutuhkan data dalam format numerik untuk pemrosesan.

Untuk variabel numerik seperti *Age*, *Work Pressure*, *Job Satisfaction*, *Work Hours*, dan *Financial Stress*, dilakukan normalisasi data menggunakan operator *Normalize* dengan metode

Z-transformation. Normalisasi ini bertujuan untuk menyeragamkan skala nilai pada seluruh variabel numerik, sehingga tidak ada variabel yang mendominasi proses pembelajaran model karena perbedaan skala yang terlalu besar. *Z-transformation* mengubah data sehingga memiliki mean 0 dan standar deviasi 1, yang membantu dalam meningkatkan performa model.

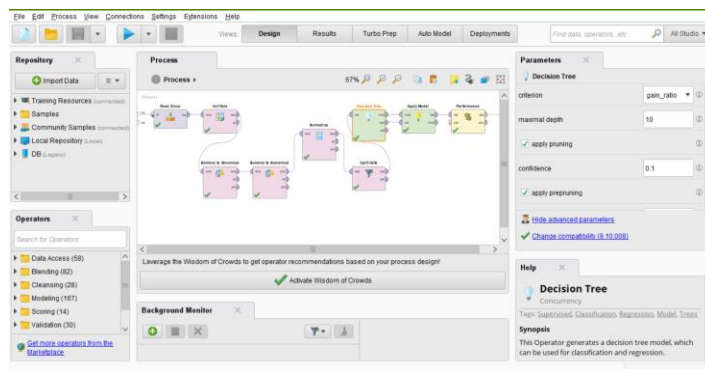


Sumber: Hasil Penelitian (2024)

Gambar 6. Pemilihan Atributes

3.3. Implementasi *Decision Tree*

Dalam penelitian implementasi algoritma *Decision Tree* untuk klasifikasi depresi menggunakan beberapa parameter kunci yang telah dioptimalkan. Parameter *criterion* diatur menggunakan *gain_ratio*, yang merupakan pengembangan dari information gain yang mempertimbangkan jumlah dan ukuran cabang saat memilih atribut untuk pemisahan. Pemilihan *gain_ratio* sebagai *criterion* bertujuan untuk mengatasi bias yang mungkin terjadi terhadap atribut dengan banyak nilai yang berbeda.



Sumber: Hasil Penelitian (2024)

Gambar 7. Tahapan Proses *RapidMiner*

Untuk mengontrol kompleksitas pohon keputusan, parameter *maximal depth* diatur pada nilai 10. Pengaturan ini membatasi jumlah tingkat maksimum dalam pohon keputusan, yang membantu mencegah model menjadi terlalu spesifik terhadap data training. Pembatasan kedalaman pohon ini penting untuk menjaga keseimbangan antara kemampuan model dalam menangkap pola data dan kemampuan generalisasi terhadap data baru.

3.4. Evaluasi Model

Berdasarkan hasil pengujian yang telah dilakukan, model menunjukkan performa yang sangat memuaskan dengan tingkat akurasi mencapai 97,50%. Hasil mengindikasikan bahwa dari keseluruhan prediksi yang dilakukan, model berhasil mengklasifikasikan status depresi dengan benar pada 97,50% kasus.

Analisis *confusion matrix* memberikan pemahaman yang lebih detail tentang kinerja model. Dari hasil pengujian, model berhasil mengidentifikasi 1463 *true negative* (kasus non-depresi yang diprediksi dengan benar) dan 139 *true positive* (kasus depresi yang diprediksi dengan benar). Terdapat 23 kasus *false positive*, di mana model salah mengklasifikasikan kasus non-depresi sebagai depresi, dan 18 kasus *false negative*, di mana model gagal mengidentifikasi kasus depresi yang sebenarnya. Jumlah kesalahan klasifikasi yang relatif kecil menunjukkan kehandalan model dalam membedakan antara kasus depresi dan non-depresi.



Sumber: Hasil Penelitian (2024)

Gambar 8. Hasil Evaluasi Model

Performa model juga dinilai melalui *weighted mean recall* yang mencapai 92,29%, mengindikasikan kemampuan model yang baik dalam mengidentifikasi kasus positif dari seluruh kasus yang sebenarnya positif. *Weighted mean precision* sebesar 93,49% menunjukkan tingkat ketepatan model dalam mengklasifikasikan kasus positif, dengan persentase *false positive* yang rendah. Kedua matrix menggambarkan keseimbangan yang baik antara sensitivitas dan spesifisitas model dalam mengklasifikasikan depresi.

4. Kesimpulan

Penelitian implementasi metode *Decision Tree* dalam mengklasifikasikan depresi menggunakan *RapidMiner* telah berhasil mencapai hasil yang sangat memuaskan. Model yang dikembangkan mampu mencapai tingkat akurasi sebesar 97,50% dengan performa yang seimbang dalam mengidentifikasi kasus depresi maupun non-depresi. Hasil didukung oleh nilai *weighted mean recall* sebesar 92,29% dan *weighted mean precision* 93,49% yang menunjukkan kemampuan model yang baik dalam mengklasifikasi kedua kelas. Keberhasilan tidak terlepas dari proses *preprocessing* data yang komprehensif, mulai dari konversi format data, transformasi variabel kategorikal, hingga normalisasi data numerik. Pengaturan parameter *Decision Tree* yang optimal, termasuk penggunaan *gain ratio* sebagai *criterion*, *maximal depth* 10, dan nilai *confidence* 0,1 berkontribusi signifikan terhadap performa model. Implementasi fitur pruning dan prepruning juga terbukti efektif dalam mencegah overfitting dan menghasilkan model yang lebih generalisasi. Penelitian membuktikan bahwa metode *Decision Tree* dapat menjadi alat yang efektif dalam mendukung proses *screening* awal kasus depresi di kalangan professional.

Untuk pengembangan penelitian saran yang dapat dipertimbangkan. Pertama, perlu dilakukan eksplorasi dengan ukuran *dataset* yang lebih besar dan lebih beragam untuk meningkatkan kemampuan generalisasi model. Penggunaan data dari berbagai kelompok demografis dan latar belakang profesi yang berbeda dapat memperkaya model dan membuatnya lebih adaptif. Kedua, dapat dilakukan perbandingan dengan metode *machine learning* lainnya seperti *Random Forest*, *Support Vector Machine*, atau *Neural Network* untuk mendapatkan perspektif yang lebih komprehensif tentang metode terbaik dalam klasifikasi depresi. Ketiga, penambahan variabel-variabel baru yang relevan dengan konteks depresi di era modern, seperti penggunaan media sosial, tingkat stress digital, atau dampak *work from home*, dapat meningkatkan relevansi model dengan kondisi saat ini.

Daftar Pustaka

- Alrasyid, H., Homaidi, A., & Fatah, Z. (2024). *Comparison Support Vector Machine and Random Forest Algorithms in Detect Diabetes*. 1(1), 447–453.
- Ariyani, E. (2022). *Cara Mencegah Terjadinya Depresi*. Victory Pustaka Media.
- Dahlia. (2022). Kesehatan Mental. In *Early Childhood Education Journal* (Issue November 2019).
- Dinasty, R. P., Widasari, E. R., & Fitriyah, H. (2023). Implementasi Metode Decision Tree untuk Sistem Pendeteksi Stres berdasarkan Detak Jantung dan Kelenjar Keringat. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 7(1), 154–161. <http://j-ptiik.ub.ac.id>
- Jalil, A., Homaidi, A., & Fatah, Z. (2024). Implementasi Algoritma Support Vector Machine Untuk Klasifikasi Status Stunting Pada Balita. *G-Tech: Jurnal Teknologi Terapan*, 8(3), 2070–2079. <https://doi.org/10.33379/gtech.v8i3.4811>
- Pratama, K., Harris, A., & Melinia W, L. (2024). *Belajar Mudah dan Singkat Machine Learning: Panduan Praktis dengan Studi Kasus, Kode Program dan Dataset* (Vol. 1). ANDI.
- Putri, U. N. H., Nur'aini, Sari, A., & Mawaadah, S. (2022). *Modul Kesehatan Mental*. CV Azka Pustaka.
- Rachmawati, S. S. P., Prakusa, K. V., & Rihastuti, S. (2023). Penerapan Data Mining dengan Metode Decision Tree untuk Prediksi Cuaca di Kota Seattle menggunakan Aplikasi Weka. *Seminar Nasional Amikom Surakarta (Semnasa)*, November, 93–100.
- Saras, T. (2023). *Mengatasi Depresi: Panduan Lengkap untuk Memahami, Mengelola, dan Menemukan Kembali Kesejahteraan Emosional*. Tiram Media.
- Wanto, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., Napitupulu, D., Negara, E. S., Lubis, M. R., Dewi, S. V., & Prianto, C. (2020). *Data Mining: Algoritma dan Implementasi*. Yayasan Kita Menulis. <https://books.google.co.id/books?id=LSPfDwAAQBAJ&printsec=frontcover&dq=Data+Mining+Untuk+Klasifikasi+dan+Klasterisasi+Data+%7C+Suyanto+%7C+Informatika&hl=en&sa=X&ved=2ahUKEwi99pHppvXtAhXW7XMBHW->

SAsIQ6AEwAHoECAUQAg#v=onepage&q=supervised&f=false

- Wasik, A., Fatah, Z., Munazilin, A., Studi, P., Informasi, S., Situbondo, U. I., Studi, P., Komputer, I., & Situbondo, U. I. (2024). Implementasi data mining untuk memprediksi penjualan accessoris handphone dan handphone terlaris menggunakan metode k-nearest neighbor (k-nn) 1. *Seminar Nasional Sains Dan Teknologi "SainTek" Seri 02, 1(2)*, 469–479.
- Yogianto, A., Homaidi, A., & Fatah, Z. (2024). Implementasi Metode K-Nearest Neighbors (KNN) untuk Klasifikasi Penyakit Jantung. *G-Tech: Jurnal Teknologi Terapan, 8(3)*, 1720–1728. <https://doi.org/10.33379/gtech.v8i3.4495>