
KLASIFIKASI MENTAL DISORDER DENGAN MENGGUNAKAN ALGORITMA RANDOM FOREST

Duwi Cahya Putri Buani¹, Ami Rahmawati²

¹Program Studi Informatika, Universitas Nusa Mandiri, Jakarta

²Program Studi Sistem Informasi, Universitas Nusa Mandiri, Jakarta

Email: *duwi.dcp@nusamandiri.ac.id, ami.amv@nusamandiri.ac.id

Abstrak

Gangguan mental merupakan masalah kesehatan global yang terus meningkat, mempengaruhi jutaan orang dan berdampak signifikan pada kualitas hidup, produktivitas, serta biaya perawatan kesehatan. WHO melaporkan lebih dari 970 juta orang di seluruh dunia mengalami gangguan mental, dengan peningkatan kasus depresi dan kecemasan selama pandemi COVID-19. Deteksi dini dan klasifikasi yang akurat sangat penting untuk intervensi yang tepat, namun tantangan muncul dari sifat data yang kompleks dan heterogen. Algoritma Random Forest menawarkan solusi potensial melalui pendekatan machine learning dalam mengklasifikasikan gangguan mental. Algoritma ini unggul dalam menangani dataset besar dan kompleks, serta mengatasi masalah data yang tidak seimbang. Penelitian ini bertujuan mengembangkan model klasifikasi gangguan mental menggunakan Random Forest, yang terbukti memiliki kinerja terbaik dengan akurasi 88,89%, precision 90%, recall 89%, dan F1-score 89%, mengungguli lima model lainnya. Model ini diharapkan dapat mempercepat diagnosis dan meningkatkan pengambilan keputusan klinis untuk penanganan gangguan mental.

Keyword : *Gangguan Mental, Machine Learning, Random Forest*

Abstract

Mental disorders are a growing global health issue, affecting millions of people and significantly impacting quality of life, productivity, and healthcare costs. The WHO reports that over 970 million people worldwide suffer from mental disorders, with rising cases of depression and anxiety during the COVID-19 pandemic. Early detection and accurate classification are crucial for appropriate interventions, yet challenges arise from the complexity and heterogeneity of data. The Random Forest algorithm offers a potential solution through a machine learning approach to classify mental disorders. It excels in handling large, complex datasets and addresses issues of data imbalance. This study aims to develop a mental disorder classification model using Random Forest, which demonstrated superior performance with an accuracy of 88.89%, precision of 90%, recall of 89%, and F1-score of 89%, outperforming five other models. The model is expected to accelerate diagnosis and enhance clinical decision-making in mental disorder management.

Keywords— *Mental disorder, Machine Learning, Random Forest*

I. PENDAHULUAN

Gangguan mental (mental disorder) merupakan salah satu masalah kesehatan utama yang mempengaruhi jutaan individu di seluruh dunia. Prevalensi gangguan mental terus meningkat dan membawa dampak besar terhadap kualitas hidup, produktivitas, serta beban biaya perawatan kesehatan masyarakat. WHO melaporkan bahwa lebih dari 970 juta orang di seluruh dunia mengalami gangguan mental, seperti depresi, gangguan kecemasan, skizofrenia,

dan gangguan bipolar. Pada tahun 2020 jumlah orang yang mengalami gangguan kecemasan dan depresi meningkat signifikan karena adanya pandemi COVID-19 dapat diperkirakan masing-masing meningkat sebesar 26% dan 28% [1].

Deteksi dini dan klasifikasi yang tepat terhadap gangguan mental sangat penting untuk memberikan intervensi yang sesuai dan mengurangi dampak negatif jangka panjang yang ditimbulkan [2]. Namun, salah satu tantangan utama dalam klasifikasi gangguan mental adalah sifat data yang kompleks dan

heterogen, mencakup gejala psikologis, data demografis, serta faktor sosial ekonomi. Pendekatan tradisional dalam diagnosis seringkali bergantung pada penilaian subjektif, yang dapat menyebabkan hasil yang bias atau tidak konsisten [3]. Oleh karena itu, penerapan teknologi kecerdasan buatan, khususnya algoritma machine learning, menjadi solusi potensial untuk mengatasi tantangan ini.

Algoritma Random Forest, yang dikenal dengan kemampuannya dalam menangani dataset yang besar dan kompleks, menawarkan pendekatan yang kuat untuk klasifikasi gangguan mental. Random Forest adalah salah satu teknik ensemble learning yang menggabungkan banyak pohon keputusan (decision trees) untuk meningkatkan akurasi dan mengurangi overfitting [4]. Algoritma ini sangat efektif dalam menangani data yang tidak seimbang dan variabel yang saling berhubungan, yang sering ditemui dalam dataset gangguan mental [5].

Penelitian terbaru telah menunjukkan potensi besar algoritma Random Forest dalam menganalisis data kesehatan mental. Sebagai contoh, penelitian dengan judul *Detection of child depression using machine learning methods* menggunakan algoritma ini untuk mengklasifikasikan gangguan depresi berdasarkan gejala klinis dan menemukan bahwa Random Forest berhasil mencapai akurasi yang tinggi dibandingkan dengan algoritma lain [6]. Penelitian lain dengan judul *Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms* juga menemukan bahwa algoritma Random Forest dapat digunakan untuk memprediksi tingkat keparahan gangguan kecemasan dengan memanfaatkan data dari survei kesehatan mental [7].

Dengan demikian, penelitian ini bertujuan untuk mengembangkan model klasifikasi gangguan mental menggunakan algoritma Random Forest. Model ini diharapkan dapat membantu dalam pengambilan keputusan klinis, mempercepat proses diagnosis, dan memberikan perawatan yang lebih tepat sasaran untuk pasien dengan gangguan mental.

II. TINJAUAN PUSTAKA

Machine Learning atau Pembelajaran mesin dapat diartikan sebagai penerapan komputer dan algoritma matematika yang belajar dari data untuk menghasilkan prediksi di masa mendatang. Proses pembelajaran ini mencakup dua tahap utama, yaitu pelatihan (training) dan pengujian (testing), yang bertujuan untuk mencapai kecerdasan. Dalam penelitian ini penulis menggunakan algoritma klasifikasi untuk menentukan model terbaik, algoritma tersebut diantaranya adalah:

a. Logistic Regression (LR)

Logistic Regression adalah metode statistik yang digunakan untuk melakukan klasifikasi pada masalah dimana variabel dependen bersifat biner atau dikotomis (misalnya, hasilnya adalah "ya" atau "tidak"). Metode ini memodelkan hubungan antara satu atau lebih variabel independen dengan variabel dependen biner menggunakan fungsi logistik. Logistic regression sering digunakan dalam ilmu sosial, kedokteran, serta bidang ilmu komputer untuk prediksi berdasarkan data kategoris [8].

b. Random Forest Classifier (RF)

Random Forest adalah pengembangan dari algoritma Decision Tree. Dalam metode ini, setiap Decision Tree dilatih menggunakan sampel individu, dan atribut dipecah berdasarkan subset acak. Random Forest memiliki beberapa keuntungan, seperti meningkatkan akurasi meskipun ada data yang hilang, tahan terhadap outliers, serta efisien dalam penyimpanan data. Selain itu, Random Forest juga melakukan seleksi fitur untuk memilih atribut terbaik, sehingga dapat meningkatkan kinerja model klasifikasi. Dengan kemampuan seleksi fitur ini, Random Forest mampu bekerja secara efektif dengan big data yang memiliki parameter kompleks. [9].

c. SVM

Algoritma Support Vector Machine (SVM) bertujuan untuk menemukan hyperplane maksimal, yaitu fungsi yang memisahkan dua kelas. Dalam prosesnya, SVM berupaya memaksimalkan margin atau jarak antara pola pelatihan dan batas keputusan. Algoritma ini memiliki beberapa keunggulan, antara lain performa yang baik baik untuk jumlah data kecil maupun besar, kinerja yang optimal pada data dengan banyak atribut, serta kemudahan dalam implementasinya. Awalnya, SVM hanya dapat digunakan untuk klasifikasi biner,

tetapi kini telah berkembang sehingga dapat mengklasifikasikan beberapa kelas sekaligus. Selain untuk klasifikasi, SVM juga dapat diterapkan dalam regresi dan deteksi outlier[10].

d. Naïve Bayes (NB)

Naive Bayes adalah metode klasifikasi probabilitas yang sederhana, yang menghitung kemungkinan dengan menggabungkan berbagai kombinasi dan frekuensi nilai dari basis data yang ada. Algoritma ini menggunakan teorema Bayes untuk memperkirakan atribut yang bersifat independen dan tidak saling bergantung, yang dapat diberikan oleh suatu nilai pada kelas variabel. Naive Bayes merupakan pendekatan klasifikasi berdasarkan probabilitas dan perhitungan yang dikembangkan oleh ilmuwan asal Inggris, Thomas Bayes, yang menghasilkan prediksi peluang berdasarkan pengalaman sebelumnya [11].

e. Decision Tree Classifier(DTC)

Decision tree adalah salah satu algoritma yang efektif untuk melakukan prediksi, baik dengan menggunakan data kategorikal maupun data numerikal. Algoritma ini membangun model dalam bentuk pohon, di mana setiap cabang mewakili keputusan berdasarkan nilai atribut, dan setiap daun menunjukkan hasil akhir atau prediksi. Kelebihan dari decision tree adalah kemampuannya untuk memvisualisasikan proses pengambilan keputusan dengan cara yang intuitif, sehingga memudahkan pemahaman bagi pengguna. Selain itu, decision tree dapat dengan mudah menangani data yang tidak terstruktur dan juga dapat digunakan dalam berbagai aplikasi, mulai dari analisis risiko hingga pemodelan dalam pemasaran. Karena fleksibilitasnya, algoritma ini sering digunakan sebagai langkah awal dalam banyak proyek machine learning [12].

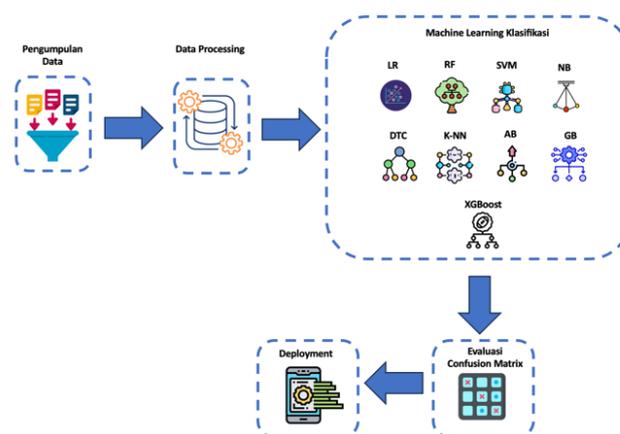
f. K-Neighbors Classifier (K-NN)

k-Nearest Neighbors (kNN) adalah algoritma klasifikasi machine learning yang sederhana dan populer, pertama kali diperkenalkan oleh T. Cover dan P. Hart pada tahun 1967. kNN mengklasifikasikan sampel berdasarkan kelas tetangga terdekatnya dan sering disebut sebagai

"lazy learner" karena tidak membangun model, melainkan hanya mengingat data pelatihan. Algoritma ini banyak digunakan dalam berbagai bidang, termasuk sains, teknik, ekonomi, dan bisnis, untuk aplikasi seperti prediksi penyakit jantung, diagnosis diabetes, dan klasifikasi protein. Selain itu, kNN juga menjadi dasar untuk pengembangan algoritma machine learning yang lebih kompleks [13].

III. METODOLOGI PENELITIAN

Gambar 1 merupakan kerangka alur penelitian dalam melakukan klasifikasi gangguan mental.



Gambar 1. Kerangka Penelitian

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data publik yang diambil dari <https://www.kaggle.com/>, data tersebut merupakan kumpulan gejala mental disorder.

2. Data Processing

Teknik-teknik umum yang digunakan dalam pengolahan data sebelum analisis, seperti pembersihan data, normalisasi, dan pengisian nilai yang hilang. Pengolahan data penting untuk mengurangi noise dalam dataset dan memastikan kualitas data yang lebih baik [14].

Dalam tahapan data processing kami melakukan analisis data, menormalisasi data dan memberikan nilai pada data yang kosong.

3. Modeling Menggunakan Algoritma Machine Learning Klasifikasi

Tahapan selanjutnya yang dilakukan dalam penelitian ini adalah kami melakukan modeling dengan Sembilan algoritma klasifikasi untuk menemukan model terbaik dalam melakukan

klasifikasi gangguan mental. Algoritma yang digunakan adalah Random Forest, Logistic Regression, Naive Bayes, Decision Tree Classifier, Support Vector Machine, dan K-Nearest Neighbors

4. Evaluasi Menggunakan Confusion

Metriks

Evaluasi yang diterapkan dalam penelitian ini menggunakan *confusion matrix* dengan mencari nilai Akurasi, Presisi, Recall dan F1-Score.

$$\text{Accuracy} = \frac{TP}{\text{Jumlah Data}}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F - 1 \text{ Score} = \frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$$

5. Deployment

Tahapan penelitian yang terakhir adalah Deployment. Setelah mendapatkan model terbaik maka model terbaik tersebut disimpan dan dijadikan aplikasi untuk melakukan klasifikasi.

IV. HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Dalam penelitian ini kami menggunakan data publik dari <https://www.kaggle.com/>. Data merupakan Kumpulan data pasien sebanyak 120 pasien yang memiliki gejala mental disorder dengan 17 gejala penting yang digunakan oleh psikiater untuk melakukan diagnosa mental disorder. 17 gejala tersebut dapat dilihat pada table 1.

Table 1. Penjelasan Data Mental Disorder

Variable	Keterangan
Sadness	Tingkat Kesedihan (Usually dan Sometime)
Euphoric	Tingkat Euphoric (Seldom, Most-Often, Usually dan Sometime)
Exhaustness	Tingkat Kelelahan (Usually dan Sometime)
Sleep	Gangguan Tidur (Seldom,

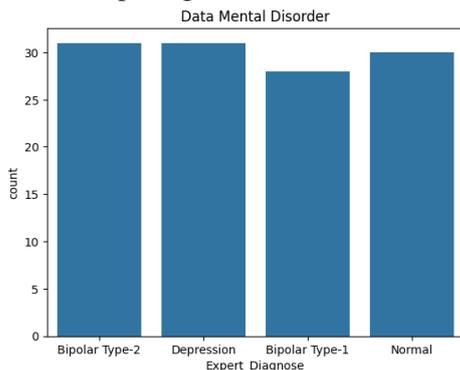
Variable	Keterangan
Dissorder	Most-Often, Usually dan Sometime)
Mood Swing	Perubahan Suasana Hati (Ya dan Tidak)
Suicidal Thoughts	Pikiran untuk bunuh diri (Ya dan Tidak)
Anorxia	Anoreksi adalah Gangguan makan sehingga menurunkan berat badan drastic (Ya dan Tidak)
Authority Respect	Menghormati Figur Autoritas (Ya dan Tidak)
Try-Explanation	Menncoba Menjelaskan (Ya dan Tidak)
Nervous breakdown	Gangguan Syaraf (Ya dan Tidak)
Ignore & Move-on	Menagbaikan dan Melupakan (Ya dan Tidak)
Admitting mistakes	Mengakui Kesalahan (Ya dan Tidak)
Overthinking	Berfikir Berlebihan (Ya dan Tidak)
Aggressive response	Respon Agresif (Ya dan Tidak)
Optimism	Optimisme (1 From 10, 2 Form 10, 3 From 10, 4 From 10, 5 From 10, 6 from 10, 7 From 10, 8 From 10 dan 9 From 10)
Sexual Activity	Aktivitas Seksual (1 From 10, 2 Form 10, 3 From 10, 4 From 10, 5 From 10, 6 from 10, 7 From 10, 8 From 10 dan 9 From 10)
Concentration	Konsentrasi (1 From 10, 2 Form 10, 3 From 10, 4 From 10, 5 From 10, 6 from 10, 7 From 10, 8 From 10 dan 9 From 10)
Expert Diagnose	Diagnosa Ahli (Normal, Bipolar Type-1, Bipolar Type-2, Depresi)

Expert Diagnose atau Diagnosa Ahli merupakan variable atau atribut kelas atau label yang digunakan untuk melakukan klasifikasi gangguan mental.

2. Data Processing/Persiapan Data

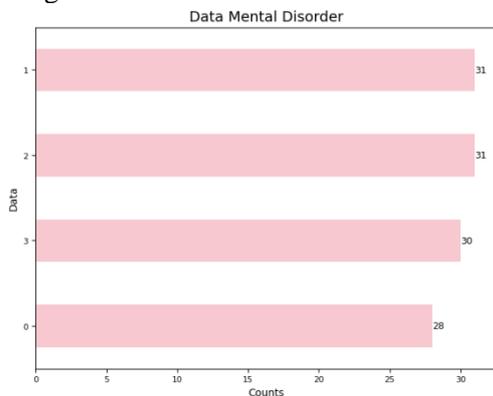
Sebelum melakukan modeling yang harus dilakukan adalah mempersiapkan data, tujuan dari persiapan data adalah agar data dapat diolah oleh

algoritma dan mendapatkan hasil akurasi yang optimal. Langkah-langkah yang dilakukan dalam proses ini adalah mengecek data yang kosong, mengecek duplikat data, dan melakukan label encoding. Data yang digunakan setelah dilakukan pengecekan tidak ada data yang kodong dan tidak ada data yang duplikat. Untuk proses leber endocer dapat dilihat pada gambar 2.



Gambar 2. Variabel Kelas Sebelum dilakukan Label Encoder

Untuk hasil dari label encoder dapat dilihat pada gambar 3:



Gambar 3. Variabel Kelas setelah dilakukan Label Encoder

Gambar 2 dan 3 merupakan gambar kelas yang dimiliki oleh data yang digunakan dalam penelitian dimana ada 4 kelas yaitu Normal, Bipolar Type-1, Bipolar Type-2, Depresi.

3. Modeling

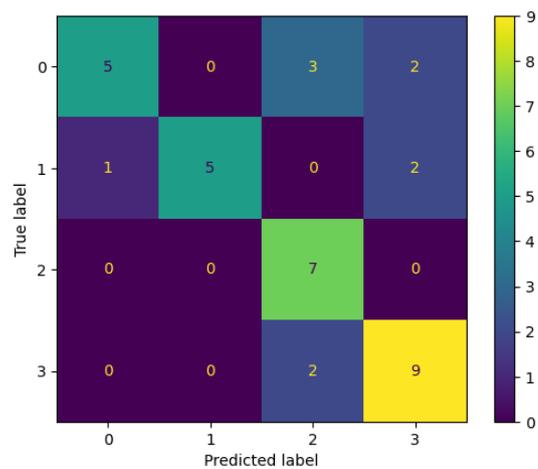
Setelah proses persiapan data selesai langkah selanjutnya yang dilakukan adalah modeling atau menerapkan algoritma klasifikasi dalam machine learning. Sebelum melakukan perhitungan dengan menggunakan algoritma machine learning hal yang dilakukan pertama kali adalah membagi data menjadi data training dan data testing.

Dalam penelitian ini data dibagi menjadi data training 70% dan data testing sebesar 30%.

Algoritma klasifikasi yang digunakan adalah Logistic Regression (LR), Random Forest Classifier (RF), SVM, Naïve Bayes (NB), Decision Tree Classifier(DTC) dan K-Neighbors Classifier (K-NN).

a. Logistic Regression (LR)

Hasil dari penggunaan LR untuk data Mental Disorder adalah sebesar 72,22% dengan grafik evaluasi dapat dilihat pada gambar 4.

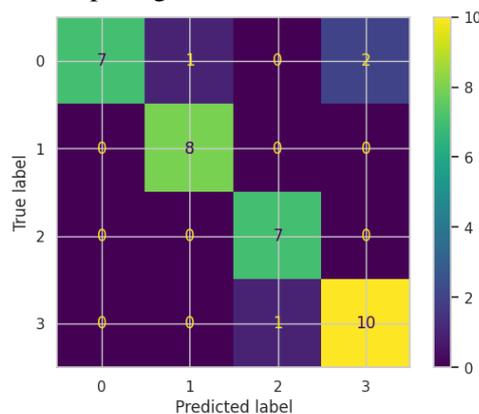


Gambar 4. Grafik hasil Evaluasi Algoritma LR

Pada gambar 4 menunjukkan bahwa masih ada data yang tidak sesuai hasil prediksinya.

b. Random Forest Classifier (RF)

Hasil akurasi dari algoritma RF adalah sebesar 88,89%. Hasil akurasi lebih besar dibandingkan model yang menggunakan algoritma LR. Dengan grafik evaluasi dapat dilihat pada gambar 5.

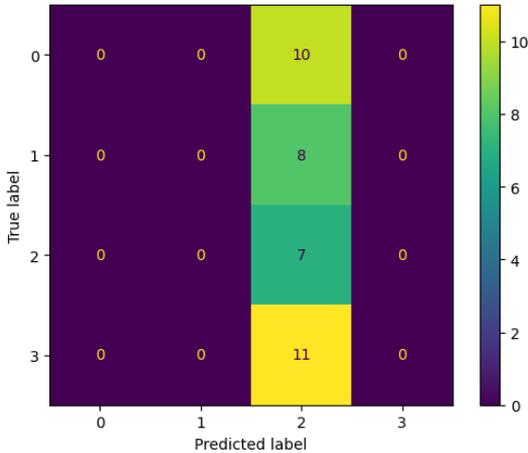


Gambar 5. Grafik hasil evaluasi algoritma RF

Dari grafik 5 dapat dilihat bahwa masih ada data yang diprediksi tidak sesuai namun sudah berkurang.

c. SVM

Hasil akurasi dengan menggunakan algoritma SVM adalah 19,44%, dan untuk grafik evaluasi dapat dilihat pada gambar 6.

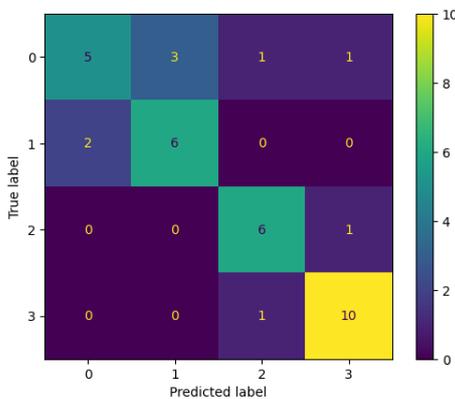


Gambar 6. Grafik hasil evaluasi Algoritma SVM

Dari grafik diatas dapat dilihat banyak data yang hasil prediksinya tidak sesuai.

d. Naïve Bayes (NB)

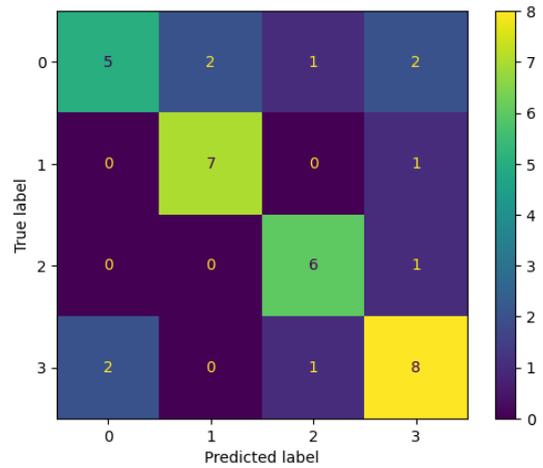
Hasil akurasi dengan menggunakan algoritma NB adalah 75% dengan grafik hasil evaluasi dapat dilihat pada gambar 7.



Gambar 7. Grafik hasil ebaluasi NB

e. Decision Tree Classifier(DTC)

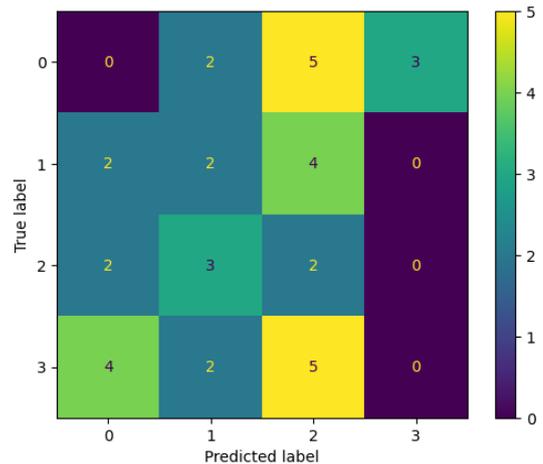
Hasil akurasi dengan menggunakan DTC adalah sebesar 75% dengan grafik evaluasi dapat dilihat pada gambar 8.



Gambar 8. Grafik Evaluasi Algoritma DTC

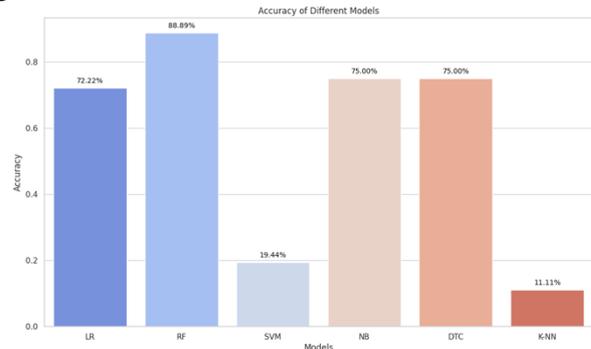
f. K-Neighbors Classifier (K-NN)

Hasil akurasi dengan menggunakan K-NN adalah 11,11%, dan grafik evaluasi dapat dilihat pada gamba 9.



Gambar 9. Grafik Evaluasi Algoritma K-NN

Untuk lebih jelas melihat model mana yang memiliki akurasi terbaik dapat dilihat pada gambar 10.

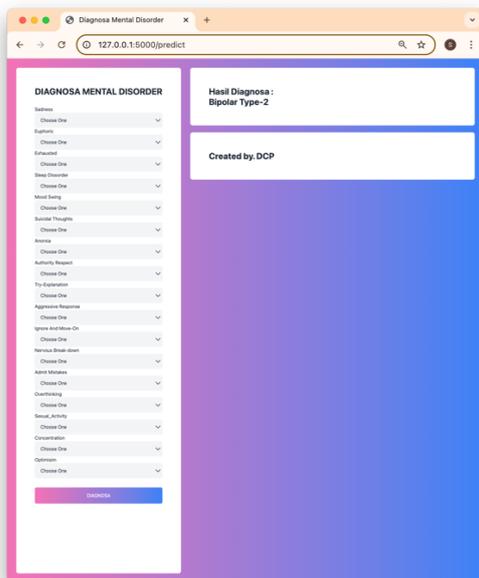


Gambar 10. Hasil Akurasi 6 Model yang digunakan

Dari gambar 10 dapat dilihat bahwa model LR dengan akurasi sebesar 72,22%, Model RF dengan akurasi 88,89%, model SVM dengan akurasi 19,44, model NB dengan akurasi 75,00%, model DTC dengan akurasi 75,00% dan KNN dengan akurasi 11,11%. Dari keenam model maka dapat di simpulkan bahwa Model dengan algoritma Random Forest yang memiliki nilai akurasi paling baik yaitu 88,89%. Tahapan selanjutnya adalah melakukan evaluasi dengan menggunakan confusion matrix, hasil dari confusion matrix dari keenam model dapat dilihat pada table 1 berikut:

Alg	TP	FP	FN	Akur	Pres	Rel	F1S
NB	27	9	9	0,75	0,75	0,74	0,75
RF	32	4	4	0,89	0,90	0,89	0,89
KNN	4	32	32	0,11	0,13	0,09	0,11
DTC	26	10	10	0,72	0,74	0,73	0,73
SVM	7	29	29	0,19	0,25	0,25	0,25
LR	26	10	10	0,72	0,74	0,78	0,76

Dari table 1 dapat dilihat bahwa model dengan akurasi terbaik adalah model Random Forest dengan akurasi sebesar 88,89%, nilai Precision sebesar 90%, Recall sebesar 89% dan F-1 Score sebesar 89%. Tahapan terakhir dari penelitian ini adalah deployment, deployment dirancang dengan menggunakan Visual Code, hasil dari program Klasifikasi Mental Disorder dapat dilihat pada gambar 3.



Gambar 5. Tampilan Aplikasi Diagnosa

Mental Disorder

Pembahasan penelitian berisi kesesuaian hasil penelitian dengan tujuan dan teori penelitian yang diuraikan sebelumnya. Pada bagian ini juga perlu dikemukakan perbandingan hasil penelitian dengan peneliti sebelumnya.

V. KESIMPULAN

Berdasarkan hasil yang diperoleh dari analisis model, dapat disimpulkan bahwa model Random Forest menunjukkan kinerja terbaik dengan akurasi sebesar 88,89%, Precision sebesar 90%, recall sebesar 89%, dan F1-score sebesar 89%. Penelitian ini mengungguli lima model lainnya, yaitu Logistic Regression, Naive Bayes, Decision Tree Classifier, Support Vector Machine, dan K-Nearest Neighbors. Model dengan kinerja terendah adalah K-Nearest Neighbors dengan akurasi hanya sebesar 11,11%, Precision sebesar 90%, recall sebesar 13%, dan F1-score sebesar 0,9%.

VI. SARAN

Diharapkan adanya pengembangan dari penelitian ini dengan melakukan eksperimen menggunakan algoritma klasifikasi lain agar hasil klasifikasi lebih akurat.

VII. DAFTAR PUSTAKA

- [1] M. Moitra *et al.*, "The global gap in treatment coverage for major depressive disorder in 84 countries from 2000–2019: A systematic review and Bayesian meta-regression analysis," *PLoS Med*, vol. 19, no. 2, Feb. 2022, doi: 10.1371/JOURNAL.PMED.1003901.
- [2] V. Patel *et al.*, "The Lancet Commission on global mental health and sustainable development," Oct. 27, 2018, *Lancet Publishing Group*. doi: 10.1016/S0140-6736(18)31612-X.
- [3] M. Bader, M. Abdelwanis, M. Maalouf, and H. F. Jelinek, "Detecting depression severity using weighted random forest and oxidative stress biomarkers," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-67251-y.
- [4] M. F. R. Aditya, N. L. Azizah, and U. Indahyanti, "Prediksi Penyakit Hipertensi Menggunakan Metode Decison Tree dan

- Random Forest,” *Jurnal Ilmiah Komputasi*, vol. 23, no. 1, Mar. 2024, doi: 10.32409/jikstik.23.1.3503.
- [5] M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim, and A. Fernández-Delgado, “Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?,” 2014. [Online]. Available: <http://www.mathworks.es/products/neural-network>.
- [6] U. M. Haque, E. Kabir, and R. Khanam, “Detection of child depression using machine learning methods,” *PLoS One*, vol. 16, no. 12 December 2021, Dec. 2021, doi: 10.1371/journal.pone.0261131.
- [7] A. Priya, S. Garg, and N. P. Tigga, “Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms,” in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 1258–1267. doi: 10.1016/j.procs.2020.03.442.
- [8] J. D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, Third Edition. John Wiley & Sons, Inc., 2023.
- [9] R. Supriyadi, W. Gata, N. Maulidah, A. Fauzi, I. Komputer, and S. Nusa Mandiri Jalan Margonda Raya No, “Penerapan Algoritma Random Forest Untuk Menentukan Kualitas Anggur Merah,” vol. 13, no. 2, pp. 67–75, 2020, [Online]. Available: <http://journal.stekom.ac.id/index.php/E-Bisnis/page67>
- [10] F. Abdusyukur, “Penerapan Algoritma Support Vector Machine (SVM) Untuk Klasifikasi Pencemaran Nama Baik Di Media Sosial Twitter,” *KOMPUTA : Jurnal Ilmiah Komputer dan Informatika*, vol. 12, no. 1, 2023.
- [11] R. Rachman, R. N. Handayani, and I. Artikel, “Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM,” *JURNAL INFORMATIKA*, vol. 8, no. 2, 2021, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [12] R. Latifah, E. Setia Wulandari, and dan Priadhana Edi Kreshna, “Model Decision Tree untuk Prediksi Jadwal Kerja menggunakan Scikit-Learn,” 2019.
- [13] M. S. Fajri, N. Septian, and E. Sanjaya, “Evaluasi Implementasi Algoritma Machine Learning K-Nearest Neighbors (kNN) pada Data Spektroskopi Gamma Resolusi Rendah,” *Al-Fiziya: Journal of Materials Science, Geophysics, Instrumentation and Theoretical Physics*, vol. 3, no. 1, pp. 9–14, Aug. 2020, doi: 10.15408/fiziya.v3i1.16180.
- [14] J. Han, M. Kamber, and J. Pei, “Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems),” 2011.

