



ISSN : 2339 - 1871

JURNAL ILMIAH BETRIK

Besemah Teknologi Informasi dan Komputer

Editor Office : LPPM Sekolah Tinggi Teknologi Pagar Alam, Jln. Masik Siagim No. 75
Simpang Mbacang, Pagar Alam, SUM-SEL, Indonesia
Phone : +62 852-7901-1390.
Email : betrik@sttpagaralam.ac.id | admin.jurnal@sttpagaralam.ac.id
Website : <https://ejournal.sttpagaralam.ac.id/index.php/betrik/index>

PERBANDINGAN ALGORITMA RANDOM FOREST DAN K-NEAREST NEIGHBOR DALAM KLASIFIKASI KESEHATAN MENTAL MAHASISWA

Farhan Pratama¹, Zikri Hadryan Nst², Zuriatul Khairi³,
Rahmaddeni⁴, Lusiana Efrizoni⁵

Program Studi Teknik Informatika Stmik Amik Riau¹²³⁴⁵

Jalan Purwodadi Indah No.KM, RW.10, Sidomulyo Barat, Tampan, Pekanbaru
City, Riau

Surel:Farhanpratamaa2444@gmail.com¹, Zikrihadryan12@gmail.com², Zuriatulkhairi241³,
rahmaddeni@sar.ac.id⁴, lusiana@stmik-amik-riau.ac.id⁵

Abstrak: Kesehatan mental telah menjadi masalah besar di dunia pendidikan, terutama di perguruan tinggi. Hal ini terlepas dari kenyataan bahwa kesehatan mental berkontribusi terhadap kemajuan akademik dan perkembangan siswa, dan bahkan kesehatan mental dapat mempengaruhi lingkungan di kampus. Kesehatan mental adalah ketika seseorang dapat menjalani kehidupan normal, bekerja dengan produktif, dan berkontribusi dengan orang-orang di sekitarnya. Tujuan dari penelitian ini adalah untuk klasifikasi Kesehatan mental Mahasiswa. Kecerdasan buatan, atau yang juga dikenal sebagai pembelajaran mesin, terdapat sejumlah algoritma. Algoritma pendekatan yang digunakan dalam penelitian ini adalah Random Forest dan KNN. Dataset yang digunakan berasal dari website Kaggle, Dataset berjumlah 1100 data. Hasil penelitian menunjukkan nilai akurasi rata-rata. Dengan menggunakan metode confusion matrix pada hutan acak, penelitian ini menemukan nilai akurasi 97% pada pembagian data 70:30 dan nilai akurasi 99% pada pembagian data 80:20. Dengan metode KNN, nilai akurasi 88% pada pembagian data 70:20 dan nilai akurasi 90% pada pembagian data 80:20.

Kunci Utama: Kesehatan mental, K-Nearest Neighbor, Random forest

Abstract: Mental health has become a major problem in the world of education, especially in colleges, despite the fact that mental health contributes to academic progress and student development, and even mental health can affect the environment on campus. Mental health is when one can lead a normal life, work productively, and interact with the people around them. The purpose of this study is to classify the mental health of students. Artificial intelligence, or also known as machine learning, has a number of algorithms. The algorithms used in this study are Random Forest and KNN. The datasets used are from Kaggle's website, which contains 1100 data. The results showed an average accuracy value. Using the metric confusion method in a random forest, this study found a 97% accuracy value on a 70:30 data division and a 99% accuracy value on an 80:20 data division. With the KNN method, an 88% accurate value on the 70:20 data segment and a 90% accurate value on 80:20 Data Division.

Keywords : Mental Health, K-Nearest Neighbor, Random forest

1. PENDAHULUAN

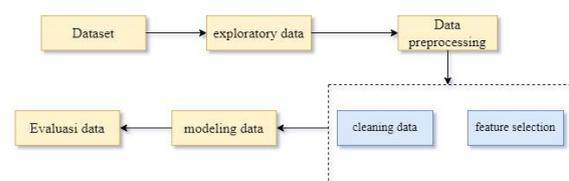
Pada tahun 1875 tepatnya di negara Jerman kajian dalam bidang ilmu kejiwaan atau kesehatan mental sudah menjadi topik penelitian terfavorit, salah satu pelopor nya ialah *Emil Kraepelin*, seorang *psikiater* Jerman yang banyak memberikan kontribusi terhadap perkembangan psikiatri modern pada abad ke 19 [1]. Dalam dunia pendidikan, terutama di perguruan tinggi, kesehatan mental telah menjadi masalah penting. Hal ini terlepas dari kenyataan bahwa kesehatan mental berkontribusi terhadap kemajuan akademik dan perkembangan siswa [2]. Bahkan Kesehatan mental dapat mempengaruhi lingkungan kampus. Kesehatan mental didefinisikan sebagai keadaan sejahtera di mana seseorang dapat menjalani kehidupan normal, bekerja secara produktif, dan berkontribusi dalam komunitasnya. Tidak adanya penyakit mental dan sejahtera psikologis ditandai dengan mental yang sehat. Penyakit mental terdiri dari gangguan kognisi, dan perilaku, sedangkan sejahtera psikologis terdiri dari unsur *afektif* dan *kognitif*, seperti stress[3]. Ketidakmampuan seseorang untuk berkembang secara fisik, intelektual, dan emosional dengan cara yang sama dan seimbang seperti perkembangan orang lain dikenal sebagai gangguan kesehatan mental [4]. Ini dapat menyebabkan depresi yang berlebihan. Perasaan sedih yang tidak henti-hentinya, kehilangan minat atau kesenangan dalam aktivitas sehari-hari, perubahan nafsu makan atau tidur, perasaan lelah yang berlebihan, penurunan energi, dan berbagai gejala fisik dan kognitif lainnya adalah gejala gangguan suasana hati yang dikenal sebagai depresi[5]. *WHO* menyatakan bahwa satu dari empat orang di dunia akan dipengaruhi oleh gangguan mental atau neurologis di beberapa titik dalam kehidupan mereka. Sekitar 450 juta orang saat ini menderita kondisi seperti itu, menempatkan gangguan mental di antara penyebab utama kesehatan buruk dan cacat di seluruh dunia. Salah satu contoh gangguan mental adalah depresi [6]. *WHO* menunjukkan bahwa depresi adalah penyakit umum di seluruh

dunia, dengan lebih dari 300 juta orang terkena dampaknya[7].

Kesehatan mental pada mahasiswa dapat dipengaruhi oleh berbagai faktor. Faktor-faktor tersebut antara lain faktor genetika, pertemanan, keluarga, lingkungan sosial, gaya hidup, dan lainnya. Orang yang mengalami masalah hidup yang ekstrem berisiko tinggi mengalami gangguan kesehatan mental [8]. Oleh karena itu penelitian ini bertujuan guna mengalih dan menentukan factor dan indicator yang dominan penyebab dari gangguan mental pada mahasiswa/i dengan menerapkan ilmu-ilmu pada kecerdasan buatan atau yang biasa disebut machine learning [9], dalam machine learning terdapat beberapa algoritma yang di unggulkan dalam memprediksi suatu kasus atau permasalahan yaitu ensemble learning yaitu pendekatan berbasis pohon keputusan yang menggabungkan beberapa algoritma klasifikasi, dan disini penulis juga akan menggabungkan beberapa algoritma lain yakni algoritma random forest dan K-Nearest Neighbors(KNN) [10]. Menurut jurnal[11]. Algoritma Random Forest dapat memprediksi akurasi mulai dari 93.33% nilai *sensitivity* dan menurut jurnal memaparkan beberapa kelebihan dari klasifikasi menggunakan metode K-NN seperti nilai akurasi yang tinggi, *insentive* terhadap *outlier*, dan tidak ada asumsi terhadap data. Pada penelitian tentang gangguan mental juga pernah diangkat oleh penelitian ini, dilakukan ujicoba dengan Pembagian 30% data uji dan 70% data latih, dengan hasil akurasi 72% dari dataset berjumlah 1254 data .

2. METODE PENELITIAN

Alur proses penelitian yang dilakukan dapat dilihat pada gambar 1.



Gambar 1 . Alur penelitian

A. Dataset

Dataset adalah kumpulan data yang terdiri dari sejumlah besar informasi yang terkait satu sama lain dan dapat diolah, dikelola, dan dianalisis. Dataset dapat mencakup berbagai jenis data, seperti angka, teks, gambar, suara, atau kombinasi dari berbagai jenis data tersebut [12]. *Beck Depression* dan *Beck Anxiety*. Gangguan kecemasan merupakan jenis gangguan mental yang umum terjadi. Mempengaruhi sekitar 1 dari empat orang dalam hidup mereka. Terdapat 30 variable serta 1100 data dari setiap variable nya.

B. Exploratory Data

Exploratory data (EDA) adalah suatu proses analisis data yang digunakan untuk memahami karakteristik, pola, dan hubungan dalam data secara visual dan deskriptif. Tujuan utama EDA adalah untuk membantu peneliti atau analis data memahami struktur data, menemukan pola yang menarik, dan mengumpulkan informasi yang dapat membantu mereka membuat keputusan.

Analisis data dapat menggunakan EDA untuk menciptakan pertanyaan penelitian atau menciptakan hipotesis untuk diuji lebih lanjut. Beberapa teknik yang umum digunakan dalam EDA termasuk visualisasi data menggunakan grafik, histogram, scatter plot, dan statistik deskriptif seperti mean, median, dan deviasi standar.

C. Data Preprocessing

Setelah Anda memiliki kumpulan data, lanjutkan dengan praproses data. Tujuan preprocessing data ini adalah untuk menghilangkan nilai yang hilang dan memilih fitur [13]. Setelah melakukan pemilihan fitur seleksi, pilih 10 variabel yang tidak dapat diandalkan untuk mencegah kegagalan. dan mengecualikan yang tidak dapat diandalkan untuk mencegah kegagalan.

Berikut ini tahapan preprocessing adalah :

a. cleaning data

Langkah selanjutnya dari penelitian ini adalah pengumpulan data. kumpulan data, penulis memiliki nilai yang hilang. data

harus dianalisis untuk mengidentifikasi informasi yang tidak *relevan*. tidak ada nilai yang hilang dalam dataset, setiap variabel di setiap baris data set memiliki nilai yang panjang .

b. Feature selection

fitur (variabel) yang tersedia dari kumpulan data untuk digunakan dalam pembuatan model dikenal dari pemilihan fitur. Tujuan utama dari pemilihan model memilih sesuai adalah untuk meningkatkan produktivitas model dengan menyesuaikan variabel yang tidak relevan model fit adalah meningkatkan produktivitas model dengan menyesuaikan variabel - variabel yang tidak relevan. Pada penelitian ini belajar, terdapat tiga puluh variabel yang akan digunakan untuk memilih fitur. ada tiga puluh variabel yang akan digunakan untuk memilih fitur. Setelah dilakukan seleksi fitur di temukan 9 variable yang relevan adalah sebagai berikut :

- a. Anxiety level
- b. Mental health story
- c. Depresi
- d. Sleep quality
- e. Academic performance
- f. Study load
- g. Teacher student relationship
- h. Peer pressure

D. Modeling data

Salah satu cara penulis menyelesaikan persoalan adalah dengan menggunakan Random forest. Dalam hal ini, Algoritma Random Forest (RF) adalah kumpulan pengklasifikasi yang mencakup berbagai jenis pohon. mempromosikan hutan acak dalam penelitian formal [8]. Metode Ini membantu menginterpretasikan hasil dan memahami elemen yang paling mempengaruhi kesehatan mental [14]. Tahapan untuk menyelesaikan random forest adalah sebagai berikut [15]:

- 1) Menentukan jumlah pohon k (dari fitur m) yang dipilih, di mana $k < m$.
- 2) Mengambil sampel acak sebanyak N dari dataset untuk setiap k .

- 3) Secara acak mengambil subset prediktor (p) untuk setiap k, di mana $m < p$.
- 4) Langkah kedua dan ketiga diulangi sebanyak k.
- 5) Hasil prediksi diperoleh dari suara terbanyak dari hasil klasifikasi sebanyak k.

Peneliti menggunakan metode Knn dan metode kedua random forest. Knn dapat memberikan prediksi berdasarkan data dari individu atau kelompok yang mirip dengan pengamatan yang diprediksi [9]. Oleh karena itu, penulis memilih pendekatan ini menghasilkan tingkat personalisasi dalam model, yang mungkin penting untuk memprediksi tingkat kesehatan mental individu, yang sangat berbeda. Langkah – langkah perhitungan pada KNN sebagai berikut :

- 1) Tentukan nilai k yang akan digunakan.
- 2) Kemudian, hitung hasil kedekatan data berdasarkan persamaan Euckidean distance terhadap data latin.

$$D(x, y) = ||x - y||_2 = \sqrt{\sum_j^N |x - y|^2}$$

- 3) Kemudian, urut hasil jarak dari urutan nilai tertinggi ke nilai terendah.
- 4) Akhir sekali, hitung jumlah setiap kelas berdasarkan K tertangga terdekat.
- 5) Selanjutnya, kelas mayoritas dibuat sebagai kelas baru untuk data uji.

E. Evaluasi model

Pada machine learning, evaluasi model adalah proses untuk mengevaluasi kinerja suatu model Klasifikasi dengan menggunakan *matrix* [16]. Evaluasi model sangat penting untuk menilai sejauh mana model tersebut mampu membuat Klasifikasi yang akurat. Dalam penelitian ini, metrix evaluasi diklasifikasikan menjadi laporannya

dalam lima bagian: akurasi, ketepatan, recall, skor F1, dukungan, dan konsistensi matriks.

Tabel *Confusion Matrix* memungkinkan penulis melihat kinerja algoritma pembelajaran yang diawasi. Setiap baris memiliki contoh class sebenarnya, tetapi setiap kolom matiks mewakili instance dalam kelas yang diharapkan [10]. Pengukuran dapat di hitung menggunakan rumus :

- 1) $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
- 2) $Recall = \frac{TP}{TP+FN}$
- 3) $Precision = \frac{TP}{TP+FP}$

3. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pada tahapan berikut ini data set diperoleh dari situ website kaggle dengan jumlah data aslinya berjumlah 1100 data yang berisikan 30 variable dan satu lable.

	basic_needs	academic_performance	study_load	teacher_student_relationship	future_career_concerns	social_support	peer_pressure	extracurricular_activities
2	3	2	3	2	3	2	3	3
2	1	4	1	5	1	4	5	5
2	2	3	3	2	2	3	2	2
2	2	4	1	4	1	4	4	4
3	4	3	1	2	1	5	0	0

Gambar 2 . Dataset

B. Preprocessing Data

Preprocessing data adalah berguna untuk merangkai langkah-langkah yang diambil untuk mengubah data mentah menjadi data yang relevan agar dapat diolah menjadi modeling data.

1) Cleaning data

Pada step ini di dataset kesehatan mental mahasiswa tidak terdapat missing value. Setelah memastikan tidak ada nya missing value maka step selanjutnya menentukan apakah data set ada yang duplikat. Pada dataset kesehatan mental mahasiswa terdapat 22 data yang bersifat duplikat oleh karena itu, disini penulis akan menghilangkan data yang bersifat duplikat dengan cara menginstall library pandas pada pyton.

2) Feature Selection

Pada tahapan ini penulis akan menggunakan variabel yang dianggap relevan untuk pembuatan model setelah

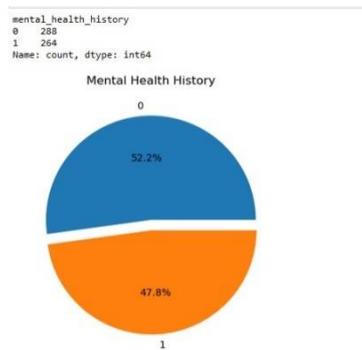
dialkukan analisis pada setiap variabel ditemukan hasil nilai mse :

MAE (percentage): 11.35%
 MSE (percentage): 16.31%
 RMSE (percentage): 28.56%

Oleh karena itu kami mengseleksi fitur. Sehingga di dapatkan 8 variable yang relavan.

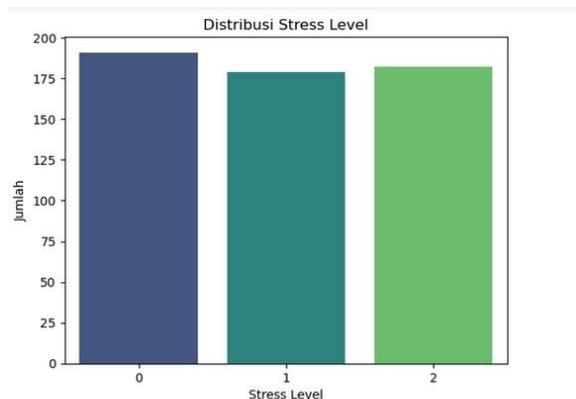
C. Exploratory data

Gambar 3. merupakan visualisasi data dari hasil Eksplorasi data yang telah dilakukan.



Gambar 3 . Diagram Riwayat penyakit mental

Gambar diatas lebih banyak mahasiswa yang yang tidak memliki mental heath histori sebanyak 52,2% mahasiswa. Dan yang memiliki mental health history sebanyak 47,8% mahasiswa.



Gambar 4 . Diagram Distribusi Stress level

Grafik diagram batang pada gambar tingkat kesehatan mahasiswa. Dari grafik di atas menunjukkan bahwa mahasiswa yang tidak memiliki tingkat kecemasan

yaitu berjumlah 191 mahasiswa, dan untuk yang memiliki tingkat kecemasan rendah yaitu berjumlah 182 mahasiswa, dan untuk yang memiliki tingkat kecemasan tinggi sekitar 179 mahasiswa.

D. Splitting Data

Pada penelitian melakukan 2 x splitting data yaitu 70:30 dan 80:20. Pada splitting data 80:20, untuk 80% digunakan sebagai data latih dan 20% digunakan sebagai data uji. Lalu pada splitting data 70:30, untuk 70% digunakan sebagai data latih dan 30% digunakan sebagai data uji.

E. Klasifikasi

1. Random forest

Table 1. merupakan hasil confusion matrix Algoritma menggunakan Random Forest.

Table 1. splitting data 70:30

	precision	recall	F1-score	support
0	0.93	1.00	0.96	52
1	1.00	0.93	0.96	58
2	1.00	1.00	1.00	56
Accuracy			0.97	166
Macro avg	0.98	0.98	0.98	166
Weightedavg	0.98	0.98	0.98	166

Berdasarkan Table 1. menggunakan metode confusion metrix maka didapatkan nilai akurasi 97% pada splitting data 70:30.

Table 2. splitting data 80:20

	precision	recall	F1-score	support
0	0.97	1.00	0.99	35
1	1.00	0.97	0.99	37
2	1.00	1.00	1.00	39
Accuracy			0.99	111
Macro avg	0.99	0.99	0.99	111
Weightedavg	0.99	0.99	0.99	111

Berdasarkan Table 2. menggunakan metode confusion metrix maka didapatkan nilai akurasi 99% pada splitting data 80:20.

2. KNN

pada klasifikasi menggunakan metode KNN menggunakan scikit-learn.

Table 3. splitting data 70:30

	precision	recall	F1-score	support
0	0.79	0.94	0.86	35
1	0.85	0.76	0.80	37
2	0.94	0.87	0.91	39
Accuracy			0.88	111
Macro avg	0.86	0.86	0.85	111
Weightedavg	0.86	0.86	0.86	111

Berdasarkan Table 3. menggunakan metode confusion matrix maka didapatkan nilai akurasi 88% pada splitting data 70:30.

Table 4. splitting data 80:20

	precision	recall	F1-score	support
0	0.83	0.83	0.83	52
1	0.80	0.88	0.84	58
2	0.98	0.88	0.92	56
Accuracy			0.90	166
Macro avg	0.87	0.86	0.86	166
Weightedavg	0.87	0.86	0.86	166

Berdasarkan Table 4. menggunakan metode confusion matrix maka didapatkan nilai akurasi 90% pada splitting data 80:20.

F. Evaluasi Model

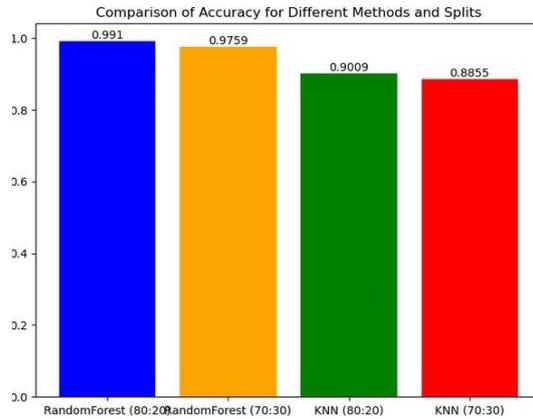
Pada penelitian menggunakan metode Random forest dan KNN dengan menggunakan splitting data 80:20 dan 70:30 dapat disimpulkan bahwasanya metode Random forest dapat menungguli metode KNN dalam melakukan tingkat accuracy model.

Splitting data	K-NEAREST NEIGHBOR				Random forest					
	precision	recall	F1-score	support	precision	recall	F1-score	support		
80:20	0	0.86	0.91	0.89	35	0	0.97	1.00	0.99	35
	1	0.85	0.82	0.86	37	1	1.00	0.97	0.99	37
	2	1.00	0.87	0.93	39	2	1.00	1.00	1.00	39
	Accuracy			0.90	111	Accuracy			0.99	111
	Macro avg	0.90	0.90	0.90	111	Macro avg	0.99	0.99	0.99	111
	Weighted avg	0.91	0.90	0.90	111	Weighted avg	0.99	0.99	0.99	111
70:30	0	0.86	0.92	0.89	52	0	0.93	1.00	0.96	52
	1	0.85	0.88	0.86	58	1	1.00	0.93	0.96	58
	2	0.96	0.86	0.91	56	2	1.00	1.00	1.00	56
	Accuracy			0.89	166	Accuracy			0.98	166
	Macro avg	0.89	0.89	0.89	166	Macro avg	0.98	0.98	0.98	166
	Weighted avg	0.89	0.89	0.89	166	Weighted avg	0.98	0.98	0.98	166

Gambar 5 . Hasil akurasi tingkat Kesehatan mental

Penelitian ini juga dapat menyimpulkan bahwasanya metode Random forest juga meungguli metode KNN dalam menentukan Precision, Recall, dan F1-score pada label 0 (Tidak berpotensi mengidap gangguan mental), 1 (Bisa jadi berpotensi

mengidap gangguan mental), dan 2 (mengidap gangguan mental).



Gambar 6 . Diagram hasil akurasi

4. SIMPULAN

Dari penelitian yang di lakukan pada dataset Tingkat kesehatan mental mahasiswa yang berjumlah 1100 data dengan menggunakan metode Random forest dan KNN. Random Forest menunjukkan kinerja yang luar biasa dalam eksperimen ini, dengan tingkat akurasi sebesar 99% pada spliting data 80:20 dan konsistensi tinggi sebesar 97% pada spliting data 70:30. Sebaliknya, K-Nearest Neighbor (KNN) menunjukkan hasil yang memadai, dengan tingkat akurasi sebesar 90% pada spliting data 80:20 dan 88% pada spliting data 70:30,.

Hal, ini menyatakan bahwa algoritma Random forest dapat membantu klasifikasi kesehatan mental mahasiswa dengan lebih baik. Agar penelitian ini lebih sempurna disarankan lebih banyak melakukan training dan testing data.

DAFTAR RUJUKAN

[1] D. F. Fakhriani, *Kesehatan Mental*, No. November 2019. 2019. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>
<http://digilib.uinsby.ac.id/918/10/daftar-pustaka.pdf>

[2] K. Yan, D. Arisandi, And T. Tony, "Analisis Sentimen Komentar Netizen Twitter Terhadap Kesehatan Mental Masyarakat Indonesia," *J. Ilmu Komput.*

- Dan Sist. Inf.*, Vol. 10, No. 1, 2022, Doi: 10.24912/Jiksi.V10i1.17865.
- [3] F. Fajrini *Et Al.*, “Environmental Occupational Health And Safety Journal Analisis Perubahan Iklim Dan Kesehatan Mental Pada Mahasiswa Fakultas Kesehatan Masyarakat Universitas Muhammadiyah Jakarta,” *Environ. Occup. Heal. Saf. J.* •, Vol. 3, No. 2, P. 95, 2023.
- [4] A. Madani, I. Prasetyowati, And C. A. Kinanthi, “Hubungan Karakteristik Mahasiswa Dengan Kesehatan Mental Mahasiswa Selama Kuliah Online,” *Ikesma*, Vol. 18, No. 2, P. 72, Jun. 2022, Doi: 10.19184/Ikesma.V18i1.25679.
- [5] H. Syahputra *Et Al.*, “Perbandingan Algoritma Support Vector Machine (Svm) Dan Decision Tree Untuk Deteksi Tingkat Depresi Mahasiswa,” Vol. 10, No. 1, Pp. 238–247, 2023.
- [6] R. Alfarezy, E. Ermatita, And R. M. B. Wadu, “Implementasi Algoritma Naïve Bayes Untuk Analisis Klasifikasi Survei Kesehatan Mental (Studi Kasus: Open Sourcing Mental Illness),” *Inform. J. Ilmu Komput.*, Vol. 19, No. 1, Pp. 1–10, 2023, Doi: 10.52958/Iftk.V19i1.4696.
- [7] A. Suryanto And S. Nada, “Analisis Kesehatan Mental Mahasiswa Perguruan Tinggi Pada Awal Terjangkitnya Covid-19 Di Indonesia,” *J. Citizsh. Virtues*, Vol. 1, No. 2, Pp. 83–97, 2021.
- [8] M. K. Sari And E. A. Susmiatin, “Deteksi Dini Kesehatan Mental Emosional Pada Mahasiswa,” *J. Ilm. Stikes Yars. Mataram*, Vol. 13, No. 1, Pp. 10–17, 2023, Doi: 10.57267/Jisym.V13i1.226.
- [9] F. A. Larasati, D. E. Ratnawati, And B. T. Hanggara, “Analisis Sentimen Ulasan Aplikasi Dana Dengan Metode Random Forest,” ... *Tekno. Inf. Dan ...*, Vol. 6, No. 9, Pp. 4305–4313, 2022, [Online]. Available: [Http://J-Ptiik.Ub.Ac.Id](http://J-Ptiik.Ub.Ac.Id)
- [10] D. Alita And A. R. Isnain, “Pendeteksian Sarkasme Pada Proses Analisis Sentimen Menggunakan Random Forest Classifier,” *J. Komputasi*, Vol. 8, No. 2, Pp. 50–58, 2020, Doi: 10.23960/Komputasi.V8i2.2615.
- [11] D. Septiani, U. Enri, And N. Sulistiyowati, “Diagnosa Tingkat Depresi Mahasiswa Selama Masa Pandemi Covid-19 Menggunakan Algoritma Random Forest,” *String (Satuan Tulisan Ris. Dan Inov. Tekno.*, Vol. 6, No. 2, P. 149, 2021, Doi: 10.30998/String.V6i2.10361.
- [12] 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.
- [13] M. Wibowo And M. R. F. Djafar, “Perbandingan Metode Klasifikasi Untuk Deteksi Stress Pada Mahasiswa Di Perguruan Tinggi,” *J. Media Inform. Budidarma*, Vol. 7, No. 1, Pp. 153–159, 2023, Doi: 10.30865/Mib.V7i1.5182.
- [14] E. R. B. Sebayang, Y. H. Chrisnanto, And Melina, “Klasifikasi Data Kesehatan Mental Di Industri Teknologi Menggunakan Algoritma Random Forest,” *Ijespg J.*, Vol. 1, No. 3, Pp. 237–253, 2023, [Online]. Available: [Http://Ijespgjournal.Org](http://Ijespgjournal.Org)
- [15] A. A. Karim, M. A. Prasetyo, And M. R. Saputro, “Perbandingan Metode Random Forest, K-Nearest Neighbor, Dan Svm Dalam Prediksi Akurasi Pertandingan Liga Italia,” *Stain. (Seminar Nas. Tekno. Sains)*, Vol. 2, No. 1, Pp. 377–382, 2023.
- [16] L. Madaerdo Sotarjua, D. Budhi Santoso, U. H. Singaperbangsa Karawang Jl Ronggo Waluyo, K. Telukjambe Timur, K. Karawang, And J. Barat, “Perbandingan Algoritma Knn, Decision Tree,*Dan Random*Forest Pada Data Imbalanced Class Untuk Klasifikasi Promosi Karyawan,” Vol. 7, No. 2, P. 2022.