

## DIAGNOSA TINGKAT DEPRESI MAHASISWA SELAMA MASA PANDEMI COVID-19 MENGGUNAKAN ALGORITMA RANDOM FOREST

Dewi Septiani<sup>1</sup>, Ultach Enri<sup>2</sup>, Nina Sulistiyowati<sup>3</sup>

Teknik Informatika, Universitas Singaperbangsa Karawang<sup>1,2,3</sup>

dewi.septiani17011@student.unsika.ac.id<sup>1</sup>, ultach@staff.unsika.ac.id<sup>2</sup>, nina.sulistio@unsika.ac.id<sup>3</sup>

*Submitted August 6, 2021; Revised November 25, 2021; Accepted November 26, 2021*

### Abstrak

Virus Covid-19 kini telah menjadi wabah penyakit diseluruh dunia, salah satu negara yang terdampak virus ini yaitu Negara Indonesia. Menurut data Satuan Petugas Penanganan Covid-19, jumlah penderita Covid-19 di Indonesia per tanggal 15 Februari 2021 mencapai 1,2 juta jiwa, Jumlah kasus yang terus bertambah setiap harinya membuat pemerintah memberlakukan kebijakan bekerja, belajar, dan beribadah dari rumah untuk meminimalkan penularan Covid-19. Adanya kebijakan tersebut dan banyaknya penderita Covid-19 di Indonesia mampu memengaruhi kesehatan mental masyarakat, tak terkecuali dengan mahasiswa Universitas Singaperbangsa Karawang. Berdasarkan hal tersebut, penelitian ini bertujuan untuk mendiagnosa awal tingkat depresi pada mahasiswa Universitas Singaperbangsa Karawang selama masa pandemi Covid-19 menggunakan data mining dengan algoritma *Random Forest*. Metodologi yang digunakan dalam penelitian ini yaitu KDD (*Knowledge Discovery in Database*) dan data yang digunakan berasal dari kuesioner PHQ-9 sebanyak 392 responden sesuai perhitungan rumus Slovin. Evaluasi model yang digunakan yaitu *10-fold cross validation*, dengan parameter *accuracy*, *sensitivity*, dan *specificity*. Hasil yang didapatkan dalam penelitian ini yaitu model prediksi tingkat depresi menggunakan algoritma *Random Forest* memiliki akurasi 85,94%, dimana dari 392 mahasiswa, 1,02% mahasiswa normal, 47,96% mengalami gejala depresi ringan, 36,73% depresi ringan, 8,16% depresi sedang, dan 6,12% depresi berat.

**Kata Kunci :** Data Mining, Depresi, Prediksi, PHQ-9, Random Forest

### Abstract

*Covid-19 virus has become a pandemic across the world, including Indonesia. Based on the data from the Covid-19 Handling Officer Unit, the number of Covid-19 sufferers in Indonesia until February 15, 2021 reaches 1.2 million people. The number of daily cases that continues to grow has forced the government to enforce policies to work, study, and worship from home to minimize the Covid-19 transmission. The policy and many Covid-19 sufferers Indonesia affect the mental health of people, including students of Singaperbangsa Karawang University. Therefore, this research aims to diagnose the initial level of depression in students of Singaperbangsa Karawang University during Covid-19 pandemic by using data mining with Random Forest algorithm. The method used in this research is KDD (Knowledge Discovery in Database) with data used come from PHQ-9 questionnaire given to 392 respondents according to calculation of Slovin formula. Evaluation model used is 10-fold cross validation, with accuracy, sensitivity and specificity parameters. The results of the research show the depression level prediction model using Random Forest algorithm has an accuracy of 85.94%. From 392 students, 1.02% of students are normal, 47.96% have mild depressive symptoms, 36.73% have mild depression, 8.16% have moderate depression, and 6.12% have major depression.*

**Keywords :** Data Mining, Depression, Prediction, PHQ-9, Random Forest

### 1. PENDAHULUAN

Virus baru bernama Covid-19 (*Coronavirus Disease 2019*) telah muncul di akhir tahun 2019. Virus ini merupakan

virus menular yang menyebabkan infeksi saluran pernafasan pada manusia [1]. Virus ini kini menjadi pandemi diseluruh dunia termasuk Negara Indonesia. Menurut data Satgas Penanganan Covid-19, jumlah

penderita Covid-19 di Indonesia per tanggal 15 Februari 2021 mencapai 1,2 juta jiwa, dan kasus aktif setiap harinya dari awal tahun hingga 15 Februari bertahan diatas 6000 jiwa, sehingga pemerintah Indonesia terpaksa memberlakukan kebijakan beraktivitas dari rumah seperti belajar, beribadah, dan bekerja secara daring untuk meminimalkan penularan Covid-19 [2].

Adanya kebijakan tersebut dan banyaknya penderita Covid-19 di Indonesia mampu memengaruhi kesehatan mental masyarakat, berdasarkan hasil swaperiksa *online* yang dilakukan oleh Perhimpunan Dokter Spesialis Kesehatan Jiwa Indonesia (PDSKJI) per tanggal 14 Mei 2020, dari 2364 responden 69% diantaranya mengalami masalah psikologis, dimana 67% diantaranya mengalami depresi dan provinsi Jawa Barat menjadi provinsi dengan masyarakat yang paling banyak mengalami masalah kesehatan mental [3].

Penelitian pada referensi [4] menghasilkan bahwa orang-orang yang berusia 18-29 lebih rentan terhadap kecemasan akibat Covid-19 dibandingkan kelompok usia lainnya. Mahasiswa termasuk kedalam kelompok usia tersebut sehingga kesehatan mental mahasiswa juga terpengaruh selama masa pandemi covid-19. Terdapat penelitian tingkat depresi pada mahasiswa selama masa pandemi pada referensi [5] dan diketahui bahwa gangguan *mood* ringan yang dialami mahasiswa sebanyak 25,7%, mahasiswa yang depresi berat sebanyak 12,2%, yang depresi rendah sebanyak 8,1%, yang depresi sedang sebanyak 0,7%, dan mahasiswa yang mengalami gangguan *mood* ekstrem sebanyak 0,7% [5]. Penelitian lain pada referensi [6], menghasilkan bahwa dari 147 mahasiswa sebanyak 21,1% mahasiswa mengalami depresi ringan, 17% depresi sedang dan 3,4% mengalami depresi berat. Depresi dapat menurunkan produktivitas mahasiswa pada kehidupan sehari-hari,

sehingga dibutuhkan diagnosa awal tingkat depresi untuk menghindari depresi kronis yang dapat memicu terjadinya bunuh diri ataupun penyakit mental lainnya.

Terdapat beberapa penelitian diagnosa tingkat depresi menggunakan algoritma data mining, pada referensi [7], data yang dikumpulkan menggunakan kuesioner PHQ-9 diolah dengan algoritma *Random Forest*, menghasilkan akurasi prediksi mencapai 93.33%. Pada referensi [8] yang menganalisis stres pada mahasiswa berdasarkan pengaruh tekanan ujian dan lama penggunaan Internet menggunakan algoritma *Linear Regression*, *Naïve Bayes*, *Random Forest*, dan SVM, menghasilkan nilai *sensitivity* algoritma *Random Forest* mencapai 100%, Sedangkan pada referensi [9] yang memprediksi penyakit mental pada pengangguran menggunakan 4 algoritma, menghasilkan algoritma *Random Forest* lebih baik dari pada *Gradient Boosted Trees* (GBT), KNN, dan *Decision Tree*.

Maka dari itu penelitian ini akan menggunakan *data mining* untuk melakukan diagnosa awal tingkat depresi pada mahasiswa pada salah satu perguruan tinggi negeri di Jawa Barat yaitu Universitas Singaperbangsa Karawang selama masa pandemi Covid-19 dengan algoritma *Random Forest* dan divalidasi menggunakan 10-fold cross validation dengan parameter evaluasi berupa *accuracy*, *sensitivity*, dan *specificity*. Hasil dari penelitian ini diharapkan dapat mengetahui hasil evaluasi model prediksi tingkat depresi menggunakan algoritma *random forest* dan mengetahui persentase tingkat depresi pada Mahasiswa Universitas Singaperbangsa Karawang selama masa pandemi covid-19.

## 2. METODE PENELITIAN

Metode pengumpulan data yang digunakan yaitu berupa kuesioner daring yang terdiri

dari 6 pertanyaan seputar biodata dan 9 pertanyaan PHQ-9 (*Patient Health Questionnaire-9*). Kuesioner PHQ-9 merupakan kuesioner yang paling sering digunakan untuk *skrining* awal depresi di fasilitas kesehatan primer dan memiliki akurasi yang lebih baik dibandingkan PHQ-2 dan WHO-5 [10],[11]. Berikut kuesioner PHQ-9 pada Tabel 1 [12].

**Tabel 1. Pertanyaan PHQ-9**

No.	Dalam 2 minggu terakhir, seberapa sering Anda terganggu oleh masalah-masalah berikut?	Tidak Pernah	Beberapa Hari	Lebih dari separuh waktu yang dimaksud	Hampir setiap hari
1	Kurang berminat atau bergairah dalam melakukan apapun.	0	1	2	3
2	Merasa murung, sedih, atau putus asa.	0	1	2	3
3	Sulit tidur / mudah terbangun, atau terlalu banyak tidur.	0	1	2	3
4	Merasa lelah atau kurang bertenaga.	0	1	2	3
5	Kurang nafsu makan atau terlalu banyak makan.	0	1	2	3
6	Kurang percaya diri, atau merasa bahwa Anda adalah orang yang gagal atau telah mengecewakan diri sendiri atau keluarga.	0	1	2	3
7	Sulit berkonsentrasi saat melakukan sesuatu, misalnya membaca koran atau menonton televisi.	0	1	2	3
8	Bergerak atau berbicara sangat lambat, sehingga orang lain memperhatikannya. Atau sebaliknya yaitu merasa resah atau gelisah sehingga Anda lebih sering bergerak dari biasanya.	0	1	2	3
9	Merasa lebih baik mati atau ingin melukai diri sendiri dengan cara apapun	0	1	2	3

Sumber : Alomedika, 2021

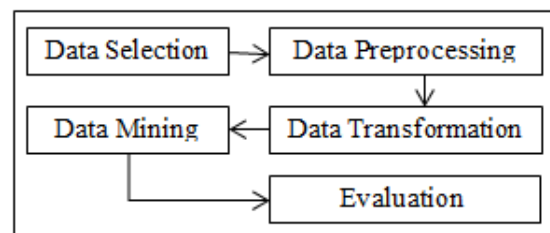
Tingkat depresi pasien ditentukan dari total skor jawaban seluruh pertanyaan. Berikut arti klinis dari masing-masing skor serta rekomendasi terapi yang dianjurkan [12]:

1. Skor 1 - 9. Gejala depresi ringan:  
Terapi yang disarankan yaitu perbanyak pengetahuan mengenai kesehatan mental dan psikologis.
2. Skor 10 – 14. Depresi ringan:  
Terapi yang disarankan yaitu mengobservasi gejala yang dialami

selama satu bulan, apabila tidak mengalami perbaikan, lakukan *konseling* psikologis singkat atau menggunakan *antidepresan*.

3. Skor 15 – 19. Depresi sedang:  
Terapi yang disarankan yaitu melakukan *psikoterapi* atau memberikan *antidepresan*.
4. Skor  $\geq 20$ . Depresi berat:  
Terapi yang disarankan yaitu melakukan *psikoterapi* secara *intensif* dan menggunakan *antidepresan*.

Metode yang digunakan dalam proses pembuatan model adalah KDD (*Knowledge Discovery in Database*). Metode ini memiliki proses yang terorganisir dalam memproses dataset besar dan kompleks hingga menghasilkan pola baru yang bermanfaat. Berikut 5 tahapan pada metode KDD.



**Gambar 1. Alur Proses KDD**

### Data Selection

Pada tahap ini, data yang telah dikumpulkan dari kuesioner diseleksi dan diambil data-data yang sesuai dengan kriteria yang diinginkan, yaitu mahasiswa aktif Universitas Singaperbangsa Karawang (UNSIKA). Setelah itu dilakukan penentuan sampel yang diambil menggunakan teknik *simple random sampling*, jumlah sampel yang digunakan dihitung menggunakan rumus *Slovin*, dengan populasi berupa jumlah mahasiswa UNSIKA tahun ajaran 2019/2020 semester ganjil yaitu 18766 mahasiswa [13]. Berikut rumus *Slovin* pada persamaan 1 [14].

$$n = \frac{N}{1+N(e)^2} \quad (1)$$

Keterangan:

- n = Sampel
- N = Populasi
- e = Tingkat Kesalahan

### Data Preprocessing

Pada tahap ini, dilakukan seleksi atribut untuk memilih atribut yang diperlukan yaitu 9 atribut PHQ-9.

### Data Transformation

Pada tahap ini, dilakukan transformasi tipe data nominal ke numerik di 9 atribut PHQ-9. Data diubah menjadi skor 0-3 sesuai ketentuan PHQ-9. Pengubahan tipe data ini dilakukan untuk perhitungan total skor tiap responden untuk menentukan tingkat depresinya. Tingkat depresi tersebut kemudian dijadikan label untuk proses data mining.

### Data Mining

*Data mining* merupakan teknik untuk menggali suatu informasi berharga dari suatu *database*. Proses *data mining* pada penelitian ini menggunakan aplikasi *RapidMiner Studio*, *dataset* yang telah ditransformasi tadi dibentuk model tingkat depresi menggunakan algoritma *Random Forest*. *Random Forest* (RF) merupakan Algoritma yang terdiri dari beberapa *tree* dimana di setiap *tree* dilakukan *training* terhadap sampel data [15]. *Criterion* yang digunakan pada penelitian ini yaitu *gain ratio* dan *voting strategy* yang digunakan untuk mendapatkan hasil prediksi yaitu *confidence vote*. Hasil data mining ini berupa prediksi tingkat depresi yaitu normal, gejala depresi ringan, depresi ringan, depresi sedang, dan depresi berat.

### Evaluation

Model yang telah dibuat kemudian divalidasi menggunakan *10-fold cross validation*, *10-fold cross validation* adalah metode *cross validation* yang membagi data sebanyak 10 bagian, dimana setiap bagiannya dilakukan pengujian, hasil dari

semua pengujian itu kemudian dirata-ratakan untuk mendapatkan nilai akhir. Hasil evaluasi penelitian ini berupa *accuracy*, *sensitivity*, dan *specificity* dari tiap tingkat depresi pada model. Hasil evaluasi tersebut didapatkan menggunakan *confusion matrix*, berikut tabel *confusion matrix* pada Tabel 2.

Tabel 2. Tabel Confusion Matrix

	True Positif	True Negatif
Prediction Positif	TP	FP
Prediction Negatif	FN	TN

Keterangan:

- TP : *True Positif*, data yang sebenarnya positif, diprediksi positif juga.
- TN : *True Negatif*, data yang sebenarnya negatif, diprediksi negatif juga.
- FN : *False Negatif*, data yang sebenarnya positif, namun diprediksi negatif.
- FP : *False Positif*, data yang sebenarnya negatif, namun diprediksi positif.

*Accuracy* adalah indikator untuk mengetahui keakuratan model yang dihasilkan, berikut rumus *accuracy* pada persamaan 2 [16].

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

*Sensitivity* merupakan persentase data yang diprediksi positif dari total data yang memang positif, berikut rumus *sensitivity* pada persamaan 3 [16].

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

*Specificity* merupakan persentase data yang diprediksi negatif dari total data yang memang negatif, berikut rumus *specificity* pada persamaan 4[16].

$$Specificity = \frac{TN}{FP+TN} \quad (4)$$

## 3. HASIL DAN PEMBAHASAN

Pengumpulan data dilakukan selama 26 hari dari tanggal 21 maret 2021 hingga 15 April 2021, menghasilkan data sebanyak 442 data responden, kemudian data

diproses menggunakan metode KDD sebagai berikut.

### Data Selecting

Pada proses ini, data 442 responden yang dihasilkan dilakukan seleksi data berupa penghilangan data ganda dan menghapus data selain mahasiswa Universitas Singaperbangsa Karawang dengan melihat kode universitas pada NPM, yaitu 1063. Berdasarkan hasil seleksi data yang dilakukan, maka 25 data dihapus karena 21 data merupakan data ganda dan 4 data tidak memenuhi kriteria yang diinginkan yaitu mahasiswa aktif UNSIKA, sehingga jumlah *dataset* yang tersisa sebanyak 417 responden. Kemudian dilakukan penentuan jumlah sampel menggunakan rumus *slovin* dengan taraf kesalahan sebesar 5% atau 0.05 sebagai berikut:

$$n = \frac{18766}{1 + 18766 \cdot 0.05^2}$$

$$n = \frac{18766}{1 + 46,915}$$

$$n = 391,65188354$$

Maka, dari 417 responden hasil seleksi sebelumnya, dipilih 392 responden secara acak untuk digunakan sebagai data sampel.

### Data Preprocessing

Data sampel yang digunakan terdiri dari 392 *record* dan 16 atribut, yaitu atribut Timestamp, *e-mail*, NPM, Jenis Kelamin, Umur, Angkatan, Fakultas, dan 9 pertanyaan PHQ-9. Pada *dataset* ini tidak terdapat *missing value*, sehingga langsung dilakukan seleksi atribut, dimana atribut yang dipilih hanya 9 pertanyaan PHQ-9, sedangkan atribut lainnya dihapus. Sehingga *dataset* saat ini terdiri dari 392 *record* dan 9 atribut.

### Data Transformation

Pada tahap ini dilakukan proses pengubahan tipe data *record* nominal ke numerik, *record-record* tersebut diubah

menjadi skor yang telah ditentukan oleh PHQ-9 seperti pada tabel 3 berikut:

**Tabel 3. Ketentuan Skor PHQ-9**

No	Jawaban	Skor
1.	Tidak Pernah	0
2.	Beberapa hari	1
3.	Lebih dari seminggu	2
4.	Hampir setiap hari dalam 2 minggu	3

Pada tahap ini juga dilakukan penambahan atribut “Tingkat Depresi” yang digunakan sebagai label dengan *record* berupa tingkat depresi setiap responden yang ditentukan berdasarkan total skor tiap responden. Pada Tabel 4 berikut menampilkan data sampel hasil transformasi.

**Tabel 4. Data Sampel Hasil Transformasi**

No	Tingkat Depresi	P1	P2	P3	P4	P5	P6	P7	P8	P9
1.	Depresi Ringan	1	2	3	3	1	1	1	0	1
2.	Depresi Ringan	1	1	1	3	0	1	2	1	0
...	...	...	...	...	...	...	...	...	...	...
391.	Depresi Sedang	1	1	2	3	3	2	1	2	3
392.	Gejala Depresi Ringan	1	0	1	0	0	0	0	0	0

Jumlah dari masing-masing tingkat depresi mahasiswa Universitas Singaperbangsa Karawang pada data sampel sangat bervariasi, pada Tabel 5 berikut menampilkan rincian jumlah masing-masing tingkat depresi mahasiswa sesuai ketentuan PHQ-9.

**Tabel 5. Penulisan Tabel**

No	Tingkat Depresi	Jumlah
1	Normal	4
2	Gejala Depresi Ringan	186
3	Depresi Ringan	124
4	Depresi Sedang	51
5	Depresi Berat	27

### Data Mining

Pada tahap ini, data sampel mahasiswa UNSIKA yang telah ditransformasi, diolah menggunakan algoritma *Random Forest*

dengan jumlah *tree* sebanyak 100 dengan subset prediktor sebesar 0.1, dimana data sampel ini dibagi menjadi 2, yaitu *data training* sebesar 0,9 dan data prediktor sebesar 0,1. *Data training* ini kemudian dibagi menjadi 100 data secara acak, lalu dibuat model *tree*-nya berdasarkan masing-masing data tersebut. 100 *tree* yang

terbentuk tersebut masing-masing memiliki hasil prediksi, untuk mendapatkan prediksi akhir, dilakukan *vote* berupa *confidence vote*, dimana hasil prediksi diambil berdasarkan nilai *confidence* yang paling tinggi. Pada Tabel 6 berikut menampilkan hasil prediksi dari proses data mining yang dilakukan.

**Tabel 6. Hasil Data Mining**

No	Prediction (Tingkat Depresi)	Confidence (Depresi Berat)	Confidence (Depresi Ringan)	Confidence (Depresi Sedang)	Confidence (Gejala Depresi Ringan)	Confidence (Normal)
1	Depresi Ringan	0.0	0.771	0.095	0.134	0.0
2	Depresi Ringan	0.003	0.455	0.179	0.363	0.0
...	...	...	...	...	...	...
391	Gejala Depresi Ringan	0.0	0.009	0.0	0.991	0.0
392	Depresi Sedang	0.051	0.295	0.591	0.063	0.0

### Evaluation

Model prediksi tingkat depresi yang telah dibuat divalidasi menggunakan 10-fold *cross validasi*, dimana data sampel dipecah menjadi 10 kelompok secara acak, kemudian diuji sebanyak 10 kali, hasil dari

masing-masing pengujian dirata-ratakan dan didapatkan hasil akurasi model prediksi sebesar 85,94% dengan rincian *class recall (sensitivity)* dan *class precision* tiap tingkat depresi pada Tabel 7

**Tabel 7. Hasil Evaluasi Model**

	True Depresi Ringan	True Gejala Depresi Ringan	True Depresi Berat	True Depresi Sedang	True Normal	Class Precision
Pred. Depresi Ringan	112	9	0	23	0	77.78%
Pred. Gejala Depresi Ringan	11	177	0	0	0	94.15%
Pred. Depresi Berat	0	0	20	4	0	83.33%
Pred. Depresi Sedang	1	0	7	24	0	75.00%
Pred. Normal	0	0	0	0	4	100.00%
Class Recall	90.32%	95.16%	74.07%	47.06%	100%	

Nilai *specificity* tiap tingkat depresi dihitung pada Tabel 8 berikut.

**Tabel 8. Perhitungan Specificity**

No.	Tingkat Depresi	Specificity = (TN) / (TN + FP)	Hasil	Persentase
1	Normal	388 / (388+0)	1	100%
2	Gejala Depresi Ringan	195 / (195+11)	0.9466	94,66%
3	Depresi Ringan	236 / (236+32)	0.8806	88,06%
4	Depresi Sedang	333 / (333+8)	0.9765	97,65%
5	Depresi Berat	361 / (361+4)	0.989	98,9%

Perhitungan nilai *accuracy* tiap tingkat depresi dilakukan pada Tabel 9 berikut.

**Tabel 9. Perhitungan Accuracy**

No.	Tingkat Depresi	Accuracy = (TP + TN) / (TP+FP+FN+TN)	Hasil	Persentase
1	Normal	$(4+388) / (4 + 0 + 0 + 388)$	1	100%
2	Gejala Depresi Ringan	$(177+195) / (177 + 11 + 9 + 195)$	0.9489	94,89%
3	Depresi Ringan	$(112+ 236) / (112 + 32 + 12 + 236)$	0.8877	88,77%
4	Depresi Sedang	$(24+333) / (24 + 8 + 27 + 333)$	0.9107	91,07%
5	Depresi Berat	$(20 +361) / (20 + 4 + 7 + 361)$	0.9719	97,19%

Berdasarkan hasil perhitungan *accuracy* dan *specificity* sebelumnya, maka pada Tabel 10 berikut memperlihatkan nilai *accuracy*, *specificity* dan *sensitivity* tiap tingkat depresi pada model prediksi tingkat depresi mahasiswa Universitas Singaperbangsa Karawang.

**Tabel 10. Nilai Accuracy, Specificity, dan Sensitivity Model Prediksi**

No.	Tingkat Depresi	Accuracy	Sensitivity	Specificity
1	Normal	100%	100%	100%
2	Gejala Depresi Ringan	94.89%	95,16%	94.66%
3	Depresi Ringan	88.77%	90,32%	88.06%
4	Depresi Sedang	91.07%	47,06%	97.65%
5	Depresi Berat	97.19%	74,07%	98.9%

Persentase *accuracy* dan *specificity* tiap tingkat depresi pada model sudah cukup baik, dimana semua tingkat memiliki persentase diatas 88%, bahkan tingkat normal mencapai 100%, hal ini dikarenakan data tingkat normal yang digunakan hanya 4 data dan prediksi yang muncul juga hanya 4 data, sehingga tidak terdapat *false negatif* dan *false positif*.

Pada nilai *sensitivity*, tingkat depresi normal, gejala depresi ringan dan depresi ringan memiliki persentase diatas 90%,

sedangkan depresi berat memiliki persentase sebesar 74,07% dan *sensitivity* terendah yaitu depresi sedang dengan persentase 47,06%, hal ini dikarenakan pada saat data mahasiswa diprediksi menggunakan model, model salah mendiagnosa tingkat depresi sedang sebagai tingkat depresi ringan dan depresi berat, sehingga mengakibatkan banyaknya data *false negatif*.

Berdasarkan hasil prediksi tingkat depresi mahasiswa UNSIKA menggunakan model, didapatkan hasil tingkat depresi dari 392 mahasiswa UNSIKA dijabarkan pada Tabel 11, yang berisi jumlah dan persentase mahasiswa Universitas Singaperbangsa Karawang pada tiap tingkat depresi.

**Tabel 11. Tingkat Depresi Mahasiswa UNSIKA**

No.	Tingkat Depresi	Jumlah Mahasiswa	Persentase
1	Normal	4	1,02%
2	Gejala Depresi Ringan	188	47,96%
3	Depresi Ringan	144	36,73%
4	Depresi Sedang	32	8,16%
5	Depresi Berat	24	6,12%

Pada tabel 11, dapat diketahui bahwa tingkat depresi terbanyak yang dialami mahasiswa Universitas Singaperbangsa Karawang selama masa pandemi Covid-19 yaitu gejala depresi ringan sebanyak 47,96% dan tingkat depresi terbanyak kedua yaitu depresi ringan sebanyak 36,73%, kemudian depresi sedang sebanyak 8,16%, depresi berat sebanyak 6,12%, dan mahasiswa yang normal sebanyak 1,02%.

Model prediksi tingkat depresi menggunakan algoritma *Random Forest* ini cukup efisien dalam melakukan *skrining* awal tingkat depresi pada mahasiswa Universitas Singaperbangsa Karawang, dimana *skrining* awal ini digunakan untuk mengetahui mahasiswa yang perlu pemeriksaan lebih lanjut dan

yang tidak perlu pemeriksaan lebih lanjut. Pemeriksaan lebih lanjut dianjurkan untuk mahasiswa yang didiagnosa tingkat depresi ringan, depresi sedang, dan depresi berat, maka dari 392 mahasiswa Universitas Singaperbangsa Karawang, 203 mahasiswa atau 51,79% mahasiswa dianjurkan untuk pemeriksaan lebih lanjut ke psikolog yang merupakan ahli/pakar dari kesehatan mental, sedangkan untuk 188 mahasiswa UNSIKA yang didiagnosa gejala depresi ringan, hanya perlu melakukan psikoedukasi apabila gejala memburuk.

#### 4. SIMPULAN

Model prediksi tingkat depresi menggunakan algoritma *Random Forest* memiliki akurasi sebesar 85,94%. Pada setiap tingkat depresi, memiliki persentase *accuracy* dan *specificity* diatas 88%, sedangkan untuk *sensitivity*, pada tingkat depresi normal, gejala depresi ringan, dan depresi ringan memiliki persentase diatas 90%, pada tingkat depresi berat sebesar 74%, dan pada tingkat depresi sedang sebesar 47%. Model prediksi tingkat depresi menggunakan algoritma *Random Forest* ini cukup efisien dalam melakukan skrining awal tingkat depresi pada mahasiswa Universitas Singaperbangsa Karawang, dimana dari 392 mahasiswa, tingkat depresi terbanyak yang dialami yaitu gejala depresi ringan sebanyak 47,96%, kemudian depresi ringan sebanyak 36,73%, depresi sedang sebanyak 8,16%, depresi berat sebanyak 6,12%, dan mahasiswa yang normal sebanyak 1,02%.

#### DAFTAR PUSTAKA

- [1] WHO, "QA For Publik," *World Health Organization*, 2020. <https://www.who.int/indonesia/news/novel-coronavirus/qa/qa-for-public> (accessed Feb. 25, 2021).
- [2] D. F. Nisa and N. K. Putri, "Bagaimana Wabah Virus Corona Mempengaruhi Kehidupan Mahasiswa? How Is The Coronavirus Outbreak Affecting The Daily Lives Of University?," *ResearchGate*, pp. 1–18, 2020, doi: 10.13140/RG.2.2.24215.19363.
- [3] PDSKJI, "Masalah Psikologis Di Era Pandemi Covid-19," *Perhimpunan Dokter Spesialis Kesehatan Jiwa Indonesia (PDSKJI)*, 2020. <http://www.pdskji.org/home>.
- [4] I. P. Jha, R. Awasthi, A. Kumar, V. Kumar, and T. Sethi, "Explainable-Machine-Learning To Discover Drivers And To Predict Mental Illness During Covid-19," *MedRxiv*, 2020, doi: 10.1101/2020.07.19.20157164.
- [5] A. Santoso *et al.*, "Tingkat Depresi Mahasiswa Keperawatan di Tengah Wabah Covid-19," *HNHS*, vol. 3, no. 1, pp. 1–8, 2020.
- [6] U. Hasanah, N. L. Fitri1, Supardi, and L. PH, "Depresi Pada Mahasiswa Selama Masa Pandemi Covid-19," *J. Keperawatan Jiwa*, vol. 8, no. 4, pp. 421–424, 2020.
- [7] S. Sawangarreerak and P. Thanathamthee, "Random Forest with Sampling Techniques for Handling Imbalanced Prediction of University Student Depression," *information*, pp. 1–13, 2020, doi: 10.3390/info11110519.
- [8] R. Ahuja and A. Banga, "Mental Stress Detection In University Students Using ML Algorithms," *Elsevier*, pp. 349–353, 2019, doi: 10.1016/j.procs.2019.05.007.
- [9] C. Gonçalves, D. Ferreira, C. Neto, A. Abelha, and J. Machado, "Prediction of Mental Illness Associated with unemployment using Data Mining," *Elsevier*, pp. 556–561, 2020, doi: 10.1016/j.procs.2020.10.078.



- [10] B. Levis, A. Benedetti, and B. D. Thombs, "Accuracy Of Patient Health Questionnaire-9 (PHQ-9) For Screening To Detect Major Depression: Individual Participant Data Meta-Analysis," *BMJ*, pp. 1–11, 2019, doi: 10.1136/bmj.11476.
- [11] C. P. Nolan *et al.*, "Depression Screening In HIV-Positive Tanzanian Adults: Comparing The PHQ-2, PHQ-9 And WHO-5 Questionnaires," *Glob. Ment. Heal.*, pp. 1–8, 2018, doi: 10.1017/gmh.2018.31.
- [12] D. Bestari, "Efektivitas Kuesioner PHQ-9 Sebagai Skrining Deteksi Dini Depresi," *Alomedika*, 2021. <https://www.alomedika.com/efektivitas-kuesioner-ph-9-sebagai-skrining-deteksi-dini-depresi>.
- [13] PDDikti, "Pelaporan Periode 2019/2020 Ganjil," *PDDikti Kemdikbud*, 2021. [https://pddikti.kemdikbud.go.id/data\\_pt/QzRENDE1QjMtOTgyRS00MDM0LUIyOUeEtRUU5NDA4MDY4MUEw](https://pddikti.kemdikbud.go.id/data_pt/QzRENDE1QjMtOTgyRS00MDM0LUIyOUeEtRUU5NDA4MDY4MUEw).
- [14] W. A. Fahrudin, "Balance Scorecard Untuk Menentukan Key Performance Indicator Di Pt Mulia Artha Anugerah," *JITMI*, vol. 3, no. 1, pp. 15–23, 2020.
- [15] K. A. Sambodo, M. I. Rahayu, N. Indriasari, and M. Natsir, "Klasifikasi Hutan-Non Hutan Data Alos Palsar Menggunakan Metode Random Forest," *Semin. Nas. Penginderaan Jauh 2014*, pp. 120–127, 2014.
- [16] T. Purwa, "Perbandingan Metode Regresi Logistik dan Random Forest untuk Klasifikasi Data Imbalanced (Studi Kasus: Klasifikasi Rumah Tangga Miskin di Kabupaten Karangasem, Bali Tahun 2017)," *JMSK*, vol. 16, no. 1, pp. 58–73, 2017, doi: 10.20956/jmsk.v16i1.6494.