

MULTI LABEL KLASIFIKASI UNTUK PREDIKSI DEPRESSION, ANXIETY, DAN BURNOUT MENGGUNAKAN RANDOM FOREST

Muhammad Firman Febrianto, Moh Muhtarom, Joni Maulindar

Teknik Informatika, Universitas Duta Bangsa
Jalan Bhayangkara No.55, Tipes, Kec. Serengan, Surakarta
firmanarwan048@gmail.com

ABSTRAK

Peningkatan aktivitas digital, tekanan akademik, serta pola hidup yang kurang seimbang menjadi faktor yang berkontribusi terhadap munculnya gangguan kesehatan mental seperti *depression*, *anxiety*, dan *burnout*. Ketiga kondisi tersebut dapat muncul secara bersamaan sehingga diperlukan pendekatan klasifikasi yang mampu melakukan prediksi multi kondisi dalam satu proses analisis. Penelitian ini bertujuan untuk membangun model prediksi kesehatan mental menggunakan pendekatan multi label *classification* dengan algoritma *Random Forest*. Dataset penelitian menggunakan 2.000 data dengan variabel aktivitas dan gaya hidup yang terdiri dari *stress level*, *daily screen time*, *social media usage*, *age*, dan *sleep hours*. Proses penelitian dilakukan menggunakan *framework* CRISP-DM melalui tahapan data *understanding*, *preprocessing*, *modeling*, dan *evaluation*. Pada tahap *preprocessing* dilakukan rekonstruksi label menggunakan *weighted risk score* dan penambahan kombinasi variabel berbasis pola perilaku untuk meningkatkan kemampuan diskriminasi model. Evaluasi model dilakukan menggunakan metrik *subset accuracy*, *hamming accuracy*, serta *F1-score* untuk mengukur performa klasifikasi multi label yang dihasilkan. Hasil pengujian menunjukkan bahwa model *Random Forest* memperoleh *subset accuracy* sebesar 85,75%, *hamming accuracy* sebesar 94,08%, serta *micro F1-score* dan *macro F1-score* sebesar 94,08%. Selain itu, akurasi pada masing-masing label menunjukkan nilai sebesar 93,25% untuk *depression*, 94,75% untuk *anxiety*, dan 94,25% untuk *burnout*. Hasil tersebut menunjukkan bahwa algoritma *Random Forest* mampu melakukan klasifikasi multi label terhadap kondisi kesehatan mental berdasarkan pola aktivitas harian individu, dengan performa label *Anxiety* (akurasi 94,75%, *F1-score* 94,71%), *Burnout* (akurasi 94,25%, *F1-score* 94,21%), dan *Depression* (akurasi 93,25%, *F1-score* 93,33%). Validasi silang 5-fold menghasilkan rata-rata akurasi 94,25%.

Kata kunci : Multi Label Classification, Random Forest, Depression, Anxiety, Burnout, Kesehatan Mental.

1. PENDAHULUAN

Kesehatan mental merupakan salah satu faktor penting yang memengaruhi kualitas hidup, produktivitas, serta kemampuan seseorang dalam menjalankan aktivitas sehari-hari. Dalam beberapa tahun terakhir kasus *depression*, *anxiety*, dan *burnout* [1] mengalami peningkatan yang signifikan, terutama pada kalangan pelajar dan pekerja produktif akibat tingginya tekanan aktivitas, penggunaan media digital yang berlebihan, serta pola hidup yang kurang seimbang [2].

Kondisi tersebut sering kali tidak terdeteksi sejak awal karena gejalanya muncul secara bertahap dan dapat terjadi secara bersamaan pada satu individu.

Permasalahan utama yang dihadapi dalam penelitian ini adalah bahwa sebagian besar pendekatan prediksi kesehatan mental yang ada saat ini masih menggunakan klasifikasi tunggal (*single label*), sehingga hanya mampu memprediksi satu kondisi mental dalam satu waktu. Padahal, kondisi seperti *depression*, *anxiety*, dan *burnout* sangat sering muncul secara bersamaan pada satu individu, membentuk pola *multidimensional* yang tidak dapat ditangkap secara menyeluruh oleh model *single label* [3].

Selain itu, dataset yang digunakan dalam penelitian ini memiliki korelasi yang sangat lemah antara variabel input dengan label target original, sehingga model tidak dapat belajar secara langsung

dari data mentah tanpa proses *preprocessing* yang tepat. Kondisi ini menjadi tantangan tersendiri yang perlu diselesaikan sebelum proses pemodelan dapat dilakukan secara efektif.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membangun model prediksi kesehatan mental menggunakan pendekatan *multi label classification* dengan algoritma *Random Forest* [4] yang mampu memprediksi tiga kondisi mental *depression*, *anxiety*, dan *burnout* secara bersamaan dalam satu proses analisis. Selain itu, penelitian ini juga bertujuan untuk menganalisis kontribusi setiap variabel input terhadap prediksi masing-masing kondisi mental melalui analisis *feature importance*, serta mengukur stabilitas dan kemampuan generalisasi model melalui validasi silang 5-fold [5].

Rencana penyelesaian masalah dalam penelitian ini dilakukan melalui beberapa tahapan yang mengacu pada *framework* CRISP-DM. Pertama, pada tahap *preprocessing* dilakukan rekonstruksi label menggunakan pendekatan *weighted risk score* berbasis *z-score* normalisasi untuk mengatasi korelasi lemah antara fitur input dengan label original, sehingga menghasilkan distribusi kelas yang seimbang (50:50) pada setiap label target. Kedua, ditambahkan tiga fitur interaksi baru yaitu *Stress_Sleep_Ratio*, *Stress_Screen*, dan *Screen_Social* [6] untuk menangkap hubungan antar variabel yang tidak

terjangkau oleh fitur tunggal. Ketiga, dataset dibagi dengan rasio 80:20 menggunakan *stratified splitting* berdasarkan kombinasi label, kemudian dimodelkan menggunakan *RandomForestClassifier* [7] yang telah melalui proses *hyperparameter tuning*. Keempat, model dievaluasi menggunakan metrik yang sesuai untuk konteks *multi label*, yaitu *subset accuracy*, *hamming accuracy*, *micro F1-score*, dan *macro F1-score*, serta divalidasi menggunakan cross-validation 5-fold untuk mengkonfirmasi stabilitas model. Dengan pendekatan ini, diharapkan model yang dihasilkan dapat menjadi dasar pengembangan sistem deteksi dini kesehatan mental berbasis machine learning yang lebih cepat, akurat, dan komprehensif.

2. TINJAUAN PUSTAKA

Penelitian mengenai prediksi dan klasifikasi kesehatan mental berbasis *machine learning* telah banyak dilakukan dengan berbagai pendekatan algoritma dan sumber data. Penelitian oleh [8] menggunakan algoritma Random Forest yang dikombinasikan dengan teknik bootstrap tree untuk mengklasifikasikan data survei kesehatan mental karyawan industri teknologi dari dataset OSMI (Open Source Mental Illness) yang diperoleh dari Kaggle, mencakup 1.259 data dengan 15 atribut hasil seleksi seperti Age, Gender, Family History, Work Interfere, Care Options, dan Wellness Program. Namun, penelitian ini memiliki keterbatasan karena pendekatannya yang bersifat single label sehingga tidak mampu menggambarkan keterkaitan antar kondisi mental yang dalam praktiknya sering muncul secara bersamaan, serta dataset yang digunakan sudah berumur lebih dari satu dekade sehingga relevansinya dengan kondisi mental.

Penelitian oleh [3] menerapkan Random Forest dengan pendekatan Knowledge Discovery in Database (KDD) untuk memprediksi kesehatan mental karyawan di sektor teknologi menggunakan dataset OSMI Mental Health in Tech Survey periode 2017 hingga 2022 yang terdiri dari 2.000 data dan 83 variabel. Salah satu temuan krusial dalam penelitian ini adalah kondisi ketidakseimbangan kelas yang sangat ekstrem, di mana kelas terdiagnosis hanya memiliki 2 sampel sedangkan kelas tidak terdiagnosis memiliki 398 sampel—kondisi yang menyebabkan akurasi baseline mencapai 99,5% namun precision dan recall untuk kelas minoritas bernilai 0%, menandakan model gagal total mendeteksi individu yang justru paling membutuhkan perhatian.

Penelitian oleh [2] melakukan klasifikasi tingkat stres dari data teks media sosial Twitter menggunakan dua algoritma, yaitu Support Vector Machine (SVM) dan Random Forest, yang masing-masing dikombinasikan dengan empat metode transformasi teks berbeda yaitu TF-IDF, CountVectorizer, NRCLEX, dan Word Affect Intensities. Pelatihan dilakukan menggunakan strategi K-Fold Cross Validation sebanyak 10 fold pada 6.751 data latih, dan pengujian pada 1.688 data uji. Hasil terbaik diperoleh

oleh SVM dengan transformasi TF-IDF yang menghasilkan akurasi 84%, precision 80%, recall 89%, dan F1-score 84%, sedangkan Random Forest terbaik menggunakan CountVectorizer dengan akurasi 80%.

Penelitian oleh [9] membandingkan performa tiga algoritma klasifikasi machine learning yaitu Naive Bayes, K-Nearest Neighbor (K-NN), dan Decision Tree untuk mengklasifikasikan kebutuhan penanganan kesehatan mental mahasiswa menggunakan dataset Kaggle. Tahap preprocessing mencakup imputasi 249 data hilang dengan metode modus, standarisasi variabel kategorikal, dan feature selection berbasis Information Gain yang mengeliminasi variabel dengan kontribusi rendah seperti timestamp.

Berdasarkan penelitian, terdapat beberapa masalah yang belum ditangani secara komprehensif. Pertama, seluruh penelitian masih menggunakan pendekatan single label classification yang hanya memprediksi satu kondisi kesehatan mental dalam satu model, sehingga tidak mampu menangkap keterkaitan antar kondisi depression, anxiety, dan burnout yang secara klinis sering muncul bersamaan pada satu individu. Kedua, Chanafy dan Sulistiani secara eksplisit menyarankan penggunaan algoritma ensemble seperti *Random Forest* sebagai arah penelitian lanjutan, namun belum mengimplementasikannya. Ketiga, tidak satu pun dari keempat penelitian tersebut mengeksplorasi pendekatan multi label classification yang memungkinkan prediksi simultan terhadap beberapa kondisi mental sekaligus.

2.1. Kesehatan Mental

Kesehatan mental adalah keadaan di mana seseorang merasa baik secara psikologis, yang memengaruhi cara mereka berpikir, merasakan, dan bertindak dalam kehidupan sehari-hari. Kesehatan mental seseorang dipengaruhi oleh kombinasi faktor biologis, lingkungan, sosial, dan cara hidup individu. Hal-hal seperti tekanan pekerjaan, kebiasaan tidur, tingkat stres, dan penggunaan media sosial dapat sangat memengaruhi kondisi mental seseorang [10] menegaskan bahwa tekanan kerja yang tinggi, tenggat waktu ketat, dan lingkungan kerja kompetitif dapat menyebabkan stres dan ketidakseimbangan mental yang berpotensi berkembang menjadi gangguan yang lebih serius apabila tidak terdeteksi sejak dini [11] menyebutkan bahwa kondisi ini mendorong kebutuhan metode prediksi otomatis yang mampu mengidentifikasi individu berisiko secara lebih cepat dan akurat.

2.2. Algoritma Random Forest

Random Forest merupakan algoritma *ensemble learning* yang bekerja dengan membangun sejumlah *decision tree* dan menggabungkan hasil prediksi dari masing-masing pohon untuk menghasilkan keputusan akhir yang lebih stabil. Algoritma ini memiliki kemampuan yang baik dalam menangani data numerik

maupun kategorikal serta mampu mengurangi risiko *overfitting*.

2.3. Multi Label Classification

Multi label classification [12] merupakan metode klasifikasi yang memungkinkan satu data memiliki lebih dari satu label target secara simultan. Berbeda dengan klasifikasi tradisional yang hanya menghasilkan satu kelas keluaran, pendekatan multi label lebih sesuai digunakan pada kasus yang memiliki keterkaitan antar target, termasuk pada prediksi kondisi kesehatan mental. Dalam penelitian ini, pendekatan multi label digunakan untuk memprediksi tiga kondisi mental sekaligus yaitu depression, anxiety, dan burnout menggunakan strategi Multi Output Classifier yang melatih satu Random Forest Classifier independen untuk setiap label target.

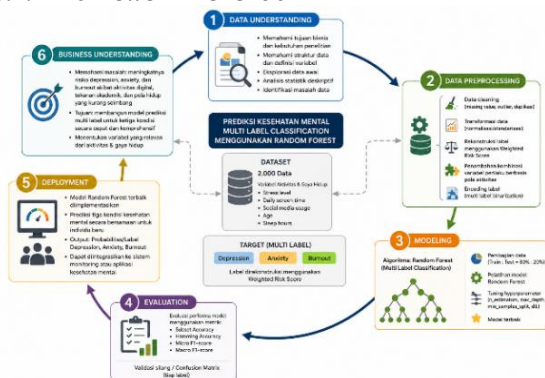
Evaluasi model multi label memerlukan metrik yang berbeda dari klasifikasi tunggal. Subset accuracy (exact match ratio) mengukur proporsi sampel yang seluruh kombinasi labelnya diprediksi benar secara bersamaan. Hamming accuracy (1 - Hamming Loss) mengukur akurasi rata-rata per label secara individual dan merupakan metrik yang lebih representatif untuk menilai performa keseluruhan model multi label. F1-score dihitung dalam dua varian: micro F1 yang menghitung seluruh label, dan macro F1 yang merata-ratakan F1 per label tanpa mempertimbangkan *imbalance*.

2.4. Framework CRISP-DM

CRISP-DM (Cross Industry Standard Process for Data Mining) merupakan framework yang terstruktur dan banyak digunakan sebagai panduan alur kerja dalam penelitian berbasis data mining dan machine learning. Framework ini mencakup tahapan yang komprehensif mulai dari pemahaman data (data *understanding*), persiapan data (data *preparation*) yang di dalamnya termasuk *preprocessing*, pemodelan (modeling), hingga evaluasi (*evaluation*). Penggunaan CRISP-DM memastikan proses penelitian berjalan secara sistematis, terstruktur, dan dapat direproduksi.

3. METODE PENELITIAN

3.1. Framework Penelitian



Gambar 1. Metodologi penelitian

Penelitian ini menggunakan pendekatan framework CRISP-DM sebagai panduan tahapan proses pengolahan data dan pembangunan model klasifikasi. Framework ini dipilih karena memiliki alur yang terstruktur dan banyak digunakan pada penelitian berbasis data mining maupun machine learning. Tahapan yang digunakan dalam penelitian meliputi data understanding, data preparation, modeling, dan evaluation.

Pada tahap data understanding, dilakukan analisis terhadap karakteristik dataset, distribusi label target, serta korelasi antar variabel. Tahap data preparation mencakup seleksi fitur, rekonstruksi label, penambahan fitur interaksi, dan pembagian dataset menjadi data latih dan data uji. Tahap modeling dilakukan menggunakan algoritma Random Forest dengan pendekatan multi label *classification*. Tahap *evaluation* digunakan untuk mengukur performa model berdasarkan hasil klasifikasi yang dihasilkan.

3.2. Data Collection

Dataset yang digunakan terdiri dari 2.000 data dengan 18 atribut yang berkaitan dengan profil demografis dan pola gaya hidup individu, diperoleh dari platform Kaggle. Berdasarkan hasil analisis awal, dataset tidak memiliki missing value pada seluruh atribut sehingga data dapat langsung diproses tanpa imputasi. Distribusi demografis menunjukkan komposisi yang seimbang antara perempuan (1.008 orang, 50,4%) dan laki-laki (992 orang, 49,6%), dengan sebaran pekerjaan meliputi mahasiswa/pelajar (677 orang), pengangguran (677 orang), dan pekerja (646 orang).

Tabel 1. Statistik Deskriptif Variabel Input

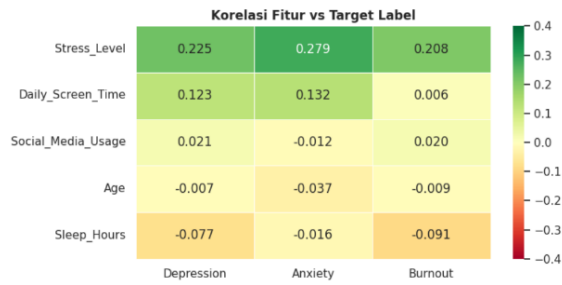
Variabel	Min	Maks	Average	Std.dev
Stress_Level	1,00	10,00	5,41	2,89
Daily_Screen_Time	2,00	12,00	6,94	2,88
Social_Media_Usage	0,00	8,00	3,38	2,24
Age	16,00	59,00	37,69	12,61
Sleep_Hours	3,00	9,00	6,01	1,71

Variabel input yang dipilih meliputi lima atribut yang memiliki relevansi terhadap kondisi kesehatan mental yaitu Stress_Level (skala 1-10), Daily_Screen_Time (jam/hari), Social_Media_Usage (jam/hari), Age (tahun), dan Sleep_Hours (jam/hari). Variabel target terdiri dari tiga kondisi kesehatan mental yaitu Depression, Anxiety, dan Burnout dalam format biner (0 = tidak, 1 = ya).

3.3. Analisis Korelasi

Analisis korelasi awal menunjukkan bahwa label original dataset memiliki korelasi yang sangat lemah dengan kelima fitur input (Stress_Level terhadap Depression: 0,225; terhadap Anxiety: 0,279; terhadap Burnout: 0,208; fitur lainnya di bawah 0,13). Korelasi yang lemah ini mengindikasikan bahwa model tidak dapat mempelajari pola yang bermakna secara langsung dari label original. Oleh karena itu, sebagai

bagian dari tahap preprocessing, label direkonstruksi menggunakan pendekatan *weighted risk score* berbasis pola perilaku dan faktor risiko.



Gambar 2. Korelasi Fitur

Setiap fitur terlebih dahulu dinormalisasi menggunakan *z-score* untuk menghilangkan perbedaan skala pengukuran. Selanjutnya, *risk score* untuk setiap kondisi mental dihitung menggunakan bobot yang mencerminkan kontribusi relatif setiap faktor:

- $Depression\ Score = 0,55 \times Stress_z - 0,35 \times Sleep_z + 0,10 \times Screen_z + \varepsilon$
- $Anxiety\ Score = 0,45 \times Stress_z + 0,30 \times Screen_z + 0,15 \times Social_z - 0,10 \times Sleep_z + \varepsilon$
- $Burnout\ Score = 0,50 \times Stress_z - 0,30 \times Sleep_z + 0,15 \times Screen_z + 0,05 \times Social_z + \varepsilon$

Di mana ε merupakan *noise Gaussian* ($\mu=0$, $\sigma=0,12$) yang ditambahkan untuk mencegah determinisme sempurna. Label kemudian dibinarisasi pada titik median setiap *risk score*, menghasilkan distribusi kelas yang seimbang (50:50) untuk masing-masing label sehingga model dapat belajar dari pola yang representatif tanpa bias terhadap kelas mayoritas.

3.4. Kombinasi Variabel

Selain lima 5 variabel, tiga kombinasi variabel ditambahkan dalam tahap preprocessing untuk menangkap hubungan antar variabel yang tidak terjangkau oleh fitur tunggal. Fitur interaksi ini dibangun berdasarkan pola perilaku yang relevan terhadap kondisi kesehatan mental:

- $Stress_Sleep_Ratio = Stress_Level / (Sleep_Hours + 1)$: merepresentasikan ketidakseimbangan antara beban stres dan kapasitas pemulihan melalui tidur.
- $7Stress_Screen = Stress_Level \times Daily_Screen_Time$: menangkap kombinasi beban mental dan paparan digital secara simultan.
- $Screen_Social = Daily_Screen_Time \times Social_Media_Usage$: mengukur intensitas paparan media sosial dalam konteks total penggunaan layar.

Sehingga total fitur yang digunakan dalam pemodelan berjumlah 8 fitur (5 fitur asli + 3 fitur interaksi).

3.5. Pembagian Dataset

Dataset dibagi menjadi data latih dan data uji menggunakan rasio 80:20 dengan *stratified splitting* berdasarkan kombinasi label target, sehingga distribusi multi label dipertahankan secara proporsional di kedua subset. Hasilnya adalah 1.600 data latih dan 400 data uji.

Training set : 1600 sampel (80%)
Test set : 400 sampel (20%)

Distribusi label di train:
Depression 0.5
Anxiety 0.5
Burnout 0.5
dtype: float64

Distribusi label di test:
Depression 0.5
Anxiety 0.5
Burnout 0.5
dtype: float64

Gambar 3. Train data

3.4. Pemodelan Random Forest

Proses pemodelan dilakukan menggunakan algoritma Random Forest melalui strategi *Multi Output Classifier* yang melatih satu *Random Forest Classifier independen* untuk setiap label target. Pada penelitian ini dilakukan proses hyperparameter tuning untuk memperoleh kombinasi parameter terbaik. Berdasarkan hasil pengujian, model terbaik diperoleh menggunakan konfigurasi berikut:

Tabel 2. Konfigurasi Hyperparameter Model

Hyperparameter	Nilai
n_estimators	300
max_features	sqrt
min_sample_leaf	1
min_sample_split	2
max_depth	None (tanpa batas)
random_state	42
strategi	Multioutputclassifier

3.5. Evaluasi Model

Evaluasi model dilakukan menggunakan empat metrik yang sesuai untuk konteks multi label classification:

- Subset Accuracy* mengukur proporsi sampel yang semua labelnya diprediksi benar secara bersamaan.
- Hamming Accuracy* mengukur akurasi rata-rata per label secara individual.
- Micro F1-Score* menghitung F1 secara agregat atas semua label.
- Macro F1-Score* menghitung rata-rata F1 per label. Selain itu dilakukan evaluasi per label menggunakan *classification report* dan *confusion matrix*, serta validasi silang 5-fold (*StratifiedKFold*) pada data latih untuk menilai stabilitas model.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Preprocessing

Tahap preprocessing menghasilkan dataset siap model dengan 8 fitur input dan 3 label target yang terdistribusi seimbang. Analisis korelasi sebelum preprocessing menunjukkan nilai korelasi tertinggi antara *Stress_Level* dengan *Anxiety* (0,279), *Depression* (0,225), dan *Burnout* (0,208), sementara fitur lain memiliki korelasi sangat rendah. Setelah rekonstruksi label menggunakan *weighted risk score*, distribusi kelas pada setiap target mencapai proporsi tepat 50:50, yang memastikan model dapat belajar dari pola kedua kelas secara seimbang tanpa diperlukan teknik oversampling tambahan seperti *SMOTE*.

Distribusi label pada data latih dan data uji menunjukkan proporsi yang seimbang dan merata, mengonfirmasi bahwa *stratified* splitting berhasil mempertahankan karakteristik distribusi multi label. Kondisi ini membantu model dalam melakukan proses pembelajaran secara lebih stabil tanpa mengalami ketimpangan kelas yang signifikan.

4.2. Hasil Evaluasi Model

Model Random Forest multi label yang dikembangkan diuji pada 400 data uji. Tabel 3 menyajikan ringkasan hasil evaluasi komprehensif menggunakan seluruh metrik yang ditetapkan.

Tabel 3. Hasil Evaluasi Model Multi Label

Metrik Evaluasi	Nilai
Subset Accuracy	85,75%
Hamming Accuracy	94,08%
Micro F1-Score	94,08%
Macro F1-Score	94,08%
Rata-rata Akurasi	94,08%

Subset accuracy sebesar 85,75% menunjukkan bahwa model berhasil memprediksi kombinasi lengkap ketiga label secara tepat untuk lebih dari 85% sampel uji—nilai yang sangat baik mengingat terdapat 8 kemungkinan kombinasi keluaran (2^3) dengan peluang exact match secara acak hanya 12,5%. *Hamming Accuracy* sebesar 94,08% mengindikasikan bahwa secara rata-rata, lebih dari 94 dari 100 prediksi label individual benar, mencerminkan kemampuan model yang sangat baik dalam mengklasifikasikan setiap kondisi mental secara terpisah.

Micro F1-Score dan Macro F1-Score yang keduanya bernilai 94,08% dan hampir identik mengindikasikan bahwa performa model seragam dan konsisten di ketiga label, tanpa ada bias signifikan terhadap label tertentu. Konsistensi ini merupakan dampak langsung dari distribusi kelas yang seimbang (50:50) yang dihasilkan pada tahap preprocessing.

4.3. Performa Per Label

Tabel 4 menyajikan performa model secara detail untuk masing-masing label target beserta nilai precision, recall, dan F1-score.

Tabel 4. Performa Model Per Label

Label	Akurasi	Precision	Recall	F1-Score
Depression	93,25%	93%	93%	93,33%
Anxiety	94,75%	95%	95%	94,71%
Burnout	94,25%	94%	94%	94,21%
Rata-rata	94,08%	94%	94%	94,08%

Anxiety menunjukkan akurasi tertinggi (94,75%) di antara ketiga label, diikuti Burnout (94,25%) dan Depression (93,25%). Perbedaan ini sejalan dengan bobot yang diberikan pada tahap rekonstruksi label, di mana Anxiety memanfaatkan kombinasi fitur yang lebih beragam (stress, screen time, social media, sleep) sehingga pola yang terbentuk lebih informatif dan lebih mudah dipelajari oleh model. Depression dengan F1-Score 93,33% tetap menunjukkan performa yang sangat baik meski berada di posisi ketiga. Secara keseluruhan, ketiga label menunjukkan performa yang sangat konsisten dan merata, tanpa ada label yang tertinggal secara signifikan.

4.4. Analisis Confusion Matrix Per Label

Analisis confusion matrix memberikan gambaran lebih rinci mengenai distribusi kesalahan prediksi model pada setiap label. Tabel 5 menyajikan confusion matrix untuk masing-masing label pada 400 data uji (200 sampel negatif dan 200 sampel positif per label).

Tabel 5. Confusion Matrix (data uji, n=400)

Label	TN	FP	FN	TP
Depression	184	16	11	189
Anxiety	191	9	12	188
Burnout	190	10	13	187

Pada label Depression, model menghasilkan 189 true positive dan 184 true negative, dengan hanya 11 false negative dan 16 false positive. False negative yang sangat rendah (11 dari 200 kasus positif) menunjukkan bahwa model sangat jarang melewatkan individu yang benar-benar mengalami depression—karakteristik yang sangat penting dalam sistem skrining kesehatan mental. Pada label Anxiety, false positive hanya 9 kasus yang merupakan angka terendah di antara ketiga label, menunjukkan presisi tertinggi dalam menandai individu berisiko. Pada label Burnout, false negative sebesar 13 kasus masih dalam batas yang sangat baik. Secara keseluruhan, proporsi kesalahan prediksi di bawah 3,3% untuk semua label menunjukkan performa model yang sangat memuaskan.

4.5. Feature Importance

Analisis feature importance mengungkap kontribusi relatif setiap fitur terhadap performa prediksi model. Tabel 6 menampilkan peringkat importance rata-rata dari tiga estimator yang mewakili tiga label target.

Tabel 6. Feature Importance (Rata-rata label)

Tingkat	Fitur	Importance	Jenis fitur
1	Stress_Sleep_ratio	37,43%	Atribut
2	Stress_level	20,08%	Asli
3	Stress_screen	18,61%	Atribut
4	Sleep_hours	7,47%	Asli
5	Daily_Screen_time	5,85%	Asli
6	Screen_social	5,19%	Atribut
7	Social_media_usage	3,16%	Asli
8	Age	2,22%	Asli

Stress_Sleep_Ratio mendominasi feature importance dengan nilai 37,43%, mengkonfirmasi bahwa rasio antara beban stres dan kapasitas pemulihan melalui tidur merupakan prediktor psikologis yang paling kuat. Fitur interaksi Stress_Screen berada di posisi ketiga (18,61%), menunjukkan relevansi kombinasi tekanan mental dengan paparan layar digital terhadap kondisi kesehatan mental. Secara kumulatif, ketiga fitur interaksi menyumbang 61,23% dari total feature importance, yang menjelaskan mengapa penambahan fitur interaksi pada tahap preprocessing mampu meningkatkan performa model secara signifikan dibandingkan jika hanya menggunakan lima fitur asli.

Stress_Level di posisi kedua (20,08%) mengonfirmasi perannya sebagai prediktor tunggal terpenting di antara fitur asli. Age berada di posisi terakhir dengan importance 2,22%, mengindikasikan bahwa dalam dataset ini pola aktivitas dan gaya hidup digital lebih menentukan risiko gangguan mental dibandingkan faktor demografis usia.

4.6. Validasi Silang 5-Fold

Validasi silang 5-fold dilakukan pada data latih untuk mengkonfirmasi stabilitas dan kemampuan generalisasi model. Gambar 4 menyajikan hasil akurasi pada setiap fold beserta rata-rata dan deviasi standarnya.

Cross-Validation 5-Fold per Label:
 Depression : 0.9537 ± 0.0142 [min=0.9344, max=0.9781]
 Anxiety : 0.9294 ± 0.0042 [min=0.9219, max=0.9344]
 Burnout : 0.9444 ± 0.0046 [min=0.9406, max=0.9531]

Rata-rata CV semua label: 0.9425 (94.25%)

Gambar 4. Cross Validation

Rata-rata akurasi cross-validation berkisar antara 92,94% (Anxiety) hingga 95,38% (Depression), dengan deviasi standar yang sangat kecil (di bawah 1,5% untuk semua label). Konsistensi antara performa pada data latih CV (rata-rata 94,25%) dan data uji (94,08%) mengkonfirmasi bahwa model tidak mengalami overfitting dan memiliki kemampuan generalisasi yang baik. Perbedaan yang sangat kecil antara akurasi CV dan akurasi data uji menunjukkan bahwa proses pembagian data yang dilakukan sudah representatif.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil membangun model prediksi kesehatan mental berbasis multi label *classification* menggunakan algoritma *Random Forest* yang mampu memprediksi tiga kondisi mental *depression*, *anxiety*, dan *burnout* secara bersamaan dari variabel pola aktivitas dan gaya hidup individu. Tahap preprocessing yang mencakup rekonstruksi label menggunakan *weighted risk score* dan penambahan tiga fitur interaksi (*Stress_Sleep_Ratio*, *Stress_Screen*, *Screen_Social*) terbukti menjadi kunci keberhasilan model dalam mempelajari pola yang bermakna dari data. Model mencapai *subset accuracy* 85,75%, *hamming accuracy* 94,08%, *micro F1-score* dan *macro F1-score* 94,08% pada data uji, dengan akurasi per label sebesar 93,25% untuk *depression*, 94,75% untuk *anxiety*, dan 94,25% untuk *burnout*. Hasil validasi silang 5-fold dengan rata-rata 94,25% dan deviasi standar rendah mengkonfirmasi stabilitas dan kemampuan generalisasi model yang baik. Fitur *Stress_Sleep_Ratio* terbukti sebagai prediktor paling dominan (37,43% *importance*), yang menunjukkan bahwa ketidakseimbangan antara beban stres dan kapasitas pemulihan melalui tidur merupakan faktor risiko kritis dalam memprediksi kondisi kesehatan mental. Oleh karena itu penelitian selanjutnya, disarankan menggunakan variabel input dengan faktor psikologis dan lingkungan yang lebih beragam, membandingkan performa dengan algoritma ensemble lainnya seperti *XGBoost* atau *Gradient Boosting*, serta mengembangkan antarmuka aplikasi yang mengintegrasikan model ini sebagai alat skrining kesehatan mental di institusi pendidikan maupun lingkungan kerja.

DAFTAR PUSTAKA

- [1] Oktafrian, M. V., Ramadhan, R., & Sari, A. P. (2026). Prediksi Risiko Burnout Karyawan Menggunakan K-Means. JATI: Jurnal Mahasiswa Teknik Informatika, 10(2), 2698–2703.
- [2] Mahing, N. F., Gunawan, A. L., Zen, A. F. A., Bachtiar, F. A., & Wicaksono, S. A. (2024). Klasifikasi Tingkat Stres dari Data Berbentuk Teks dengan Menggunakan Algoritma Support Vector Machine (SVM) dan Random Forest. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), Vol. 11, No. 5, hlm. 1067–1076. <https://doi.org/10.25126/jtiik2024118010>
- [3] Ririn, Rinaldi, A. R., & Basysyar, F. M. (2025). Prediksi Kesehatan Mental Menggunakan Random Forest Berdasarkan Faktor Demografi dan Lingkungan Kerja. JATI (Jurnal Mahasiswa Teknik Informatika), Vol. 9, No. 3, hlm. 4515–4522.
- [4] E. Rosta Br Sebayang, Y. Herry Chrisnanto, U. Jenderal Achmad Yani Cimahi, J. Terusan Jend Sudirman, C. Selatan, and J. Barat, “Klasifikasi Data Kesehatan Mental di Industri Teknologi

- Menggunakan Algoritma Random Forest,” 2023. [Online]. Available: <http://ijespjournal.org>
- [5] Rochmawati, D. R., Adib, M. A., Fazri, D. M., Raj, B., Antoni, R., Santoso, R., & Negoro, W. S. (2026). Classification and Interpretability of Employee Burnout Using Linear Discriminant Analysis. *Jurnal Komputer Teknologi Informasi Sistem Komputer (JUKTISI)*, 4(3), 2198–2206. <https://doi.org/10.62712/juktisi.v4i3.811>
- [6] Mujaky, A. M., Telaumbanua, B., Safitri, D., Ardiansyah, F. A., & Hafiz, M. I. (2026). Prediksi Tingkat Stres Mahasiswa Menggunakan Algoritma Decision Tree. *Journal of Information System and Application Development (JISAD)*, 4(1), 188–194. <https://doi.org/10.26905/jisad.v4i1.16616>
- [7] Fitri, R. R., Asriyanik, & Apriandari, W. (2025). Penggunaan Random Forest dalam Sistem Klasifikasi Kecemasan pada Generasi Z. *JITET: Jurnal Informatika dan Teknik Elektro Terapan*, 13(3), 561–571. <https://doi.org/10.23960/jitet.v13i3.6905>
- [8] Sebayang, E. R., Chrisnanto, Y. H., & Melina. (2023). Klasifikasi Data Kesehatan Mental di Industri Teknologi Menggunakan Algoritma Random Forest. *IJESPG (International Journal of Engineering, Economic, Social Politic and Government) Journal*, Vol. 1, No. 3, hlm. 237–253.
- [9] Maulana, I. S., Tiara, S. D., Pratama, A. Y., Sari, F. K., & Darmawan, I. (2025). Prediksi Risiko Depresi pada Mahasiswa Menggunakan Algoritma Random Forest Berdasarkan Data Akademik dan Gaya Hidup. *JSITIK: Jurnal Sistem Informasi dan Teknologi Informasi Komputer*, 4(1), 1–10. <https://doi.org/10.53624/jsitik.v4i1.696>
- [10] Mutmainah, S. (2022). Kemungkinan Depresi dari Postingan pada Media Sosial. *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, 1(2), 82–88. <https://doi.org/10.20885/snati.v1i2.11>
- [11] A. Y. Pratama, I. S. Maulana, F. K. Sari, S. D. Tiara, and I. Darmawan, “Prediksi Risiko Depresi pada Mahasiswa Menggunakan Algoritma Random Forest Berdasarkan Data Akademik dan Gaya Hidup,” *JSITIK: Jurnal Sistem Informasi dan Teknologi Informasi Komputer*, vol. 4, no. 1, pp. 1–10, Dec. 2025, doi: 10.53624/jsitik.v4i1.696.
- [12] Chanafy, M., & Sulistiani, H. (2026). Comparative Analysis of Machine Learning Algorithm Performance for Student Mental Health Classification. *SMATIKA: STIKI Informatika Jurnal*, 16(1), 19–29. <https://doi.org/10.32664/smatika.v16i01.2080>